IEEE
Control Systems
Society™

An IEEE Control Systems Initiative

CONTROL FOR SOCIETAL-SCALE CHALLENGES: ROAD MAP 2030

Anuradha M. Annaswamy
Karl H. Johansson | George J. Pappas

CONTROL FOR SOCIETAL-SCALE CHALLENGES: ROAD MAP 2030

EDITED BY:

Anuradha M. Annaswamy | Karl H. Johansson | George J. Pappas

PUBLISHED BY:

IEEE Control Systems Society

May 2023

SPONSORS:

IEEE Control Systems Society

International Federation of Automatic Control (IFAC)

National Science Foundation (NSF)

Digital Futures

This report is available online at
https://ieeecss.org/post/control-societal-scale-challenges-roadmap-2030

CONTROL FOR SOCIETAL-SCALE CHALLENGES: ROAD MAP 2030

EDITED BY:

Anuradha M. Annaswamy | Karl H. Johansson | George J. Pappas

AUTHORED BY:

Andrew Alleyne, Frank Allgöwer, Aaron D. Ames, Saurabh Amin, James Anderson, Anuradha M. Annaswamy, Panos J. Antsaklis, Neda Bagheri, Hamsa Balakrishnan, Bassam Bamieh, John Baras, Margret Bauer, Alexandre Bayen, Paul Bogdan, Steven L. Brunton, Francesco Bullo, Etienne Burdet, Joel Burdick, Laurent Burlion, Carlos Canudas de Wit, Ming Cao, Christos G. Cassandras, Aranya Chakrabortty, Giacomo Como, Marie Csete, Fabrizio Dabbene, Munther Dahleh, Amritam Das, Eyal Dassau, Claudio De Persis, Mario di Bernardo, Stefano Di Cairano, Dimos V. Dimarogonas, Florian Dörfler, John C. Doyle, Francis J. Doyle III, Anca Dragan, Magnus Egerstedt, Johan Eker, Sarah Fay, Dimitar Filev, Angela Fontan, Elisa Franco, Masayuki Fujita, Mario Garcia-Sanz, Dennice Gayme, Wilhelmus P.M.H. Heemels, João P. Hespanha, Sandra Hirche, Anette Hosoi, Jonathan P. How, Gabriela Hug, Marija Ilić, Hideaki Ishii, Ali Jadbabaie, Matin Jafarian, Samuel Qing-Shan Jia, Tor Arne Johansen, Karl H. Johansson, Dalton Jones, Mustafa Khammash, Pramod Khargonekar, Mykel J. Kochenderfer, Andreas Krause, Anthony Kuh, Dana Kulić, Françoise Lamnabhi-Lagarrigue, Naomi E. Leonard, Frederick Leve, Na Li, Steven Low, John Lygeros, Iven Mareels, Sonia Martinez, Nikolai Matni, Tommaso Menara, Katja Mombaur, Kevin Moore, Richard Murray, Toru Namerikawa, Angelia Nedich, Sandeep Neema, Mariana Netto, Timothy O'Leary, Marcia K. O'Malley, Lucy Y. Pao, Antonis Papachristodoulou, George J. Pappas, Philip E. Paré, Thomas Parisini, Fabio Pasqualetti, Marco Pavone, Akshay Rajhans, Gireeja Ranade, Anders Rantzer, Lillian Ratliff, J. Anthony Rossiter, Dorsa Sadigh, Tariq Samad, Henrik Sandberg, Sri Sarma, Luca Schenato, Jacquelien Scherpen, Angela Schoellig, Rodolphe Sepulchre, Jeff Shamma, Robert Shorten, Bruno Sinopoli, Koushil Sreenath, Jakob Stoustrup, Jing Sun, Paulo Tabuada, Emma Tegling, Dawn Tilbury, Claire J. Tomlin, Jana Tumova, Kevin Wise, Dan Work, Junaid Zafar, Melanie Zeilinger

Preface

The world faces some of its greatest challenges of modern time and how we address them will have a dramatic impact on the life for generations to come. Control systems, with its advanced toolbox for understanding and designing feedback, robustness, dynamics, decision-making, and complex systems in general for a wide set of industrial and societal contexts, has an opportunity to play a central role in the development of technologies and solutions for many of the key problems. Many control researchers are already engaged in such activities, and some are even driving the rapid technological evolution we are currently experiencing. We believe however that much more can be done. Awareness needs to be raised about societal drivers, technological trends, and emerging methodologies that are most essential for control systems to gain critical mass and true scientific and global impact. And we need to outline how control education and technology translation should evolve in the new landscape and for the years to come, considering also issues related to limited natural resources, ethics, regulations, etc.

The initiative to develop this road map on control for societal-scale challenges was taken by us more than two years ago. Inspired by encouraging colleagues in the community and by past efforts, we engaged a group of leading researchers to help organize the community around this activity. We have had many meetings and workshops to discuss and prepare this document. It should be emphasized that the entire road map exercise has been a true community effort. In the acknowledgement section in the end of the document, the large number of contributors to the road map and the various events are listed. We received significant and focused contributions from various leaders in the community as well as input from the community at large. This insight was gathered through open invitations to articulate grand visions and broad roles for the control systems community to participate in.

While we believe the road map will be of a strong and broad interest for many individuals and organizations, we have prepared it with two principal audiences in mind: (1) Young researchers who are in a nascent stage in their career. We want to draw their attention to open and important problems that can lead to significant breakthroughs. (2) Funding agencies. We want to underscore the need for support to develop methodologies that can lead to groundbreaking results and can be transitioned to societal-scale systems.

Finally, we would like to express our deepest appreciation for the commitment and dedication from all our colleagues who contributed to the road map, the activities that led up to it, and reviewed various versions of it. We are grateful to the leadership of IEEE Control Systems Society for launching this initiative and supporting it from the beginning. Without the significant financial funding from the IEEE Control Systems Society, IFAC, US National Science Foundation, and Digital Futures, this project simply would not have been possible to execute. We would also like to thank Amritam Das, Angela Fontan, and Linnea Sundling for their help with many editorial and organizational tasks.

We hope this road map will provide guidelines for participation in the overall scientific endeavor for developing methods for the betterment of humanity.

Anuradha M. Annaswamy
Karl H. Johansson
George J. Pappas

Executive Summary

The field of control systems applies specific principles and methods to control dynamic systems so that they produce desired outcomes. The scope of this field has been gradually enlarging as the world embraces a digital way of life. Information enriched by various degrees of analytics to inform decision-making is moving beyond engineering into financial services, socioeconomic analysis, entertainment and sports, and political and social sciences. Increased levels of automation are sought after in various sectors and being introduced into new domains. All of these advances and transformations urge a shift in the conversation toward how control systems can meet grand societal-scale challenges. This is the focus of this Control for Societal-Scale Challenges: Road Map 2030.

The document seeks to chart a road map for the evolution of control systems, identifying social areas where our discipline can have an impact over the next decade. We begin with a discussion of major societal drivers, the role of control systems therein, and emerging technological trends that enable the implementation of control systems. We delineate a few methodological directions in control systems that have recently emerged, as well as needs and pathways for ensuring validation of emerging methods and technology transition. We also provide insight into new workforce education and training curricula to address and implement the solutions and methods identified.

The societal drivers we address span climate change, healthcare for quality of life, smart infrastructure systems, the economy of sharing, and resilience in societal-scale systems. While these societal drivers discussed in Chapter 2 introduce challenges, technological trends outlined in Chapter 3 provide opportunities for control systems to address some of these challenges and outline new ones in AI and big data, electrification of everything, engineering biology, and robotics in the real world. The challenges and opportunities in Chapters 2 and 3 lead to the investigation of new methods and concepts in control systems, some of which are outlined in Chapter 4. These include learning and data-driven approaches, methods for safety-critical systems, resilient cyber-physical systems (CPS), cyber-physical-human systems (CPHS), and novel control architectures. In each case, near-term and far-term challenges are outlined.

Chapter 5 articulates technology validation with a focus on control implementation and demonstration of impact. This chapter also points out the importance of engaging industry in the overall conversation of advanced control technologies and their tangible benefits in various applications that could benefit society. Recognizing that education is the cornerstone for the growth and prosperity of the control systems field, Chapter 6 discusses university curriculum changes. It provides specific suggestions on what concepts and methodologies should be emphasized and how to adapt to a new generation of students. Chapter 7 emphasizes that the control systems community has an important role to play in future technology designs that respect human rights and human values, ensure ethics and fairness, and meet regulatory guidelines, even while safeguarding our environment and our natural resources. Finally, Chapter 8 summarizes the recommendations made in previous chapters. We end with Epilogue, a section that describes the genesis and evolution of this roadmap, Acknowledgments that list all contributors and sponsors, a Glossary of keywords, and an Index.

Table of Contents

Foreword

As president, I'm delighted to express my full support for this flagship initiative led by the IEEE Control Systems Society. The scope of IEEE is *Advancing Technology for Humanity*, and the Control Systems Society is dedicated to advancing the theory and practice of systems and control.

2030 is the year chosen for the 2030 Agenda for Sustainable Development by the United Nations. It is therefore most appropriate for the field of control systems to lead the way in solving societal-scale challenges over the next decade. In this document, we propose novel scientific challenges that the community should pursue, investigate workforce education and training curricula to address these challenges, and suggest strategies for translation to innovation and transition to practice leading to future disruptive technologies.

It is important to acknowledge the editors of this milestone document: Anuradha M. Annaswamy, Karl H. Johansson, and George J. Pappas. Their leadership, vision, passion, and hard work made this project a reality. Anu, Kalle, and George coordinated an inclusive effort by a large number of leading members of our community. Thank you all! I also acknowledge the generous support from the International Federation of Automatic Control, the National Science Foundation in the U.S., and Digital Futures in Sweden. This road map is the culmination of a collaborative effort by the systems and control community. I am sure that control systems will play a pivotal role in enabling new technologies that will benefit society and will no longer be perceived as "the hidden technology."

Thomas Parisini
2022 President of the IEEE Control Systems Society

CHAPTER 1

Introduction

Introduction

Anuradha M. Annaswamy, Karl H. Johansson, George J. Pappas

The field of control systems pertains to specific methods and principles to control dynamic systems and produce desired outcomes despite uncertainties in the system and in the environment. The development of these principles and methods has been both broad and deep, from aerospace systems and wireless networks to bioengineering and traffic control. These methods are deeply embedded in many sectors, including energy, transportation, healthcare, manufacturing, and robotics.

The scope of control applications has continuously increased over the past few decades. It has moved from feedback control of a single device or system to optimizing and making decisions in large-scale systems, system-of-systems, and infrastructure systems. This impressive growth of scope and scale is enabled by large paradigm shifts in science and technology across the global landscape. Digital transformation has been pervasive, enabling significant advances in mature engineering systems and introducing new concepts and computational constructs in non-engineering systems such as economic and social sectors. This increased cyber footprint, and related exponential advances in computational, communication, and actuation technologies, have introduced the concept of automation into new domains and in many cases accelerated its implementation. Applications such as self-driving vehicles and automatic control of aerial robots have entered the social lexicon and captured the imagination of the general population. Automation does not only underpin these applications but is being introduced into a host of other domains as well. All of these advances and transformations urge a shift in the conversation toward how control systems can solve grand societal-scale challenges. This is the focus of this **Control for Societal-Scale Challenges: Road Map 2030**.

This road map guides the reader on a journey into the future of control systems, illustrating new societal areas where our discipline can have an impact over the next decade. It identifies major societal drivers and emerging technological trends, proposes novel scientific challenges that the community should pursue, delineates needs and pathways for ensuring validation of emerging methods and technology transition, and investigates workforce education and training curricula to address and implement the solutions and methods identified.

Similar efforts with a large, ambitious scope have been carried out in the control community in the past [1, 2, 3, 4]. While [1] focused primarily on scientific research challenges, efforts in [2] and [3] concentrated on both success stories and research challenges across a range of applications. The reference [4] described several critical societal challenges in various sectors, including transportation, energy, water, healthcare, and manufacturing.

In 2019, discussions began in the CSS Executive Committee for a need for a document delivering a broad picture of how control systems can effectively impact societal challenges. This led to organizing of two workshops, one in June 2021 and another in June 2022, both of which focused on *Control for Societal-Scale Challenges*. The workshop in June 2021 was held virtually, over two days. On both days, the first half of the day was an open session attended by about 300 people worldwide. The second half of each day was dedicated to creating the road map. The open sessions consisted of six panels organized around a range of topics, including safety-critical autonomous systems, resilient infrastructure systems with AI and IoT, decision-making with real-time and distributed data, control with human-in-the-loop, control for climate change mitigation and adaptation, and education and training. About six panelists participated in each panel, with presentations focused on the topic at hand, the role of control, the underlying challenges, barriers to entry, and the potential societal challenges that this topic will impact. The topics were expanded on and organized around suitable headings, setting the stage for creating the road map. Gaps and overlaps were identified with steps taken to allow better exposition and more comprehensive coverage of the major topics.

Over the following 12 months, regular conversations were organized to brainstorm the overall structure of the road map. This brainstorming included appropriate classification of topics to suitably capture the major societal challenges to be presented in the road map. An overall structure emerged, with about six to seven chapters, spanning societal drivers, technological trends, key methodologies, technology transition, and education and training. The road map shows how main control methodologies can be brought to bear on the societal drivers of global challenges and the emerging technologies that can help address and potentially mitigate these challenges.

In addition to discussing drivers, challenges, and methodologies, we also discuss the transition from control research to products and solutions in various application domains. Considerations and pathways for technology transfer, understanding barriers and roadblocks in various application domains, and cost-benefit of automations and other control solutions are articulated.

Readers will note that the title of this document specifies the year 2030. As this target is only eight years away from the time of writing, the time horizon for the methods outlined in this chapter is not that far away. We have chosen 2030 to emphasize that many of the glide paths that we need to embark on in order to meet the global challenges should be created to meet this time frame. An additional point to note is that in many of these cases, whether they concern realizing a high penetration of renewables, aggressive reduction of greenhouse gas emissions, or phasing out of internal-combustion engine-based automobiles, the year 2030 figures prominently. We have therefore targeted this year as the focus for advancing and transitioning the requisite control systems technologies.

A few details follow regarding the structure of this road map. In Chapter 2, we examine a few major societal challenges that face the world in the 21st century. The first of those challenges is climate change, a problem of increasing urgency that we are continuously exposed to. Healthcare issues that affect quality of life are next addressed. Infrastructure systems that provide fundamental services such as energy, water, and transportation are discussed, especially the impact of digitalization and empowered consumers. As we proceed deeper into the 21st century and resources become scarce, new concepts of resource sharing (including the Sharing Economy) are shaping society, and these are also discussed. Finally, concerns about global security and risks that affect the socioeconomic landscape are delineated.

In Chapter 3, we outline emerging technological trends that support a large fraction of the efforts of the scientific community at large. These include biological engineering, robotics in the real world, large-scale electrification, and the roles of artificial intelligence and big data. Each one of these trends presents a range of opportunities for the control systems community to play a role in addressing and mitigating the societal challenges presented in Chapter 2.

Meeting the challenges and opportunities presented in Chapters 2 and 3 require new methodologies in control systems, some of which are outlined in Chapter 4. These include learning and data-driven approaches, methods for safety-critical systems, methods for resilience, analysis and synthesis of cyber-physical-human systems (CPHS), and novel control architectures. In each case, near-term and far-term challenges associated with each of these methods are outlined.

Chapter 5 articulates the need for a validation infrastructure to illustrate the performance of new control methods and demonstrate their impact. This chapter also points out the importance of engaging industry and the public sector in the overall conversation of advanced control technologies and their tangible benefits in various applications. The important problem of bridging the gap between industrial needs and economic and financial expectations on the one hand and relevant and significant research advances on the other is addressed.

Recognizing that education is the cornerstone for the growth and prosperity of the field of control systems, Chapter 6 discusses university curriculum changes. It provides specific suggestions on what concepts and methodologies should be emphasized and how to adapt to a new generation of students. Finally, in Chapter 7, we emphasize that the control systems community has an important role to play in future technology designs that respect human rights and human values, ensure ethics and fairness, and meet regulatory guidelines while safeguarding our environment and our natural resources.

A glossary of keywords and an index are included at the end of the document. Also included as appendices are details associated with some of the societal drivers. Visuals are provided throughout the document to highlight key ideas and underlying concepts. References are listed at the end of each chapter for further reading.

We have prepared this document with two principal audiences in mind: **(1)** Young researchers who are in a nascent stage in their career. We would like to draw their attention to the societal needs and technological trends that drive the development of the field of control systems, emerging methodologies, open problems, and key challenges in applying control systems to address emerging needs. **(2)** Funding agencies. Given the emerging scientific challenges and opportunities for control systems to impact those challenges, we want this road map to lead to discussions with funding agencies and other sponsors about new research initiatives that critically rely on advances in control systems. We want to underscore support needed to develop various methodologies and validate and transition them to societal-scale systems.

We have made every effort to include all major imperatives that will shape the research directions that the control systems community engages in. An effort is also made to outline a few other directions that our community ought to engage in, such as ethics, fairness, socially responsible automation, and intersections with regulatory agencies and policy makers.

Finally, we note that this entire road map exercise has been a community effort, with significant and focused contributions from various leaders in the community as well as inputs from the community at large. This information was gathered through open invitations to articulate grand visions and broad roles for the control systems community to participate in. We hope that this road map will provide guidelines for participation in the overall scientific endeavor for developing methods for the betterment of humanity.

BIBLIOGRAPHY

[1] K. J. Åström, S. P. Boyd, R. W. Brockett, G. Stein, and R. M. Murray, *Control in an Information Rich World: Report of the Panel on Future Directions in Control, Dynamics, and Systems*, R. M. Murray, Ed. Society for Industrial and Applied Mathematics, 2003. https://epubs.siam.org/doi/abs/10.1137/1.9780898718010

[2] T. Samad and A. M. Annaswamy, Eds., *The Impact of Control Technology*, 1st ed. IEEE Control Systems Society, 2011.

[3] T. Samad and A. M. Annaswamy, Eds., *The Impact of Control Technology*, 2nd ed. IEEE Control Systems Society, 2014.

[4] F. Lamnabhi-Lagarrigue *et al.*, “Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges,” *Annual Reviews in Control*, vol. 43, pp. 1–64, 2017. https://www.sciencedirect.com/science/article/pii/S1367578817300573

CHAPTER 2

Societal Drivers

Societal Drivers

Society faces a number of major global challenges. More attention needs to be paid by the research community to analyze and assess these global challenges and develop scientific tools and engineering solutions to mitigate their effects.

This chapter highlights how the control systems community can provide tools and technologies for five societal drivers. The chapter starts with a section on climate change mitigation and adaptation. It highlights the existential threat climate change poses to humanity and the crucial role the control community plays in developing methods and technologies to reduce greenhouse gas emissions and their negative consequences. The second section presents challenges in healthcare and quality of life, including equitable access to high-quality healthcare. Several opportunities connected to new medical technologies and capabilities are discussed. The third section covers smart infrastructure systems. It focuses on how the digital transformation of transportation, water, energy, food, and other infrastructures enables new capabilities and opportunities for control due to advances in communication, sensing, and data analysis. The disruptive paradigm of the Sharing Economy is covered in the fourth section. Here it is argued that new notions of ownership and market interactions would benefit from new control-theoretic concepts.

The chapter concludes with a fifth section on the resilience of societal-scale systems, discussing the need for better planning, management, and control of global systems such as supply chains and networks. This section also highlights interdependencies between the societal drivers in this chapter, such as infrastructure failures due to climate change and extreme events, which lead to complex vulnerabilities.

2.A Climate Change Mitigation and Adaptation

Pramod Khargonekar, Tariq Samad, Saurabh Amin, Aranya Chakrabortty, Fabrizio Dabbene, Amritam Das, Masayuki Fujita, Mario Garcia-Sanz, Dennice Gayme, Gabriela Hug, Marija Ilić, Iven Mareels, Kevin Moore, Lucy Y. Pao, Akshay Rajhans, Jakob Stoustrup, Junaid Zafar, Margret Bauer

> This section is focused on the issues of climate change mitigation, adaptation, and resilience. A comprehensive and diverse collection of research opportunities for the control systems community is discussed along with considerations that provide a broader framework for this research.

Abstract In this section, we explore the enormous societal-scale challenges driven by climate change and global warming and discuss possible research opportunities for the control systems community. We begin with a synopsis of the current understanding of climate change processes and their impacts, including the major challenges pertaining to replacement of fossil fuels and decarbonization of energy systems as well as agriculture and land-use management. We also recognize that the impact of global warming is already being felt. This motivates efforts on climate change adaptation and resilience. We present research opportunities in the following topics: Infrastructures and Communities; Electric Energy Systems; Electric Power Generation; Transportation, Homes, Buildings, and Facilities; Industry and Manufacturing; Hydrogen, Ammonia, and Renewable Fuels; Food and Agriculture; Water; Artificial Intelligence Computations; Negative Emissions Technologies; Geoengineering; and Environmental Monitoring. In each of these topic areas, we present three concrete research directions where control systems can play a significant role, although the opportunities are even more numerous. We then discuss several contextual issues that provide a broader frame for pursuit of these research opportunities. We discuss the considerations for sustainable economic growth. Because of the urgency and time-sensitive nature of climate change, we discuss the need for going from research to large-scale deployment and how that might be accomplished. We discuss the role of education and awareness among students and the broader communities. We underscore the need to consider the paramount issues related to equity and justice (economic, regional, global, and generational). We then posit that the control systems community will need to engage in transdisciplinary collaborations with experts from other fields of science, engineering, health, social sciences, law, humanities, and arts for meaningful research that has a desired, positive impact.

2.A.1 Introduction

Climate change poses an existential threat to humanity. It is now indisputable that the primary cause of this threat is human activity resulting in high greenhouse gas emissions, which began during the Industrial Revolution and has continued to rapidly accelerate. The first warnings of impending and irreversible climate change were sounded decades ago, when governmental and intergovernmental policy makers had sufficient time to enact the changes needed to avoid the dire situation we find ourselves in today. It appears that global warming (of more than 1.5°C) is all but inevitable—and its adverse impacts are already being felt all around the world.

The global community must now focus on actions that will save the planetary ecosystem from the most dramatic potential consequences. Scientists are increasingly shifting away from the single-minded goal of avoiding climate change and toward the twin goals of **(a)** reducing additional global warming beyond the 1.5°C threshold and **(b)** mitigating and adapting to the effects of climate change.

With the emergence and convergence of powerful new technologies in the last few decades, the control systems community is well-positioned to play a crucial role in the worldwide effort to tackle the challenges posed by anthropogenic climate change.

On the information side, these promising technologies include advanced communications, the industrial internet of things, artificial intelligence, machine learning, data analytics, and smart computing capabilities embedded in edge devices. In collaboration with experts in these and other fields, the control systems community can make vital contributions to climate resilience, including through the integration of key climate variable into the planning and operations of critical infrastructure systems; the leveraging of feedback to manage multiscale processes in resource-constrained settings; and the investigation of cross-sector interdependencies in societal systems.

Success will also depend on making major methodological changes—in particular moving beyond the traditional insularity of controls research. For better and faster solutions, control design must be integrated with overall system design rather than being relegated to later stages of development. And we cannot neglect the human element—users and stakeholders need to be engaged through appropriate mechanism design and behavioral incentives.

In this chapter, we will first review current data and projections on climate change and its anthropogenic basis, noting in particular the recent assessment by the Intergovernmental Panel on Climate Change (IPCC). Next, we will discuss several distinct targets for research and innovation for the control community, including infrastructure systems, power generation and grids, industrial processes, transportation and logistics, and food and agriculture. The urgency of the climate crisis means that research-as-usual approaches must be eschewed. To that end, we discuss several important considerations related to sustainability, social justice, deployment-at-scale, education, and transdisciplinary collaborations. General sources we have relied on for this chapter include (IPCC, 2021, [1]) and (IEA, 2021, [2]). Unless otherwise noted, emissions information is from (Our World in Data, 2020 [3]; see Figure 2.1). Additional references are included where appropriate.

2.A.2 Climate Change: the 2022 View

Earth's climate system is undergoing rapid changes not seen in at least the last 2,000 years. There is evidence that global average temperatures now are higher than the warmest period in the last 100,000 years, which was 6,500 years ago. The fundamental physical principles that govern the climate system are well established. It is important to note that the climate system is an increasingly coupled natural-human system. Human behavior plays a large and essential role in fundamental climate physics. Earth system models for climate are being improved and used to predict (along with uncertainty in the predictions) its future course.

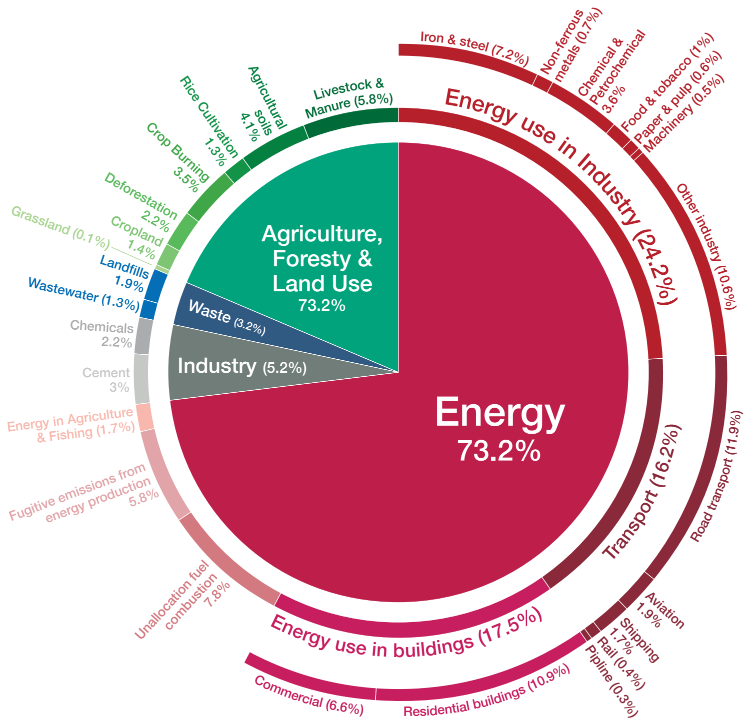


Figure 2.1: Global greenhouse gas emissions by sector [3]; the data is for 2016.

The main causes of human activity-induced global warming include increases in CO_2, CH_4, N_2O, volatile organic compounds and CO, and black carbon. There is strong evidence that in 2019, atmospheric CO_2 concentrations were higher than at any time in at least 2 million years, and concentrations of CH_4 and N_2O were higher than at any time in at least 800,000 years. It is estimated that CO_2 concentration has increased from preindustrial levels of 280 ppm to 412 ppm in 2020.

The concept of climate sensitivity is defined as the expected increase in global average temperature in response to the doubling of atmospheric CO_2 from preindustrial levels. The recent Sixth Assessment report by IPCC ([1]) concludes that the (equilibrium) climate sensitivity is in the range of "2°C (high confidence) and 5°C (medium confidence)." This provides a possible range of global warming as future greenhouse gas (GHG) concentrations increase.

The Sixth Assessment report provides very strong evidence that the global average surface temperature in the 2011-2020 decade was 1.09°C (confidence interval 0.95 – 1.2) higher than the 1850-1900 baseline. In fact, each of the last four decades has been successively warmer than any decade that preceded it since 1850. The report leverages extensive research on natural drivers of climate change such as changes in incoming solar radiation, volcanic activity, and global biogeochemical cycles, among other drivers. It also shares observational data on earth's climate system variables to find unequivocal evidence that the human activity-induced increases in greenhouse gases have "warmed the atmosphere, ocean, and land. Widespread and rapid changes in the atmosphere, ocean, cryosphere, and biosphere have occurred."

Global warming has already had significant consequences. The emerging techniques from climate change attribution science allow us to causally connect global warming to specific observed events. For example, the report concludes that globally averaged precipitation over land has increased since 1950 and that human influence contributed to this phenomenon. The frequency and intensity of flooding events have increased. There are also increases in agricultural and ecological droughts in certain parts of the world, including west, central, and south Africa; west and central Europe; the Mediterranean; and western North America.

It is also virtually certain that human-induced climate change is the main driver of hot extremes (including heat waves) that have become more frequent and more intense since the 1950s. It is also very likely that human activity is causing ocean warming, the observed retreat of glaciers, and the decrease in Arctic ice. This, in turn, has increased global mean sea levels by 0.2m between 1901 and 2018.

Increased levels of greenhouse gasses caused by human activities have also resulted in ocean acidification and reduction in ocean oxygen levels, changing the migration patterns of large fish and sea mammals.

Climate change is happening around the world with increased frequency and magnitude of extreme weather events. (Cold extremes, including cold waves, have become less frequent and less severe.) The future evolution of the climate system will depend on how human behavior and policies will affect natural systems. The current human population stands at 8 billion and is expected to increase to 9.7 billion by 2050. Much of this increase will occur in Asia and Africa, with modest increases in the Americas. Underdeveloped economies and societies in Asia, Africa, and other parts of the world will naturally and justifiably aspire to higher standards of living and food-energy-water security. Thus, demand for energy, water, food, cement, metals, and other natural resources is expected to increase. Thus, there is a strong imperative to meet these demands while avoiding further climate change damage.

There are numerous scenarios for how socioeconomic-technological systems will evolve to impact climate. The Paris climate agreement, which was formally adopted in 2016, commits the world "to limit global warming to well below 2, preferably to 1.5°C, compared to pre-industrial levels." Nevertheless, the 1.5°C target appears highly likely to be exceeded, hence the need to adapt strategies even as global coordination is harnessed to limit further greenhouse gas emissions and global warming. Both strategy and global coordination will be required to ensure a livable, sustainable planet.

2.A.3 Targets of Opportunity for Control Systems Scientists and Engineers

The urgency and scale of the climate change problem should be an immediate call to arms for collaboration between scientists, policymakers, and experts from many disciplines. Control scientists and engineers have a crucial role to play.

In this section, we address a number of topics in which research and innovation are required and where the control community can play a leadership role. Figure 2.2 provides an overall schematic illustrating the interconnections between these topics and some of the associated technology enablers and applicable control methods and tools.

In each case, we provide a capsule summary of the opportunity target, followed by three research priorities for control scientists and engineers. (The limitation to three opportunities is solely because of space restrictions, not a lack of additional research areas with high-impact potential.) We have also indicated the timeline in which each opportunity could be expected to have broad societal impact—by which we mean that the technology would be fully matured and deployment at scale underway—as follows:

- Short-term: Societal impact by 2030
- Medium-term: Societal impact during 2030 – 2040
- Long-term: Societal impact post-2040

The authors of this section of the report plan to prepare a more comprehensive document in which the opportunities noted below and others will be discussed at greater length.

Infrastructures and Communities

An outcome of climate change is that extreme weather is becoming an increasingly serious threat to critical infrastructures such as electric power grids, transportation (road, rail, and air) systems, water treatment and distribution, communication networks, and public safety. Some of these infrastructures are critical targets of opportunity on their own and are discussed in sections below. However, many infrastructures have dynamic interdependencies. For example, repairing electric grids after a hurricane requires that the transportation system be available to move equipment and people. Similarly, traffic and water systems may depend on the availability of electric power. Such sector couplings and interdependencies need to be carefully modeled and addressed. In general, the toolkit of the control theorist or engineer on decision-making under uncertainty can help in better integration of weather and climate data for both assessment and management of climate risks to critical infrastructure.

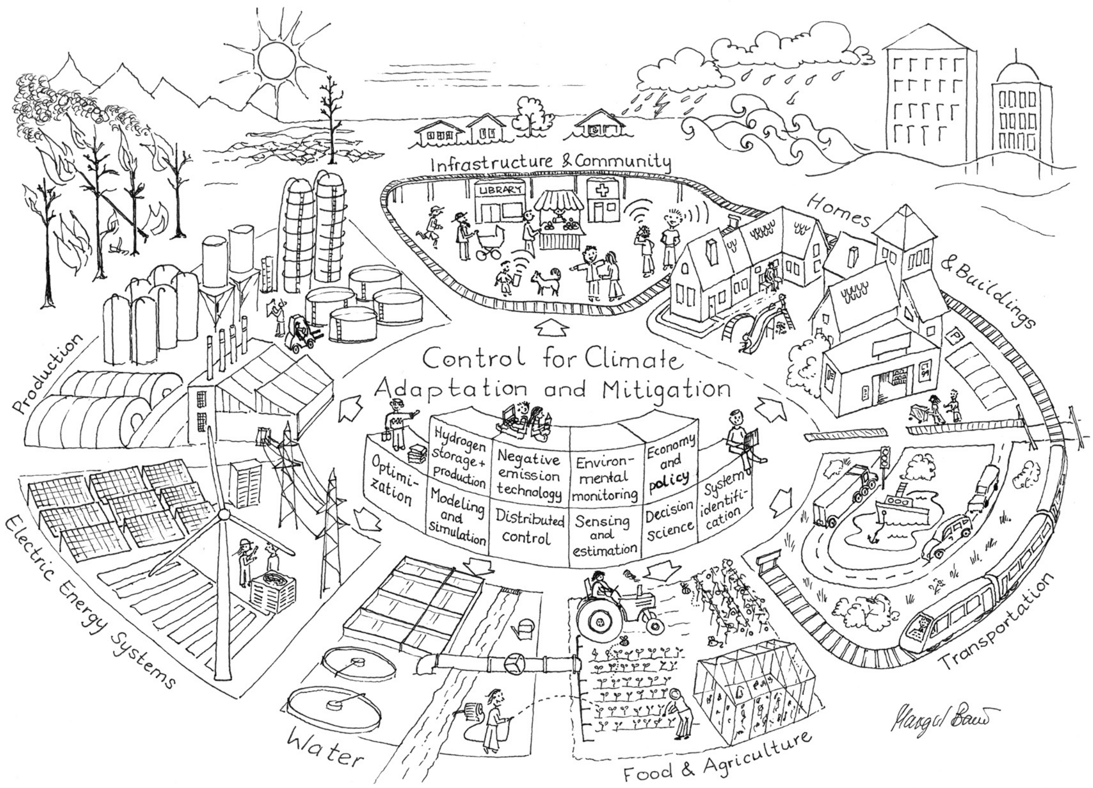


Figure 2.2: Control technology and tools can play a central role in addressing a wide range of challenges associated with climate change adaptation and mitigation. A number of these challenges involve infrastructure and communities (food and agriculture, water, transportation, energy systems, production, homes and buildings) that directly impact humans and their interactions with the natural environment.

We also include the greening of communities and civil works in this section. Example opportunities:

- Strategically allocate distributed resources for resilience in infrastructures and infrastructure networks, failure identification, adaptation, and repair (medium-term)
- Design net-zero (or net-positive) smart cities with holistic models encompassing multiple vertical functions (medium-term)
- Design resilient coastal structures and management systems capable of withstanding expected sea-level rise (long-term)

See also Section 2.C of this report for further discussion of control-relevant infrastructure opportunities.

Electric Energy Systems

A dramatic transformation of power grids is underway. The world is moving from the historical model of centralized, dispatchable power generation based on fossil fuels to a distributed system in which the most power generation is from renewable sources. The vision is to create an autonomous, continuous, secure, and decarbonized power grid that is both efficient and affordable. Today's grid technologies and operations are poorly suited for massive-scale penetration of renewable generation. New advances in control technologies and architectures transitioning from conventional centralized controls to more distributed and hierarchical controls will be essential to ensure that tomorrow's grids are stable, reliable, scalable, and cost-effective. The challenges and opportunities span transmission and distribution networks, and that division itself may need to be revisited. Example opportunities include:

- Develop control using new power electronics technologies for enabling renewable integration (short-term)
- Create new economic models for dynamic electricity pricing across different timescales (medium-term)
- Create distributed optimization and control for deep integration of renewable generation, storage, and demand management (medium-term)

Electric Power Generation

A major priority is to replace fossil fuel-based electricity production with renewable energy such as wind, solar, and geothermal. Wind and solar energy costs have decreased dramatically in many regions and nations, with the levelized cost of energy at parity with or better than fossil fuel generation. Additional renewable sources, such as marine and riverine hydropower, are also promising avenues for enhancing carbon-free generation. Opportunities for the control community exist at multiple levels, from turbines to generator assemblies to grid interfaces. Example opportunities include:

- Dynamically manage and optimize large-scale renewable energy generators (short-term)
- Develop control co-design for floating offshore wind, tidal, and wave power generators (medium-term)
- Optimize coupled cross-sector electricity generation, including hydrogen and biofuels (long-term)

Transportation

Decarbonizing road, rail, air, and marine transportation, which accounts for about 16% of global greenhouse gas emissions, is another high priority for climate change mitigation and adaptation. In the last few years, electric vehicles have made dramatic progress, but more is needed: The IEA projects that oil share will need to drop to slightly over 10% in the transport sector worldwide if net-zero emissions are to be achieved by 2050 [2]. The decarbonization of air and marine transport is in the exploratory stage. In addition to electrification, other energy strategies, such as hydrogen and ammonia, also need to be developed. This opens up a gamut of research opportunities for the control community, not only at the level of individual vehicles but also for logistics and freight networks. Example opportunities include:

- Create sensing and control for batteries, fuel cells, and emerging powertrains such as hydrogen and ammonia (short-term)
- Develop novel CPHS designs for energy-efficient smart mobility, including vehicle-to-facility/-grid power flows (medium-term)
- Control decarbonization of multiscale transportation networks (long-term)

Homes, Buildings, and Facilities

Homes and buildings account for about 18% of global greenhouse gas emissions. Some level of decarbonization will automatically occur as renewable generation is expanded, but other opportunities in the built environment also exist. The focus should not be limited to individual buildings, whether residential or commercial; larger scales of habitation are also promising targets. We need to remember that these are places where people live and work, and thus human behavior must be considered. A major opportunity here is to match the energy a community needs with renewables in a flexible manner, i.e., using demand management. Example opportunities include:

- Develop greenhouse gas emissions-aware energy management for data centers (short-term)
- Automate and control intelligent microgrids and demand response (short-term)
- Control co-design and operation of energy-efficient geothermal heat pumps for heating and cooling (medium-term)

Industry and Manufacturing

The carbon impact of industrial and manufacturing processes arises from the extensive use of power and heating, and both must be addressed. In addition, the diversity of facilities implies that some opportunities are sector-specific. However, some sectors are large enough carbon emitters (e.g., iron and steel, chemicals and petrochemicals, cement) to require dedicated attention. In some cases, new industrial processes may need to be developed as alternatives to energy-intensive and hard-to-decarbonize processes. Almost 30% of global greenhouse gas emissions arise from industry. Moves toward circular economies, which emphasize resource efficiency, reduced waste, and reuse, can also help the sustainability of manufacturing—and the inherent feedback loops involved imply strong control connections. Example opportunities include:

- Ensure energy efficiency of industrial processes and equipment, integrating heat and power (short-term)
- Optimize and control “greening” energy-intensive manufacturing sectors (medium-term)
- Optimize and control CO_2 capture, utilization, and storage (long-term)

Hydrogen, Ammonia, and Renewable Fuels

Energy-dense fuels with low carbon footprints are needed to replace conventional oil, gas, and coal for applications beyond the performance limits of current or near-term electric energy storage technologies. Clean hydrogen and hydrogen-based clean-burning liquid fuels such as ammonia and low-carbon biofuels are potential energy carriers for such applications. Hydrogen and its derivatives are expected to contribute 10% of the total emissions reductions at the lowest cost [4] by 2050, equivalent to 80 GT of CO_2. Because it is at an incipient stage, this sector presents numerous opportunities and challenges for design, innovation, and technology development. Example opportunities include:

- Design and implement electrolyzer control systems with extensive buildout for large-scale hydrogen production (short-term)
- Produce hydrogen from natural gas with carbon capture and sequestration (blue hydrogen), including advanced process control, dedicated SCADA, and reservoir modeling [5] (medium-term)
- Control co-design for the production of synthetic hydrogen-based fuels such as ammonia, methanol, e-methane, and e-kerosene, especially for the transportation sector (long-term)

Food and Agriculture

Agriculture, forestry, and land use activities account for over 18% of global greenhouse gas emissions. In addition, the impact of climate change on crop cultivation is expected to be drastic. This sector will need to rapidly evolve in order to gain flexibility and adaptability to environmental changes and to optimize production. Recently, the development of the Agriculture 4.0 paradigm—the integration of sensors, actuators, algorithms, and digitalization—has provided a platform for control systems to play a key role in sustainable agriculture. The entire food supply chain also presents opportunities for control. Example opportunities include:

- Coordinate uninhabited aerial and ground vehicles for agriculture and other land use (short-term)
- Manage food supply chain for sustainable product distribution integrating greenhouse gas footprints (medium-term)
- Control agrivoltaics for integrated agriculture and power generation (medium-term)

Water

The deleterious impacts of climate change on water are numerous and evident. Water levels are already being depleted in rivers, streams, aquifers, and freshwater lakes, even as sea levels are rising in oceans. Supplies of—and demand for—potable water need to be better managed. Water scarcity is affecting agriculture worldwide too. If left unaddressed, the adequacy of food supplies for a growing global population will be in question. Increased frequencies of floods and storm surges raise a different kind of critical concern, one that relates to climate change adaptation as well as mitigation. Example opportunities include:

- Manage and optimize agricultural irrigation and its associated infrastructure (short-term)
- Create dynamic models and control strategies for water conservation in communities, including behavioral change (medium-term)
- Control and optimize water treatment, including resilient and autonomous mobile facilities (medium-term)

Artificial Intelligence Computations

The estimated amount of computation used to train deep learning neural network models increased 300,000 times between 2012 and 2017 [6]. Computational workloads for AI and digital transformation are expected to grow up to 100-fold, outstripping GPU scaling and Moore's law [7]. The climatic repercussions of large-scale information and computing technologies (ICT), especially machine learning, can no longer be ignored. Pursuing efficient computation and communication is an urgent need, even as renewable sources are increasingly being relied upon for power. Example opportunities include:

- Develop energy optimization for training algorithms, covering both model architectures and hardware-aware execution control (short-term)
- Optimize computing infrastructure and hardware, encompassing data storage, parallelization, and repurposing (medium-term)
- Develop distributed learning algorithms, leveraging large-scale edge and cloud computing and optimizing resource allocation based on greenhouse gas considerations (medium-term)

Negative Emissions Technologies

Completely cutting out carbon emissions in the near term is virtually impossible—some forms of transport, industrial activity, and agriculture and husbandry cannot be made greenhouse gas-neutral yet. Other actions must be taken to offset these emissions. i.e., negative carbon technologies are essential. The opportunities range from landscaping to CO_2 capture, utilization, and storage. Example opportunities include:

- Measure, track, and optimize emerging carbon markets (medium-term)
- Design and control direct-from-air CO_2 capture systems (medium-term)
- Design and operate artificial photosynthesis systems (long-term)

Geoengineering

Geoengineering is a controversial topic but one that offers opportunities for controls research. By making large-scale interventions in the Earth's ecosystem, including in its oceans and atmosphere, greenhouse gasses may be absorbed, and solar radiation reaching the planet's surface may be reduced. Such interventions may involve extensive aerosol releases in the atmosphere, launching and unfurling large reflective mirrors in space, and enhancing the CO_2 absorption capacity of oceans through massive chemical doping. As these examples suggest, considerable caution needs to be exercised in even contemplating some geoengineering projects—and control scientists can help sound alarm bells where needed. On the other hand, today's large-scale projects on generation and storage also have geoengineering implications that need to be better assessed. Example opportunities include:

- Plan large-scale renewable generation sites to minimize adverse local ecosystem and climate impacts (medium-term)
- Model ultrascale spatiotemporal ecosystems, taking into account the interconnections among terrestrial, oceanic, and atmospheric dynamics (long-term)
- Develop risk-sensitive optimization and control under high long-term uncertainty, with awareness of the potential for catastrophic unmodeled effects (long-term)

Environmental Monitoring

The truism that we cannot control what we cannot measure is relevant for climate change mitigation and adaptation. Sensing, estimation, and monitoring will be necessary to assess continuing greenhouse gas emissions and the effectiveness of control strategies for reducing them. Instrumentation on the ground, on and underwater, in the sky, and in space will be required, along with integrated analytics incorporating spatiotemporal dynamic models. To be effective, environmental monitoring must be undertaken as an international collaboration. Example opportunities include:

- Develop fault detection and alarm management in large-scale instrumentation networks (short-term)
- Undertake regional and planetary-scale spatiotemporal estimation and filtering (medium-term)
- Cooperatively monitor the environment with fixed and mobile sensors (medium-term)

2.A.4 Broader Perspectives

Climate change mitigation and adaptation is complex and not just an engineering or technology problem. It requires a collaborative effort among engineers, technologists, economists, and those developing social, regulatory, and public policies. We need to be aware of this broader context as we develop frameworks for impactful research, education, and real-world translations.

Sustainable Economic Growth

Sustainability requires a balanced interplay between society, the economy, and the environment. The key question to address is how we can create and maintain a prosperous society with high quality of life for all, without the negative impacts that have historically harmed our environment and communities in the name of development. Economic growth is part of the solution, particularly for developing nations that need to raise standards of living and improve health, nutrition, and education for billions of people. Economic growth must be accomplished while effectively managing and natural resources and preserving them for current and future generations. Our vision needs to shift from consumption and waste to regeneration and recirculation—a shift that will enable future generations to thrive.

From a climate change mitigation perspective, we must minimize the need for fossil fuel energy sources while meeting economic growth targets. This is a major challenge and requires a rapid reduction in the costs of technologies. However, given that such cost-competitive technologies do not currently exist, we can expect that fossil fuel-based energy sources will be used for at least the next decade.

It is also important to steer economic growth toward a sustainable future, i.e., green growth. Investments in sustainable development, elimination of fossil fuel in various sectors, creation of circular economies, and climate change adaptation and resilience can enable economic growth and long-term sustainability.

From Research to Large-Scale Deployment

Climate change mitigation and adaptation require solutions that can be implemented at scale. Typically, energy-generating and energy-consuming systems are large-scale systems, e.g., electric grids, transportation, buildings and cities, and manufacturing. Because we are in a race to decarbonize the energy system before global warming exceeds 2°C or even goes well beyond, the transition from research to large-scale deployment is a major challenge. Dramatic reductions in the costs of wind and solar electricity are, therefore, inspiring developments.

Fortunately, there is an increasing understanding of research-based innovations. Also, government funding agencies are focusing on the need for high-impact innovations. The controls research community should systematically, creatively, and aggressively think about the research-to-real-world transition. National-scale testing and experimental infrastructures will be needed. The business case for investment in new technologies also needs to be understood.

Tight coordination and collaboration between academic researchers and industrial communities can be especially helpful. Developing collaborative networks with shared goals, datasets, and mutually reinforcing activities can also be a powerful approach to ensure research results have real-world impacts. Closer ties with policymakers and regulatory bodies will also be needed to expedite the adoption and scale-up of solutions.

Education and Awareness

The younger generation of students is deeply interested in and motivated by climate change mitigation and adaptation. Many of them see these as “their problems” that their generation will have to deal with. There is little doubt that many students interested in control systems are also interested in the challenges posed by climate change. As such, there are opportunities to include topics related to climate change in undergraduate and graduate-level controls education. These can range from using examples and projections of climate change impacts in undergraduate courses to multidisciplinary advanced courses at the M.S. and Ph.D. levels. A specific opportunity is a course on control methods for sustainability.

Awareness of the ecosystem, and especially climate, as a dynamic system, and the relevance of control science and engineering for its stewardship, need to be instilled in the broader public—an audience that educators within the discipline rarely reach out to.

This topic, which goes beyond the specific societal challenge of climate change mitigation and adaptation, is also discussed in Chapter 6 in this report.

Equity and Energy Justice

Historically, the effects of climate change and environmental degradation have dramatically impacted the disadvantaged and poorer sections of society worldwide. For example, the impact of fossil fuel plants on air quality has been disproportionately borne by poor and minority communities. Similarly, climate change impacts such as flooding, droughts, and wildfires disproportionately impact these same communities. Heat waves affect poor households without adequate air-conditioning, leading to loss of life. Sea level rise is expected to affect poorer nations, leading to large migrations. Developing nations have borne the brunt of climate change yet have made a relatively small contribution to global warming. These are the nations that need better energy infrastructures in order to develop their economies.

Development of new climate-related technologies can further exacerbate these problems. For example, increased government investments in research, development, and commercialization can further widen economic inequality. Perversely, the increased costs of climate change mitigation and adaptation might worsen the burdens on disadvantaged communities.

The control community must be fully cognizant of these issues of energy justice, intergenerational equity, and global economic development as it engages in work related to climate change mitigation and adaptation (see also Chapter 7 for additional discussion).

Transdisciplinary Collaborations

It is important to note that climate change mitigation and adaptation problems are beyond the remit of any single discipline. Progress on climate change problems will require collaboration among engineering, business, social-economic-behavioral sciences, and humanities. The control community does not necessarily have the right models to deal with these broader perspectives, particularly where socioeconomic-technical intersections occur.

Thus, it is not helpful to think of these problems as control problems. Rather, the control community should partner with experts from other domains and fields to form collaborative teams to address the large-scale, urgent, and daunting challenges. Forming such teams is a challenge considering the additional time and resources needed to develop effective and functional teams. Fortunately, there is a large body of literature, tools, and techniques for convergent transdisciplinary research and innovation. The control community should also be proactive in rewarding and recognizing its members working on such collaborations, as their work does not easily fit into the traditional framework for publications, presentations, and professional advancement.

Concluding Comments

The enormous challenges posed by climate change mitigation and humanity's sustainable adaptation to its effects offer signature opportunities for systems and control scientists and engineers. As is evident from this section, the scope and scale of opportunity for the control community are broad and deep, spanning numerous technology and industry sectors. Multiple, specific challenges in these sectors offer opportunities for examination by experts in systems, design, modeling, controls, decision-making, optimization, and related topics. A summary of the opportunities examined in this chapter—recognizing that these are but a subset of the vast array of opportunities open to the control community to showcase its capabilities to help address climate change and mitigation—is shown in Figure 2.3.

The climate change challenge is of urgent and existential importance. It renders historical paradigms of control-centric research, translation, and development inadequate. Transdisciplinary collaborative partnerships are necessary, and these must extend beyond science and engineering disciplines to also embrace the humanities and social sciences. Early engagement with industry and government will also be crucial so that deployments can take place at scale and as rapidly as possible. We have humanity's future to gain but little time to lose.

SHORT-TERM OPPORTUNITIES

- Power electronics for renewable integration
- Large-scale renwable energy generators
- Batteries, fuel cells, and emerging powertrains
- Data centers
- Automated microgrids and demand response
- Industrial processes and equipment integrating heat and power
- Electrolyzer systems for hydrogen production
- Coordinated uninhabited vehicles for land use
- Argicultural irrgation and associated infrastructure
- Machine learning algorithms
- Large-scale instrumentation networks

MEDIUM-TERM OPPORTUNITIES

- Strategic allocation of distributed resources for infrastructures and networks
- Net-zero/net-positive smart cities
- Dynamic electricity pricing across different timescales
- Floating offshore wind, tidal, and wave power generators
- Energy-efficient smart mobility, including V2X
- Geothermal heat pumps for heating and cooling
- "Greening" energy-intensive manufacturing sectors
- CO_2 capture, utilization, and storage
- Blue hydrogen
- Food supply-chain management for integrating GHG footprints
- Agrivoltaics for integrated agriculture and power generation
- Water conservation in communities, including behavioral change
- Water treatment, including for mobile facilities
- Computing infrastructure and hardware
- Distributed learning algorithms based on GHG considerations
- Emerging carbon markets
- Direct-from-air CO_2 capture systems
- Renewable generation sites minimixing adverse local impacts
- Regional and planetary-scale spatiotemporal estimation and filtering
- Cooperative environmental monitoring with fixed and mobile sensors

LONG-TERM OPPORTUNITIES

- Resilient coastal structures for rising sea levels
- Deep integration of renewable generation, storage, and demand management
- Coupled cross-sector electricity generation
- Decarbonized transportation networks
- Synthetic hydrogen-based fuels, especially for transportation
- Artificial photosynthesis systems
- Ultrascale spatiotemporal ecosystem modeling
- Risk-sensitive optimization and control under high long-term uncertainty

LEGEND

- Infrastructures and Communities
- Electric Energy Systems
- Electric Power Generation
- Transportation
- Homes, Buildings, and Facitilities
- Industries and Manufacturing
- Hydrogen, Ammonia, and Renewable Fuels
- Food and Agriculture
- Water
- Artificial Intelligence Computations
- Negative Emissions Technology
- Geoengineering
- Environmental Monitoring

Figure 2.3: Selected opportunities for control scientists and engineers to address climate change and mitigation. The color coding indicates the subsection in this chapter where the topic is further elaborated.

Recommendations

For young researchers: Climate change is arguably among humanity's most important challenges, if not the single most important. It will affect our lives and our communities for decades to come. It encompasses a very large and diverse set of technical topics. Because of its inherently transdisciplinary nature, it will catalyze continued intellectual growth and ever-expanding perspectives. You will have opportunities to work with colleagues from other fields of knowledge on collective goals. You will also have opportunities to work in a variety of organizations, including academic, private industry, nonprofit, policy, and government.

For funding agencies: The importance of climate change cannot be emphasized enough. To address this challenge and develop urgent and effective solutions, robust financial and programmatic support from government, private, and philanthropic organizations is essential. There is a great variety of research topics within this overall theme. Thus, one or more of these themes is likely to align with any given funding agency's goals. It is very important to keep in mind that climate change research is inherently transdisciplinary. Control systems experts can make important and essential contributions. But to realize this potential, research funding programs should be designed with care, encourage transdisciplinary collaborations, and integrate control systems where appropriate.

2.B Healthcare and Ensuring Quality of Life

Marcia K. O'Malley, Philip E. Paré, Sri Sarma, Tommaso Menara, Fabio Pasqualetti, Timothy O'Leary, Rodolphe Sepulchre, Francis J. Doyle III, Eyal Dassau, Sarah Fay, Anette Hosoi, Dalton Jones

> Healthcare and quality of life are important societal-scale challenges that will increase in importance as populations grow and age. The control system community has played and can play even more of a key role in the development of technologies and systems that meet these challenges and enhance healthcare outcomes.

Abstract Quality of life is critically important to all human beings. As we age and experience declining health, disease, and injury, our perceived quality of life is intertwined with issues of healthcare. There are significant challenges to achieving equitable access to high-quality healthcare globally. These challenges include the rising numbers of older populations worldwide; increasing healthcare costs; decreased trust in science and public health advisors; and the increasing occurrence, awareness, and attention to mental health issues. Despite these challenges, new technologies and capabilities can and will significantly and positively impact human health and quality of life. Current perspectives on healthcare and quality of life have been shaped over years of slow-moving, often costly progress and were then recently transformed by the COVID-19 pandemic. In this section, we discuss how pandemics and epidemics historically, currently, and continually affect the world and our field. We also provide some context for medical device development and the regulatory aspects that must be understood. We then proceed to expound upon target opportunities for the control systems community in the areas of systems-level medical technologies, neuroengineering solutions to modulate the central and peripheral nervous system, and medical devices that incorporate closed-loop control to regulate and treat disease. The section concludes with recommendations for young researchers and relevant funding agencies to help the control systems community address the societal-scale challenges of healthcare and quality of life.

2.B.1 Introduction

Quality of life is critically important to all human beings, and as we age and experience declining health, disease, and injury, our perceived quality of life is intertwined with issues of healthcare. A recent report by Deloitte examined significant drivers that are influencing healthcare globally, including the worldwide COVID-19 pandemic, rapid advances in medical science, and exponential growth in digital technologies and data analytics (see Figure 2.4). These drivers are coupled with increasingly informed and empowered healthcare consumers and a significant shift from thinking only about disease care to more broadly addressing prevention and well-being [8].

There are significant challenges to achieving equitable access to high-quality healthcare globally. In the U.S., the number of Americans 65 and older is expected to nearly double by 2060 [9], and this trend is even more apparent in other countries. The costs of treatments and prescription medications are escalating. Decreasing trust in science, healthcare workers, and public health initiatives has resulted in vaccine hesitancy. There is a shortage of healthcare professionals due to the ongoing stressors of the COVID-19 pandemic. Mental health concerns are also on the rise, with increased attention on pandemic- and age-related isolation, and the impact of social media on mental health and well-being.

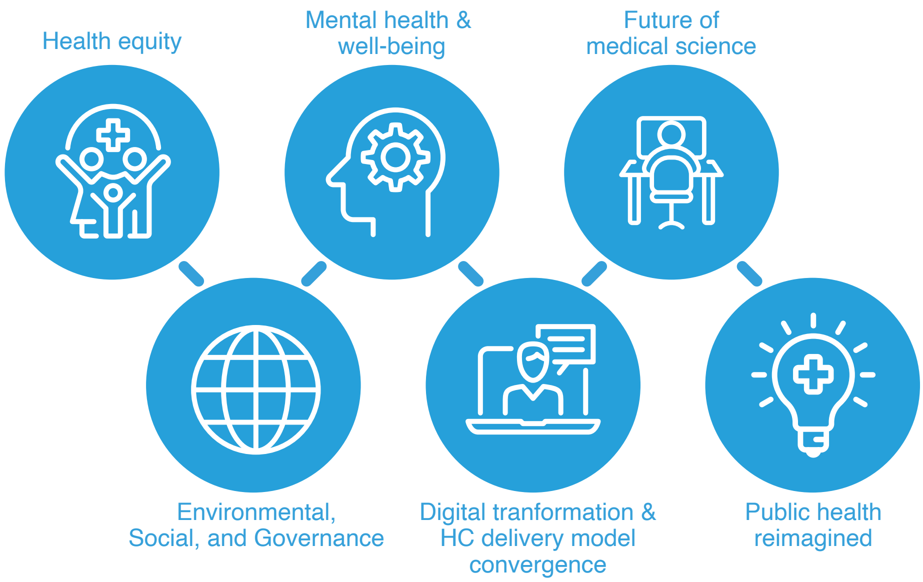


Figure 2.4: Six primary issues for the global healthcare sector [8].

But all is not dire. New technologies and capabilities have the potential to positively impact human health and quality of life in significant ways. For example, there is an increasing focus on disease prevention and health monitoring using wearable technologies and fitness trackers. Compared to other nations, the U.S. excels in common prevention measures like flu vaccinations and breast cancer screenings [10]. The COVID-19 pandemic showed the potential for novel mRNA technologies to enable rapid vaccine development. New therapeutic techniques are showing promise to deliver gene-based therapies tuned to an individual. There are government initiatives to expand access to healthcare, life expectancy is on the rise worldwide. The emergence of enabling technologies such as data analytics, novel sensors, and advanced imaging techniques creates opportunities for further advancing the quality of healthcare and health-related quality of life. The control systems community is strongly positioned to address these challenges and further advance our capabilities and technologies.

2.B.2 Current Perspectives on Healthcare and Quality of Life

There are a number of stressors that shape healthcare and quality of life. Some stressors are not directly related to healthcare but impact it nonetheless, particularly in terms of equity. For example, extreme poverty worldwide (exacerbated by the pandemic), climate change, and emerging geopolitical conflict affect healthcare access and quality [8]. Many countries struggle to offer universal health coverage to their people, leading to large out-of-pocket expenses for those seeking healthcare and treatment for disease and prevention. There are also large differences in the quality and access to healthcare worldwide, including access to facilities, state-of-the-art technologies, and even appropriately skilled clinicians. Telemedicine for disease diagnosis and counseling has emerged as one way to reach rural and underserved populations, but issues of data privacy and confidentiality have not fully been addressed [8]. Further, unintentional but systematic biases exist based on age, race, gender, caste, sexual orientation, disability, mental illness, and more, resulting in inconsistent healthcare experiences across populations [8]. Clinical trials often fail to truly and meaningfully represent the heterogeneity of populations. Artificial intelligence and machine learning techniques have great potential to transform healthcare, from offering new insights and capabilities for treatment decisions and disease detection. However, there are ingrained biases in algorithms in these systems that have the potential to result in inaccurate or missed diagnoses, leading to undesirable clinical outcomes and poor experiences for patients [8]. Finally, trust in healthcare providers varies across race, culture, and ethnicity, and this can affect whether or not people take advantage of preventive healthcare interventions like vaccines.

The previous section of this chapter focused on climate change, which has been recognized as a health emergency by leading organizations worldwide [11]. Healthcare costs related to climate change and pollution topped $820 billion per year in the U.S. in 2021 [12]. Healthcare impacts related to climate change and pollution include an increase in heat-related deaths, malnutrition due to extreme flooding and drought, and a climate that is becoming more suitable to the spread of infectious diseases [8]. Natural disasters are increasing in frequency, intensity, and variability, leading to negative downstream effects that challenge healthcare infrastructure, supply chain, and workforce [8].

Mental health and well-being are increasingly being recognized as a healthcare and quality-of-life priority. It has been stated that one-quarter to one-half of the population will deal with mental health issues at some point in their lives. The direct and indirect costs of mental illness are estimated at over 4% of global GDP, more than the cost of cancer, diabetes, and chronic respiratory disease combined [8]. The World Economic Forum identified five priorities for improving mental health, including improving access to mental healthcare and encouraging the development of novel therapies to treat mental health conditions [13]. Digital technologies hold much promise for transforming global mental and behavioral health systems, improving accessibility and affordability while also enabling scalable and fit-for-purpose solutions [8].

Advances in medical science offer many opportunities for contributions from the control engineering and science community. Increasingly, digital therapeutics are being explored for mindfulness and pain management. These technologies integrate virtual reality, gamification, and novel sensing technologies to improve patient outcomes. Digital companions can improve healthcare, quality of life, and medical outcomes by encouraging compliance with medical interventions that are used in the home, such as inhaler use or insulin delivery. In the remainder of this section, we highlight two topics that are particularly relevant today and shape the way that we think about how we can lend our expertise to improve healthcare delivery and quality of life for the world. First, we discuss pandemics and endemics in greater detail. Then, we provide some context for medical device development and its regulatory aspects.

Pandemics and Endemics

As the SARS-CoV-2 pandemic made clear, infectious diseases pose serious risks to public health. Mitigating such risks is imperative but difficult for several reasons, including uncertainty in disease dynamics and the nuances of human behavior. Societal leaders must use the information available to make decisions that minimize the costs of an epidemic. These costs include the direct health effects of the disease and secondary societal effects such as unemployment, loss of educational opportunities, and mental health concerns. There is an endless variety of infectious diseases that menace humanity, each of which requires different interventions to control. A 6-foot social distancing mandate, while effective for slowing the spread of airborne diseases, does very little to stop the spread of cholera transmitted through the water supply. As policies to prevent future epidemics are put in place, various interventions must be understood and weighed against their effectiveness for a given disease and any externalities they may pose on individuals and society.

Applying policies to mitigate the spread of disease is not new to the SARS-CoV-2 pandemic. Quinto Tiberio Angelerio, a physician in late 16th century Sardinia, advocated for lockdowns, social distancing, and quarantines to stop the spread of the plague [14]. Even without fully understanding the nature of disease transmission (the germ theory of disease spread would not be proven until Leeuwenhoek developed a strong enough microscope in 1675), Angelerio intuited that by instituting preventive measures, the spread of plague could be slowed. Angelerio's success reaffirmed one of the most powerful maxims of control theory: coarse models can be used to develop effective and robust feedback control.

From a modern perspective, understanding how to apply systems and control theory to use measurements, synthesize models, and derive effective policy recommendations will be instrumental in preventing future epidemics. Projections indicate that, while pandemics have occurred regularly over time, data indicates that they arise more frequently now than in the past [15]. Moreover, because of changes in globalization and travel, once a pandemic starts, there is a greater risk that it will spread throughout the population. Modern techniques that can be used to limit the spread of disease include testing, contact tracing, vaccination, social distancing, masks, and others. Understanding how these approaches work together and how they might be implemented in the real world is imperative. For more details and discussion on pandemics, please see Appendix 2.A.

Medical Devices and Regulatory Considerations

Medical devices contain hardware and software elements, like other devices, but might have a different product life cycle and are subject to regulatory approvals per their classification. A consumer electronic device design and development process might be as short as a few months from early design to manufacturing. In contrast, medical devices may take several years from conceptual design to a cleared product. Safety, reliability, and security are fundamental principles in the design, testing, and clinical evaluation of medical devices as well as clinical effectiveness of a device.

In the United States, the Food and Drug Administration (FDA) is the body that is responsible for protecting public health. It accomplishes this by assuring the safety (the device does not expose the user or others to unnecessary risk from its use), efficacy (the device is effective in achieving the claims that are made for its intended use), and security (the device has been designed to protect the users and others from unintentional harm and misuse) of medical devices [16, 17]. Software as medical device (SaMD), software in medical device (SiMD), and software that is used in manufacturing or maintenance of medical devices are three types of software that are related to medical devices and will need to follow the regulatory guidelines and be subject to regulatory review [17, 18]. The FDA has published guidance related to medical device development and testing for medical software and control algorithms for medical devices such as:

1. Software as a Medical Device (SAMD): Clinical Evaluation Guidance for Industry and Food and Drug Administration Staff [19]
2. The Content of Investigational Device Exemption (IDE) and Premarket Approval (PMA) Applications for Artificial Pancreas Device Systems Guidance for Industry and Food and Drug Administration Staff [20]
3. Investigational Device Exemptions (IDEs) for Early Feasibility Medical Device Clinical Studies, Including Certain First in Human (FIH) Studies Guidance for Industry and Food and Drug Administration Staff [21]
4. Policy for Device Software Functions and Mobile Medical Applications Guidance for Industry and Food and Drug Administration Staff [22]

2.B.3 Targets of Opportunity for Control Systems Scientists and Engineers

Ensuring equitable access to high-quality healthcare, advanced diagnostics, preventive methods, and infrastructure is critical to realizing continued longevity and quality of life as we age. While this work relies on the contributions of a broad team of experts, including scientists, policymakers, and clinicians, control systems scientists and engineers play a critical role in advancing healthcare and quality of life.

We have identified a number of specific targets of opportunity related to healthcare and quality of life. These are grouped in three high-level categories: **a)** medical technologies that work at the systems level, **b)** neuroengineering solutions for modulating the activities of the central and peripheral nervous system, and **c)** medical devices that incorporate closed-loop control for regulating or treating diseases. We offer a summary of the opportunity and suggest high priority research opportunities for each category.

Medical Technologies

Medical Internet of Things (IoT): Medical internet of things (IoT) is a growing collection of sensors, actuators, and clinical and personal decision support systems that are capable of sharing information, providing a connected infrastructure, a system of systems that can monitor, diagnose, and provide clinical insight or therapy to individuals. The IoT domain is large and growing, including sensors like continuous glucose monitors (CGM), wearable activity trackers or wellness devices, and GPS and calendar applications (apps) that are part of a mobile smartphone (see Figure 2.5). The information fusion of these IoT devices brings new opportunities for making medical diagnoses. It also delivers new tetraphonic modalities where control and monitoring strategies can lead to data creation, analytic engines, and therapy control. One example of such a system is automated insulin delivery (AID) or artificial pancreas for people with Type 1 diabetes (T1D). A CGM provides information to a control algorithm that is embedded in a medical device such as an insulin pump or a medical app in a phone that controls insulin delivery from a pump. Such an AID can be enhanced with information from an activity tracker, sleep monitor, or another wearable device to improve the automation of the system beyond single-input single-output (SISO) control. This can include additional predictive capabilities, disturbance estimation and rejection, and behavioral control opportunities such as targeted coaching to users [23, 24, 25].

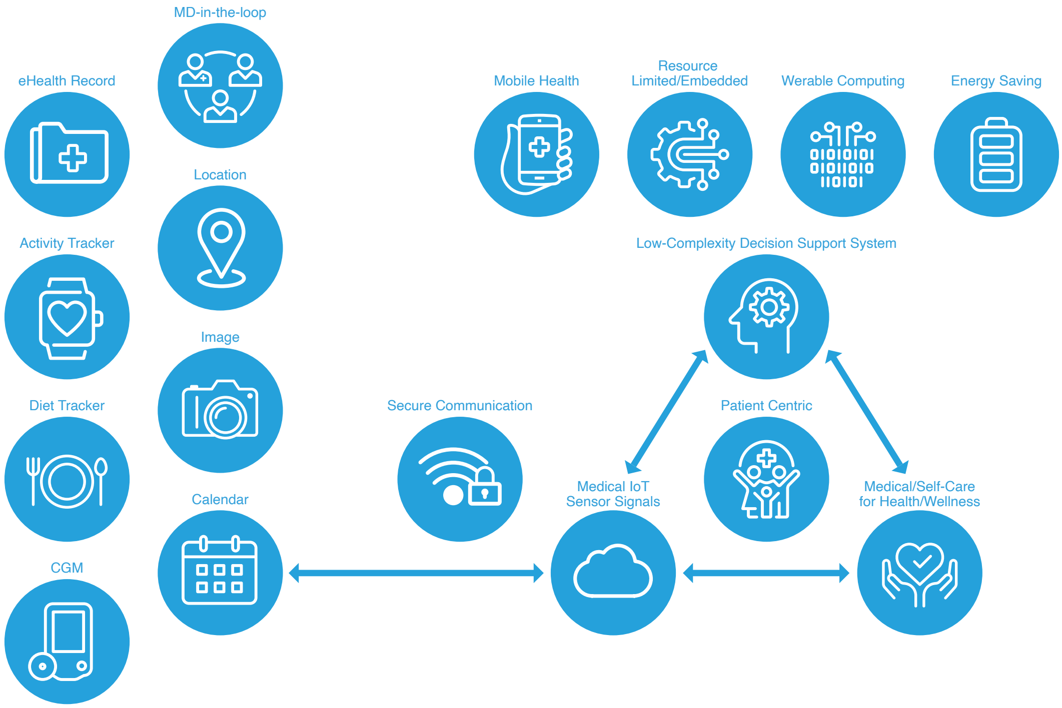


Figure 2.5: Schematic diagram of a medical IoT clinical system [23].

Decision Support Systems (DSS): Decision support systems (DSS) can operate in a stand-alone manner, or they can be embedded in closed-loop systems to provide enhanced clinical outcomes. For example, while commercial closed-loop systems are increasingly finding acceptance among individuals with Type 1 diabetes, the vast majority of individuals managing their glucose levels are doing so without closed-loop automation [26]. A good fraction of those individuals use devices that provide guidance and support for glucose management in the form of smart pens, continuous glucose sensors, advanced insulin pumps, and web portals [27]. These technologies can augment human decision-making in the form of warnings and alarms, glucose predictions, insulin dosing guidance, and more. Unlike closed-loop systems that automatically adjust insulin dosing, these systems can provide valuable guidance on fine-tuning an individual's insulin regimen (e.g., adjustments to basal profiles or updates to carbohydrate ratios). These technologies can be embedded in web portals that integrate e-health and telemonitoring programs. Often, an engineering research system will be evaluated in a decision support framework prior to testing in a fully automatic mode. Such systems have demonstrated measurable improvements in clinical outcomes [28].

Neuroengineering

The new bridges between brain science and brain technology offer novel control science opportunities. Brain signals are dynamic and multiscale. Processing brain data and modeling brain functions must acknowledge those properties. Brain circuits are plastic, variable, and uncertain. Designing machines that interact with brain circuits must acknowledge those properties. Feedback is central to brain operation at any temporal and spatial scale, so controlling brain circuits must acknowledge those properties. Brain sensing and actuator technologies are booming, but the control theory of machines that interact with the brain is still in its infancy.

Progress in brain medicine and in brain-inspired technology is expected but will be increasingly limited by the absence of a control theory of brain-machine interaction. For example, deep brain stimulation (DBS) is an approach to treating neurological and psychiatric disorders (such as Parkinson's disease) by stimulating neurons in the brain with an implanted electrode. Broadly stated, neurostimulation relies on both measurements of patient state (e.g., limb accelerometers, EEG, single neuron recording) and an actuation stimulus (e.g., DBS, pharmacological agent) [29]. Current clinical approaches rely on open-loop stimulation protocols (duration, amplitude, frequency of a pulse train). However, recent short-term clinical studies show promise for closed-loop strategies to automate the treatment [30, 31].

Over the years, a wide range of algorithms have been investigated, ranging from Proportional-Integral-Derivative (PID) to geometric control to adaptive controllers [30]. At the core of most algorithms being investigated is a method for learning and adapting the algorithm in response to the patient's changing characteristics over the course of days, months, or years [31]. Commercial DBS systems from multiple companies (including Medtronic, NeuroPace, and Boston Scientific) have been approved by the FDA. Some of these are for open-loop treatment, and at least one is for closed-loop treatment of Parkinson's disease, dystonia, and obsessive-compulsive disorder. Numerous challenges remain in developing safe closed-loop DBS systems, including the determination of reliable biomarkers (measurements), battery design, decision support systems, and control algorithm optimization.

Opportunities for control systems scientists and engineers exist within the field of neuroengineering beyond the application of DBS. We categorize these in three areas.

1. **Neuroscience:** New measurement devices provide neuroscientists with an unprecedented wealth of data covering a broad range of spatial and temporal scales. New modeling and analysis methodologies are needed to acknowledge the multiscale and dynamic nature of brain function.
2. **Neuroengineering:** New actuation devices lead to increased interaction between machines and brains that provide novel opportunities to restore impaired brain functions. Control theory is needed to acknowledge the closed-loop nature of brain-machine interaction, and new methodologies are needed to account for the spiking nature of brain signals.
3. **Neuromorphic Engineering:** Brain-inspired sensors, actuators, and computing devices offer novel opportunities to make artificial machines more like animal machines in their energy efficiency and attention capabilities. Reducing the carbon footprint of digital AI will be a major challenge in the forthcoming decades.

For more details and further discussion on neuroengineering, please see Appendix 2.B.

Closed-Loop Control of Medical Devices

The opportunities for closed-loop control of medical devices is much broader than the snapshot of applications summarized here, and they continue to grow. Researchers continue to explore novel control strategies for drug dosage for applications ranging from HIV to cancer. New directions are also revealed, motivated by the challenges of the COVID pandemic, such as the work of Sri Sarma and colleagues at Johns Hopkins University. They propose a solution to automate ventilator control to optimize levels of peripheral capillary oxygen saturation. Using data-driven linear parameter-varying (LPV) dynamical systems models, which relate patient clinical state and ventilator inputs to the output oxygen levels, the team proposes the design of controllers using linear matrix inequality (LMI)-based methods [32]. Below are a few scenarios that offer unique opportunities for our community and potentially impact healthcare and quality of life for broad populations.

Automated Anesthesia: Closed-loop anesthesia systems have been researched, designed, and evaluated in the clinical setting for over 20 years. The operating room environment is quite complicated, with the need to monitor many critical variables on a real-time basis. However, a distinct advantage of this setting is that nonportable sensors can be employed that would be too cumbersome or impractical for ambulatory applications. The physiological variables that must be monitored include the depth of anesthesia (DOA), metabolic status, hemodynamic state, and other physiological conditions. The DOA has been the subject of intense research over the last two decades, and sensors are available to determine the depth of anesthesia through correlations. These sensors are inferential in nature. They do not directly measure the medical state of anesthesia, which is characterized by such patient responses as hypnosis, amnesia, analgesia, and muscle relaxation. Instead, they measure the state of electrical activity in the patient's brain. A promising method is the bispectral index, derived from signal analysis of an electroencephalograph (EEG) [33, 34]. A commercial tool for translating EEG signals using wavelets to indicators for closed-loop anesthesia control is available (WAVCNS from NeuroSENSE) [35].

A variety of manipulated inputs are also available, resulting in an intrinsically multivariable control problem. Some candidate-manipulated variables include vasoactive drugs such as dopamine and sodium nitroprusside, as well as anesthetics (isoflurane, propofol, isofluorane, remifentanil, etc.). Like the developments with the artificial pancreas, a variety of control algorithms have been investigated with the primary categories being PID control, model predictive control (MPC), and rule-based control [33]. Pharmacokinetic/pharmacodynamic (PK/PD) models have been used directly in model-based schemes and indirectly in PID and rule-based designs that employ predictions of drug levels and use override safety systems [35]. Multiple simultaneous PID controllers have been clinically evaluated, with demonstrated improvement over manual control (as determined by postoperative cognitive measures), although the differential benefit of one control loop over the other could not be quantified [34]. Recognizing the need to have robust solutions across populations, recent work has also begun to explore the use of scaling principles to create a personalized PID algorithm that led to reduced interpatient variability [36].

Cardiac-Assist Devices: Cardiac-assist devices are mechanical pumps that provide cardiac output at an appropriate pressure that allows normal circulation of blood through the patient's body, subject to the changing demands for cardiac output as a function of the patient's state (e.g., level of exercise, emotion, posture, etc.). These devices are critical to the support of patients in the end stages of congestive heart failure [37]. The ideal device would mimic the body's own mechanisms for maintaining cardiac output at target levels. However, the first-generation devices were rather primitive in terms of automation, requiring the patient to adjust the set point. One of the more challenging aspects of the control design problem for ventricular-assist devices is the correct anatomical placement of the sensors and actuators for individual patients and doing so to limit susceptibility to infection. The current generation of devices span catheter-based interventions for short-term circulatory support all the way to long-term implantable pumps [37]. Most of the current devices on the market incorporate proprietary pump and flow control algorithms with scant details on the controller. However, recent results point to the growing availability of patient data that can be used to build predictive models of hemodynamic variables using AI-based and other methods (see, for example, Abiomed's devices [38]).

Diabetes Devices—Connected and Automated: Insulin management or delivery can be divided into two main classes. Automated insulin delivery systems (AIDs) combine an insulin pump, CGM, and algorithm (see Figure 2.6 for a taxonomy of automated insulin delivery systems). Smart insulin pens or caps report the amount of insulin dialed and/or delivered to a clinical decision-support system (CDS). AID systems and smart pens are designed to improve glycemic control for people with diabetes and reduce the day-to-day burden and the decision-making around the amount of insulin that should be delivered. The FDA has already cleared several systems for use and others are under development by both commercial entities and individuals or communities like the DIY movement. Different feedback-control strategies ranging from PID and fuzzy logic to model-based and optimal control, are used and are embedded in the insulin pump or run as a mobile medical device in a smartphone [39].

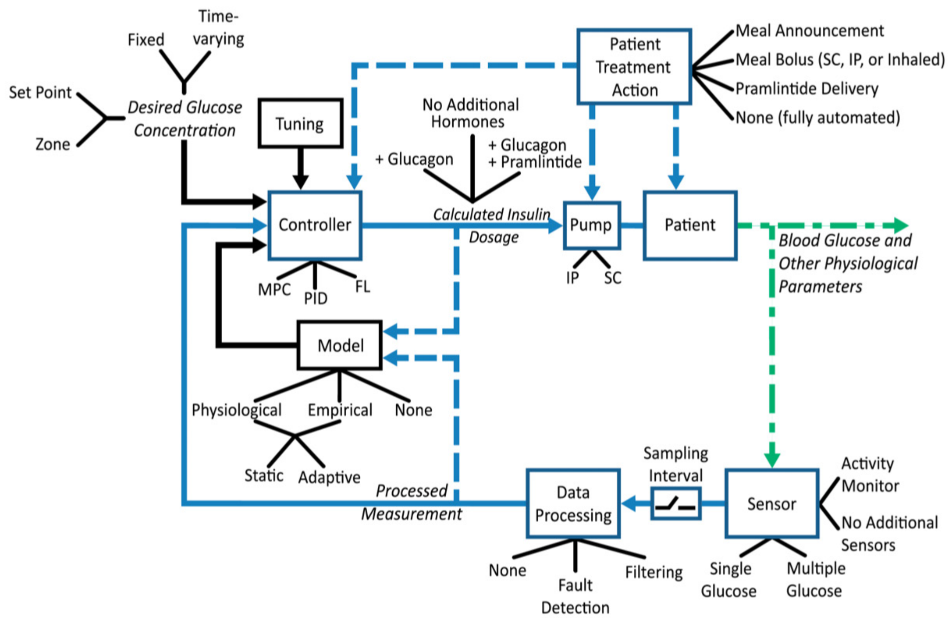


Figure 2.6: Taxonomy of automated insulin delivery core design and options [39].

While great strides have been made with the artificial pancreas, numerous open challenges remain for continued research and development. All the current products are "hybrid" systems. They require the person with diabetes to estimate and program the bolus associated with meals. It is also fair to say that the current devices were designed to compensate for overnight excursions (i.e., avoiding hypoglycemia) as well as meal compensation (in a hybrid model). However, they have not been optimized to handle exercise and strenuous activities.

There are myriad challenges with designing a dosing strategy during exercise, including the availability of suitable sensors and how to customize the strategy to the fitness of the individual. Another major challenge is distinguishing between exercise in a low-stress training mode versus exercise during a stressful competition. The former tends to deplete and lower blood glucose levels, while the latter can lead to elevated glucose levels, which block insulin receptors [40]. Consequently, an identical signature in activity-based sensors might require diametrically opposite compensation strategies in the two scenarios.

There are limitations to what a device can do that is extracorporeal, but there have been promising results with devices like the DiaPort that point to the promise of full automation [41]. Another striking observation about the artificial pancreas is that it is intrinsically a single-input (insulin), single-output (glucose) control system. Further advances are needed in developing sensors for implicated physiological variables (insulin, ketones, cortisol, lactate, etc.) that are sensitive at physiologic levels, can report in real-time, and can be designed for on-body use [42]. On the input side, there have been promising early results with glucagon as a counter-regulatory hormone that might be dosed in tandem with insulin. There have also been studies of amylin in closed-loop systems where amylin modulates the meal (i.e., disturbance) response and allows insulin to more fully compensate for meal excursions [43, 44].

Recommendations

For young researchers: To help improve healthcare and quality of life, young researchers must do more than acquire strong analytical, quantitative, and qualitative skills. We encourage you to seek out and embrace transdisciplinary collaborations, continue to learn about equity and social justice, and always be inquisitive. Just because you can formulate and solve an optimization problem does not mean that you have the correct or best cost function. Further, it is essential to work hand in hand with professionals who have domain expertise, especially in healthcare. Finally, while this section has explored numerous research directions related to healthcare and quality of life, do not limit yourselves. Be creative and seek out other important, interesting problems that will improve the lives of individuals, communities, and/or society.

For funding agencies: The control community has played essential roles in innumerable fields over the last several decades and will continue to contribute, especially with respect to healthcare and quality of life. It is essential that funding mechanisms be of the scale and scope necessary to support cross-disciplinary teams that can cultivate complete, holistic solutions.

2.C Smart Infrastructure Systems

Dan Work, Carlos Canudas de Wit, Masayuki Fujita, Hideaki Ishii, Karl H. Johansson, Anuradha M. Annaswamy

> Infrastructure systems such as those in transportation and energy are primed for digital transformation that will enable new capabilities through advances in communication, sensing, and methods for interpreting data. We discuss novel challenges in building smart infrastructure due to the complexity of societal-scale problems as well as opportunities to contribute to solutions through interdisciplinary research

Abstract Infrastructure systems are the basic physical and organizational structures and facilities that a society or enterprise needs to operate. They support planning, operation, and safety. In this section, we will highlight how advances in communication, sensing, and methods for interpreting data are enabling new capabilities for such systems. Built infrastructure systems are increasingly connected and are actuated, offering opportunities to increase resilience, sustainability, and performance. Digital transformation will make future infrastructure systems more interactive, dynamic, and integrated. By combining sensing, communication, and actuation at scales that have previously been out of reach, new high-performance infrastructures will emerge that can adapt and respond to the changes in the environment and needs of society. Such dramatic changes are primed for many societal-scale infrastructure systems. We provide a brief overview of the prospects and future challenges in several sectors, including transportation, energy, water, and food and agriculture. We further discuss why solving these societal-scale problems with smart infrastructure will require interdisciplinary collaboration across multiple domains beyond control systems. This may range from other engineering disciplines to public policy and administration.

2.C.1 Introduction

Infrastructure systems are the basic physical and organizational structures and facilities (e.g., buildings, roads, and power generation) that a society or enterprise needs to operate. Infrastructure systems have been around as long as urban centers, supporting a society's needs for its planning, operation, and safety. Advances in communication, sensing, and methods for interpreting data enable new capabilities for designing, managing, and operating infrastructure systems. Built infrastructure systems are increasingly connected and are actuated, offering opportunities to increase resilience, sustainability, and performance of critical systems that support society. Facilitating this transformation while recognizing the unique needs of each community is important to meet the growing complexities and needs of future infrastructure systems.

While today's infrastructure systems are largely static, infrastructure systems of the future will need to be interactive, dynamic, and integrated. While today's systems are physically coupled, they are often managed independently, which does not allow interdependencies to be exploited. Future infrastructure systems will combine sensing, communication, and actuation at scales that have previously been out of reach, enabling new high-performance infrastructure that can adapt and respond to a changing environment and changing societal needs.

Figure 2.7: Multimodal transportation systems are undergoing a significant technology transition to meet mobility system performance needs of society. Image Source: USDOT Strategic Plan.

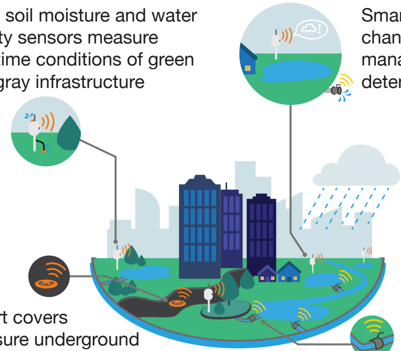


Figure 2.8: Water networks are beginning to include sensing and actuation to improve sustainably. Image Source: USDOT Water Lab, University of Michigan Strategic Plan.

2.C.2 Drivers: Infrastructure Systems Undergoing Digital Transformation

Many societal-scale infrastructure systems are primed for digital transformation, including transportation, energy, water, and food and agriculture (see Figures 2.7 and 2.8 for a few examples). The same technologies used to advance smart infrastructure will also help develop a smart society.

Transportation

Smart transportation systems are well-positioned for dramatic breakthroughs in performance. The widespread adoption of smartphones, which began about 15 years ago, helped unlock personal navigation, real-time traffic monitoring, new ride-sharing services, and more. While commuters are connected through advances in smartphone technology, the vehicles themselves and the infrastructure (e.g., traffic signals) do not share this same level of sophistication. Commercial vehicles are sold worldwide with Society of Automotive Engineers level 1 and level 2 automation capabilities, but true "brain-off" driving has not yet matured into a commercial technology. There is a large gap between simple automation technologies that are easy to use and offer correct modalities and advanced machine learning approaches that can enable transformative performance but often introduce brittle failure modes. Moreover, today's commercially available automated vehicles operate independent of the infrastructure and of other vehicles. Emerging 5G technologies will soon break the current paradigm and open possibilities where cooperative driving and cooperative infrastructure can improve safety, reduce emissions, and improve overall transportation efficiency. Creating the estimation and control techniques to manage mobility infrastructure at a societal scale is a core challenge for the coming decade.

Water

Population growth, urbanization, and climate change are dramatically increasing the demand for water systems and driving the need for a holistic, intelligent decision-making framework for managing water infrastructures. At the same time, the availability of new, cheap, and accurate sensing and actuation devices opens the possibility to significantly improving our understanding of the needs and operational constraints of today's water networks. By increasing the availability of data describing the water network, it becomes possible to create new data-driven tools for automating decision-making in the design, monitoring, control, and security of urban water systems. The multidisciplinary design and analysis of monitoring and control methods for smart water systems will open new possibilities to extract efficiencies that are out of reach for today's water networks. Possible enabling technologies include new approaches for state estimation of water hydraulics; detection and isolation of water leaks; water pressure control; water quality state estimation; water quality control; water contamination fault diagnosis; emergency response to water contamination events; and cyber-physical security of water systems.

Energy

Meeting the global energy demand while simultaneously addressing the environmental impact of energy production infrastructure is a key challenge of our time. Power systems, where innovations in the use of solar and wind renewable energy started in the late 2000s, may lead in the transformation to smarter infrastructures.

Power networks are becoming increasingly hard to manage because of the large number of suppliers. This makes it difficult to plan, dispatch, and balance supply and demand. Demand management solutions like dynamic pricing can help influence consumer electricity use to respond to the intermittent generation of renewable energy. While electric vehicles can contribute by bringing more batteries to the system, power grids must prepare for a world in which large portions of the ground transportation infrastructure will move from internal combustion to plug-in electric technologies—a dramatic shift that will occur over the next decade. Better modeling, connectivity, sensing, and communication are needed to address a significant expansion in the complexity of the power grid while simultaneously meeting a wider range of performance targets.

Agriculture & Food

Key challenges facing agriculture and food stem from the need to improve crop yield and increase the reliability of the food supply chain to reduce food loss from the farm field to the table. Food production must dramatically increase to meet the needs of a global population that continues to grow.

The importance of the food supply chain is easy to identify by recognizing that agriculture plays an important role in many of the 17 Sustainable Development Goals adopted by the United Nations [45]. Notably, Goal 2 is to "End hunger, achieve food security and improved nutrition, and promote sustainable agriculture." One major challenge to meet this goal is that current resource-intensive farming paradigms have proven to not be scalable due to their enormous environmental impacts that cause deforestation, water scarcity, and greenhouse gas emissions. Consequently, new approaches are needed to revolutionize the current practices of agriculture. This might result in completely new forms of agriculture such as agriculture forestry, intercropping, and vertical farming. Alternatively, it may require that agriculture become more efficient by moving to technology-enhanced precision agriculture in order to increase the production of highly nutritious foods at minimal environmental impact. These improvements should go hand in hand with increasing the sustainability and efficiency of the food production chain. Control technologies can contribute to these goals in the development of advanced machinery, data-gathering and processing infrastructure, automated decision-making schemes, and precision agriculture.

Beyond Infrastructure: Toward a Smart Society

A smart society is a human-centered society that aims to solve societal problems through a system that integrates cyber and physical resources. It is enabled by many of the driving technologies facilitating smart infrastructure, but it is more broad in scope and potential impact. A smart society connects all people, things, and systems through information networks, and a variety of knowledge and information is shared to enable the society to overcome various challenges. In a smart society, advanced methods will provide people with information when needed, and automated technologies are expected to solve various problems. For example, automated technologies will do tasks and make adjustments that humans have performed. The new value created by innovation will eliminate disparities based on region, age, gender, language, and other factors, enabling a finely tuned response to individuals' diverse and latent needs. Social well-being will be improved for everyone by promoting fairness, inclusivity, and trust. When data is used, the privacy of the individual is respected. When services are designed, they support the needs of all citizens rather than exacerbating the digital divide. Although smart society initiatives are beginning, most are still in the verification stage within individual cities/regions. Many opportunities are yet unexplored.

2.C.3 Opportunities for Control

Resilience & Robustness

While these and other societal infrastructure systems are facing extreme challenges that drive needed innovation, it is important to recognize that they do not operate independently. For example, future freight transportation systems will be electrified, placing new demands on the power grid. These same transportation systems are responsible for moving goods and people, including vital assets in the food supply chain. The increased demands on the power grid will increase demands on the water needed for cooling. Water networks require more energy to deliver water to population centers affected by climate change. Building methods to advance the management of coupled infrastructure systems without simultaneously creating destabilizing feedback loops will require new design and control approaches to manage the critical infrastructure systems of the future.

Convergence

One of the key opportunities of control is to study how societal drivers such as smart infrastructure systems or their integration influence individuals, groups, or society. These systems are highly interactive and affect people's everyday lives, requiring new techniques in building algorithmic models of human behavior at different levels of fidelity (such as individual or societal-scale models). This also creates an opportunity for interdisciplinary research between control theory and other fields in behavioral economics, cognitive science, or human computational modeling. Techniques that capture human bounded rationality, rational inattention, or models of decision-making along with data-driven approaches such as imitation learning or preference-based learning can enable building computational models of human behavior. Although these models are effective for prediction, they are often not sufficient for interaction as they do not capture the closed-loop dynamics of interaction between humans (or groups of humans) and smart infrastructure systems. Control theory can potentially provide approaches for leveraging and improving human behavior models in a closed-loop manner, as well as a formal approach for analyzing equilibria that can emerge from interactions between humans and these intelligent systems during closed-loop, continual, and repeated interactions.

Multiscale Modeling & Control

The development and operation of smart infrastructures often require the orchestration and control of the emergent collective behavior of large ensembles of entities or agents (e.g., sources and sinks in a power network or intersections and autonomous vehicles in a transportation network) interacting with one another in a nontrivial way over a complex graph. Steering the behavior of these large-scale complex systems can be achieved by controlling the nodes of the network, endowing the edges with some dynamics or communication protocol, or manipulating the structure of the network itself. The problem then becomes understanding how to close the feedback loop across different scales. Typically the problem is to control the emergent behavior of the network system or infrastructure of interest at the macroscopic scale while acting on the nodes, the edges, or the network structure at the microscopic scale. A feedback connection needs to be established between the sensed or estimated macroscopic behavior of the ensemble and the action (or actuation) that needs to be taken at the microscopic level to steer the resulting collective dynamics in a desired direction. From a control viewpoint, such complex network systems are examples of large-scale dynamic systems consisting of many continuous- or discrete-time units interacting over a web of complex interconnections that can be either static or time-varying.

A key opportunity for control is to address the open challenge of understanding how to close the loop across different scales. This can be achieved by developing new decentralized and distributed estimation, control, and optimization strategies to endow agents at the microscopic level with the ability to sense and control the emergent macroscopic behavior stemming from their collective dynamics. This requires extending to this class of systems analysis and synthesis strategies from classical control theory (e.g., controllability, observability) without neglecting the often nonlinear, uncertain nature of the agent dynamics and the time-varying, evolving nature of the underlying network of interest. Another fundamental problem is to analyze and prove convergence of the controlled network system to certify its stability and robustness. A further opportunity for control is the use of data-driven and learning-based control strategies (e.g., multi-agent reinforcement learning). These may become fundamental in solving these types of problems because model-based approaches can be unable to cope with the multiscale, uncertain, and complex nature of the problem of interest.

2.C.4 Recommendations

- **Promote convergence research.** Solving smart infrastructure problems will require collaboration across multiple domains beyond control systems. These collaborations will include other engineering disciplines, public policy, and administration. Acknowledging the needs and role of humans in developing these systems will also expand the interdisciplinary frontier.
- **Expand R&D funding and experimentation sites.** Many smart infrastructure systems are critical to the well-being of society, and it is not possible to easily experiment on these systems without degrading day-to-day operational performance. R&D funding is needed to expand realistic testbeds and test facilities to produce critical datasets and validate new methodologies.
- **Expand collaboration with industry and government partners.** Many critical infrastructure systems are operated by government agencies in the public interest but depend on scalable technologies from industrial partners. Collaboration by academic, government, and industry organizations can leverage complementary areas of expertise that can help accelerate the pace of designing and deploying new innovations for these systems.

2.D The Sharing Economy

Christos G. Cassandras, Robert Shorten, Anuradha M. Annaswamy

> The Sharing Economy is a disruptive paradigm based on replacing traditional notions of ownership with mechanisms based on sharing/on-demand access. Ideas central to this emerging discipline resemble many system-theoretic concepts. Thus, systems and control concepts and methods have an important role to play in managing shared resources and in efficient market design.

Abstract Sharing Economy is an economic model that transitions the current economy from owning (sole possession) resources to accessing them (sharing) only when required. While the benefits are numerous, this transition also entails multiple challenges and considerations, including privacy, safety, security, fairness, risk, market distortion, and strategic behavior. Control-theoretic methodologies are well suited to address many of these challenges, specifically those that involve effectively managing constrained resources at both the aggregate (societal) and individual user levels. We argue that the Sharing Economy can be viewed as a feedback control problem in which individual agents make up the plant and the control allocates the constrained resource among agents. In such a framework, markets may be designed so consumers can exchange goods and services in a peer-to-peer context using mechanisms such as transactive control. System and component-level decisions would be made with economic transactions negotiated between the components of the system. The management and design of shared services, while preserving privacy, is also an area familiar to control engineers, where the goal is to develop infrastructure and algorithms to support sharing a pool of objects in a community. A unique feature of the Sharing Economy is the importance of fostering compliance through the design of social contracts using new cooperative control tools and modeling human behavior in a closed-loop system. Game theory has an obvious role in developing compliance mechanisms, including access control and adversarial settings.

2.D.1 Introduction

The Sharing Economy is an economic model that transitions from owning resources to accessing them only when required. This transition is a consequence of dwindling resources and a higher cost for those resources. This in turn makes it necessary to embrace circularity, monetizing assets that are not being fully used. It requires a multitude of concomitant changes in how business is conducted, how the infrastructure is designed to support collaborative consumption, and how user behavior adapts at both individual and aggregate levels (see Figure 2.9). Examples of the Sharing Economy abound in every sector of the economy: transportation network companies (e.g., Uber, Lyft) and vehicle sharing (e.g., Zipcar), peer-to-peer lodging platforms (e.g., Airbnb, Vrbo) in the service industry, microfinance institutions (e.g., Sardex, Prosper), co-working to share office or product storage space, and reselling/trading platforms (e.g., eBay, Craigslist).

In the energy sector, the Sharing Economy occurs, albeit indirectly, between the power grid infrastructure and the building infrastructure, through the notion of flexible consumption serving as a negative load. Peer-to-peer trading mechanisms are also emerging between electrical energy and thermal energy networks to ensure adequate resources for both heating and electricity. Potential exists in introducing such transactive mechanisms between the electrical network and the train network to leverage the same notion of flexible consumption. The interest these new economic paradigms has attracted for both industrial applications and broader societal adoption is well documented [46, 47, 48, 49, 50, 51].

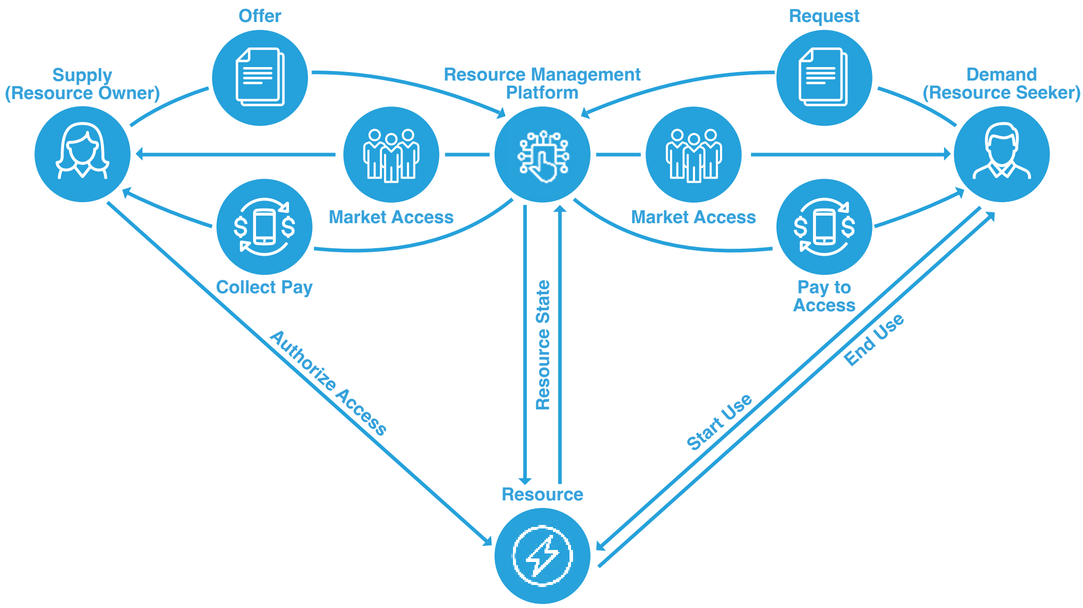


Figure 2.9: Overview of a Sharing Economy system.

The motivation for a Sharing Economy is multifaceted. For starters, many consumer resources have extremely low use (e.g., the average car use is about 4%). Sharing this resource for on-demand use can dramatically reduce the cost of ownership, including insurance and maintenance. Globally, competition for resources is intense, driven by an expanding gray community and middle class. In the face of such pressure, sharing resources would appear to be one tool that would allow living standards to be maintained, or even increased, in the face of constrained resources.

From a societal standpoint, reducing the number of certain resources, such as pollution-generating vehicles, can benefit the environment. It is worth noting that managing access to shared resources, as opposed to building more efficient devices, will be a central pillar in the fight against global warming. William Jevons noticed in the 19th century that developing new, more efficient technologies often stimulates more aggregate demand and counteracts the impact of new technology. Pollution in our cities and our insatiable demand for energy readily yield, at least in part, examples of this effect, known as the Jevons paradox. For example, the fact that cars are rated on per device emissions factors ignores the impact of this effect and allows major technology providers to ignore the aggregate impact of their devices. Using control theory to manage access to shared assets is a fundamental tool to combat the effects of Jevons paradox.

A Sharing Economy also offers access to resources that might not be feasible or practical to own (such as a car) or to obtain (such as a bank loan) without assuming prohibitive amounts of risk. At the same time, features intrinsic in sharing platforms, such as ratings and reviews, contribute to building stronger user communities. This keeps both providers and consumers honest, promoting interpersonal collaboration and even using a platform's influence to assist those in need (e.g., Airbnb coordinating free accommodations for people affected by natural disasters). Moreover, a Sharing Economy is greatly facilitated by technological developments that enable ubiquitous networking and easy access to platforms that coordinate resource sharing and build constructive relationships where none existed before.

It is important to recognize and pinpoint the fundamental differences between the operation of a Sharing Economy (also referred to as collaborative consumption) and current practice. One such difference is that many resources and services are not centrally created but may often be produced by consumers, leading to the emergence of prosumers [52, 53, 48]. An example is the use of a parking space consumed when one uses it for their own vehicle and then produced as a resource for somebody else. Although joint production and consumption itself is not new, the scale at which it is now being carried out is unprecedented in the volume of transactions, number of people involved, and the granular nature, both in space and time, of the resources shared.

With the benefits of a Sharing Economy come numerous challenges, many of which constitute problems that can be squarely addressed using approaches intimately familiar to the control systems community. Those challenges include:

Privacy, Safety, and Security: In the networked environment required for information sharing, privacy and safety concerns arise similar to those already extensively studied in the context of CPS. These take center stage in the context of the Sharing Economy as they become more complex due to the presence of multiple owners.

Market Distortion: The introduction of multiple ownership, as well as new mechanisms such as peer-to-peer markets, is a complex systems problem. When a peer-to-peer marketplace is established within an already existing traditional marketplace, the latter often experiences new kinds of short-term disruptions (i.e., instabilities), as seen in some housing markets (due to Airbnb) or urban transportation (due to Uber and Lyft). Such short-term instabilities may have to be analyzed and understood.

Fairness: Since the Sharing Economy critically depends on access to information through networking technology, this can result in fairness concerns regarding the availability of this technology to all societal groups and the way this could be manipulated. This and related concepts, such as energy justice, are becoming increasingly important and can be analyzed through the systems and control lens. In addition, fairness includes respecting the preferences of individual agents or groups when resources are allocated.

Risk: In many ways, replacing ownership with sharing entails a much higher degree of risk. As a prosumer, one faces the risk of not being properly compensated for a resource or suffering damage to that resource by those who consume it. The effective management of risk is a critical issue in a Sharing Economy. This includes risk to the individual and collective actions, which have the potential to spread risk over groups of prosumers.

Gaming, Collusion, Cooperation, and Other Strategic Behavior Considerations: Cooperation is synonymous with sharing and is a cornerstone of the multi-agent dynamic system that forms the backbone of an effective Sharing Economy. At the same time, the presence of multiple owners and prosumers introduces several challenges, including gaming, collusion, and overall strategic behavior. All of these aspects are important features of a multi-agent dynamic system and have to be investigated. Moreover, this has to be done while recognizing the time-varying nature of the decision processes involved.

Despite these challenges, it is clear that the market forces driving the emergence of a Sharing Economy are irresistible, and many large corporations are moving quickly to embrace the opportunity to develop new shared services. A striking example is the automotive manufacturer Riversimple, whose vehicle components may be owned by different entities.

Many ideas that are familiar to the control community are central to these market forces. Control-theoretic methodologies are well-suited to address many of the challenges that arise in Sharing Economy applications and business models, specifically those that involve the effective management of constrained resources at both the aggregate (societal) and individual user levels. The use of feedback, as an example, is a tool that can be used to deal with Jevons paradox, where technological innovation may result in unintended consequences due to increased aggregate demand for certain resources that would counteract the effect of sharing them [54] (also sometimes referred to as the rebound effect).

2.D.2 The Sharing Economy and Control

Interest in the Sharing Economy is being accompanied by innovation in three broad areas: new business models, enabling technologies, and analytics. Systems and control theory have a role to play as new ideas emerge across all three areas. At a very basic level, control is an essential component of the design of Sharing Economy systems. Sharing a constrained resource involves regulation where there may be additional complexities, such as determining the set-points. Sometimes, these problems rise to brand new classical problems in control or classical problems but new human-centric constraints that may seem unfamiliar to practitioners.

At a fundamental level, the need for control arises for two reasons. First, the Jevons paradox (the fact that new technologies can stimulate demand to improve efficiency) manifests itself in many of the important problems that we are interested in solving (such as pollution in our cities, greenhouse gas emissions, and even poverty). Such examples can be also readily found in other domains, and one approach to address the Jevons paradox is to manage the aggregate effect of human behavior as part of the system being designed. This involves acknowledging the fact that resources are always constrained and should be shared, and that sharing them can be managed by treating sharing as a combined multilevel optimization and control problem. In the case of vehicle pollution, the safe level of aggregate emissions could, for example, be treated as the resource to be shared among drivers in the city [55, 56].

In terms of novelty from a control perspective, the feedback loop must be designed to account for the human-centric constraints of the ensemble being controlled—such as fair access to the resource and guaranteed quality of service for individual users. In general, most problems that arise in a Sharing Economy can be cast in a framework where a large number of agents, such as people, cars, or machines, often with unknown objectives, compete for a limited resource. The challenge is to allocate this resource in a manner that is not wasteful, gives an optimal return on the use of the resource for society, and gives a guaranteed level of service to each of the agents competing for that resource. The primary problem in this framework is one of matching supply with demand and doing so in a way that provides adequate stability and resiliency.

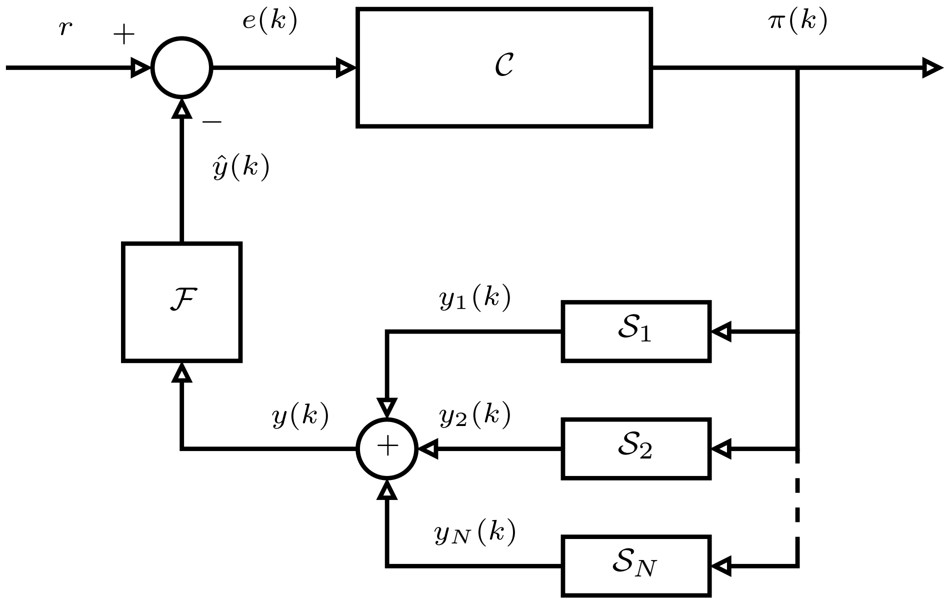


Figure 2.10: Sharing Economy system as a feedback control problem. Individual agents make up the plant and the control allocates the constrained resource among agents [57].

This gives rise to a whole host of problems, which, in principle, are best addressed in a control-theoretic manner. First, we want to fully use the resource, which is a regulation problem. Second, we want to make optimal use of the resource. While both of these objectives are concerned with the aggregate behavior of an agent population, they do not attempt to control how agents individually orchestrate their behavior to achieve this aggregate effect. Therefore, the third objective must focus on the effects of the control on the microscopic properties of the agent population. Ultimately, this third objective can be phrased in terms of properties of the (generally stochastic) process capturing the share of the resource that is allocated to an individual agent. For example, we may wish that each agent, on average, receives a fair share of the resource over time. At a more fundamental level, we wish the average allocation of the resource to each agent over time to be a stable quantity that is entirely predictable, does not depend on initial conditions, and is not sensitive to noise entering the system. From a control perspective, the challenge is to design the feedback mechanism in a way that the stochastic system is ergodic and in a manner in which the invariant measure can be shaped to have desirable properties (see Figure 2.10 for a schematic of the overall feedback control problem).

Regarding the design of the feedback control system, these latter concerns are related to the unique invariant measure (assuming agents behave stochastically) that governs the resource distribution among the agents. Thus, the design of feedback systems for deployment in multi-agent applications must consider not only the traditional notions of regulation and optimization but also the guarantees concerning the existence of this unique invariant measure. This is not a trivial task and many familiar control strategies, even in very simple situations, do not necessarily give rise to feedback systems that possess all three features [57]. This is especially the case when one considers the design of two-sided markets.

Finally, it is important to point out that often non-Sharing Economy agents seek to optimize their individual (selfish) performance with no concern for the aggregate (social) system performance. In other words, agents compete for resources rather than cooperate, as is in a Sharing Economy. It is well known that a cooperative optimum (maximizing a social welfare function) is better than the equilibrium achievable through individual optimality. This difference is reflected through what is known as the *Price of Anarchy*, which can be explicitly quantified in instances where data are readily available, e.g., in transportation systems [58].

2.D.3 Where Can Control Make a Difference?

The Sharing Economy gives rise to many issues that are of deep interest to the systems and control community. After all, sharing is both about regulation of a resource and giving a community (or pool of agents) fair access to this shared resource. In terms of mathematical system construction, several distinct directions of Sharing Economy systems have emerged that are directly relevant to the control and systems community. We list some of these which are perhaps less conventional from a systems and control perspective.

Peer-to-Peer Market Design: This type of Sharing Economy system emerges in the context of prosumer markets where consumers can exchange goods and services in a peer-to-peer context. The design of peer-to-peer markets is a fertile ground to apply ideas from mechanism design, game theory, distributed ledger technology, consensus, distributed optimization algorithms, optimal transport, and the stable marriage problem. Work on the design of such markets will be familiar to most control engineers and is being undertaken by many, especially given the contemporary interest in consensus problems. [59, 60]. A critical issue and frontier topic in the design of peer-to-peer markets concerns the exchange and sharing of data. Data is unlike other shared resources because it is easily copied and replicated. While recent work by eminent researchers [61, 62] has attempted to design such markets, much work remains to be done. This sharing of and shared valuation of data problem is of considerable interest to industry, and is perhaps one of the problems that is currently most pressing. It brings together ideas from peer-to-peer market design [60], distributed ledgers, and economics, like the Shapley value [63].

Transactive Control: This distributed control strategy is another emerging tool that may be used to design shared economies. Because the consumer plays an active role and can carry out decisions that impact the overall Sharing Economy system, the question is about the actual signal from the infrastructure to which the consumer responds. Transactive control is defined broadly as a mechanism through which system and component-level decisions are made through economic transactions negotiated between the components of the system [64]. It is being explored in depth in the context of a smart grid [65, 66] and smart city infrastructures [67, 68]. Given that transactive control provides an opportunity to design a closed-loop system, the specific feedback mechanism of the underlying transactions needs to be suitably designed. It also needs to accommodate both the consumers' behavioral model and the infrastructure managers' economic objectives. A plethora of challenges remains to be addressed, as a result, pertaining to cooperation and coordination of multiple consumers and organizations, dynamic traffic design, constraints on all entities that are involved in the underlying transactions, and dynamic modeling and accommodation of multiple time scales associated with the problem. No matter the market structure used to design the overall shared economy (peer-to-peer, bilateral, auction, or others), the underlying coordination between multiple stakeholders may be through economic transactions, and therefore through transactive control.

Management and Design of Shared Services: The second emergent area very familiar to control engineers concerns products and services that are designed to be shared. Here, the basic idea is to develop both the infrastructure and algorithms to support the sharing of a pool of objects in a community. The most basic elements of designing a shared service are its dimensioning and management. At a very high level, the key questions here concern how many shared items are needed to deliver a certain quality of service (QoS) to the community and/or which users should collaborate to create a sharing system.

On the one hand, problems of this nature often reduce to classical problems in queuing theory which are familiar to the control theory community, but with additional constraints [69, 70]. First, the dimensioning problem is often coupled with the supply-demand statistics, and in some cases, can even influence them. Additional complications arise when sharing systems are themselves coupled.

Further complications are intrinsic to the nature of sharing. For example, a fundamental problem is ensuring that all agents, on average, receive an equal QoS over time. Fairness problems of this nature are common in the networking and AI communities but are rarely discussed in control, despite their obvious connections to consensus and optimization. Further complications emerge as algorithms are developed to manage access to the shared resource, especially when requests for this resource are synchronized over its users. Pricing strategies to manage situations of this nature, especially in the context of polluter-pays (personalized pricing)-type models, are an interesting opportunity for the control community, especially when considering the need for personalized interventions [71].

Management and Design of Social Contracts: The issue of compliance and the design of social contracts using new technology bricks such as digital ledgers, digital companions, and digital twins are central to the design of Sharing Economy systems. Compliance refers to a situation in which we try to encourage (rather than fully control) how agents interact with their environment. For example, in a Sharing Economy application, one might wish to ensure that a shared vehicle is returned to the shared pool at a certain time. Another example of a social contract is the requirement that plastic bottles be returned to the point of sale after use. In Germany and other countries, bottle deposits are used to encourage consumers to adhere to this social contract. However, the minimal deposit is not large enough to encourage good behavior, and the plastic bottles end up in waste bins. This effectively creates a bounty for anybody willing to sort through public rubbish bins to find discarded bottles. So rather than incentivizing responsible behavior, the enforcing this social contract incentivizes others (who often belong to the most vulnerable in society) to sort rubbish. Often, they must collect many bottles just to make minimum wage, and the sorting ecosystem becomes a modern form of exploitation. In the case of bottle deposits, the enforcement mechanism is fragile. It creates a de facto currency that anyone can redeem. It does not place a high enough price on the deposit to incentivize compliance. It also does not apply differential penalties to miscreants depending on levels of personalized compliance.

A further issue in such systems is that they can be attacked by resource-rich nefarious actors. Anyone with enough resources and bad intent could attack the entire system by simply dumping many plastic bottles in the ocean, resulting in an increased deposit for everyone. Game theory has a prominent role to play in developing compliance mechanisms, where uncertainties and variations in the players' preferences, changes in decision-making through observation of other players' behavior, and collective player actions are features that need to be considered. Ultimately, such compliance mechanisms are implemented through the use of specific incentives, to achieve a good equilibrium. Recently, ideas from control theory have emerged as an alternative to these classical mechanisms and offer very interesting directions for the control community [71].

Privacy Preservation: The major impediment in solving the above-mentioned problems is related to the inability to assign deposits to individuals. If one could do this in a privacy-preserving manner, we could create systems in which only those purchasing goods could redeem deposits. We could also create personalized interventions based on differential penalties and be able to price the risk of misbehavior accordingly. However, enforcing social contracts and nudging people to comply with rules for the greater good is a delicate and complex task. In particular, using privacy-invasive mechanisms to micromanage behavior is unacceptable and might lead to dangerous scenarios in which a centralized authority violates the rights of individuals. How is it then possible to mitigate the risks while still maintaining the advantages for the community? One technology that offers great potential in this context is distributed ledger technology, especially when combined with control theory to manage group behavior. Building on these new disruptive technologies to implement secure personalized control strategies may be of great value in several nontraditional application domains for the control community.

Modeling of Human Behavior In a Closed Loop: A significant opportunity for control theory in the context of the Sharing Economy concerns modeling ensembles of human behavior and designing both recommendation platforms for users and algorithms to orchestrate their behavior [57]. The need for suitable microscopic and macroscopic models, of ensembles of human behavior, is essential for control design. Interesting challenges arise in this context as the scale of such ensembles is often large (much larger than typically addressed by the control community), but often not large enough to be treated as a fluid. In many cases, these systems also involve very large populations of humans interacting with the assistance of algorithmic systems. Fundamental questions arising in this context include how to model such systems, understand their equilibrium behavior, and design robust strategies to optimize their fairness and efficiency.

Because these systems are often inherently stochastic, it is important to ensure that they give rise to "good" invariant distributions. This requires distributions that are reasonably fair (e.g., the same user is not always denied access to a bike when there is a shortage of supply at a given location). It is also important to understand how to design control algorithms for such systems. Furthermore, in sharing systems, consumers and producers often make decisions on the basis of predictions and under the influence of software agents. However, such predictions must consider the fact that people will be taking actions based on them, since these actions themselves affect the predictions [72].

Two specific modeling methods for human behavior are worth mentioning here: utility theory [73] and prospect theory [74, 75, 76]. Human decision-making is normally assumed to be based on the optimization of an inherent function of utility. When significant uncertainties are present in the underlying problem (e.g., random shocks), the modeling framework needs to be stochastic and employ distorted perceptions of the underlying outcomes and probabilities. These are the tenets proposed in prospect theory [74] leveraged for modeling empowered consumers in power grids and shared mobility [75, 76].

The challenge of including humans in closed-loop control systems goes beyond the Sharing Economy and comes up in several chapters of this report where societal drivers and specific methodologies are discussed. It remains to be determined whether explicit models of human behavior should be developed, as opposed to learned human behavior characteristics that can be used to design controllers without the need for detailed models.

Access Control and Adversarial Settings: The final, and perhaps the greatest, opportunity for control engineers is connected to the technology platforms that enable much of the Sharing Economy. Technologies such as blockchain or other distributed ledgers enable platforms for peer-to-peer sharing and exchange of resources in a trustworthy manner without the need for a centralized authority. They also enable a host of personalized signaling strategies to achieve regulation of large-scale ensembles of humans. The opportunity for the control community here is immense, both in using and designing such technologies. The use of distributed ledgers in a control context is in its infancy, and much of the technology itself is based on control theoretic concepts (access control). The de facto mechanism for achieving access control in such systems is called proof of work (PoW). This mechanism is now giving way to advancement in FPGAs. Revisiting PoW-type mechanisms and the security of such ledgers from a control perspective may greatly benefit in the design of such systems. Analysis of these systems from a system-theoretic perspective could provide guarantees for use in control design [77, 78, 79].

2.D.4 Long-Term Challenges

In a modern context, sharing is a very disruptive paradigm for businesses across many domains. The impact of the Sharing Economy in the peer-to-peer mobility and accommodation markets and the proliferation of startups focusing on collaborative consumption are testimony to the potential impact that sharing can have on our daily lives. However, despite of the successes, many challenges remain to be addressed as the Sharing Economy matures. A much-discussed side effect of the Sharing Economy is the potential for worker exploitation and poor ethical standards that have emerged due to the so-called *gig economy* [80]. While this is undoubtedly a result of poor behaviors on the part of some Sharing Economy business owners, it is also due to the fact that advances in technology, analytics, and business models have moved at a much faster pace than the legal models providing oversight for Sharing Economy systems. Legal scholars are grappling with new notions of ownership, federated and collective risk and responsibility, insurance, taxation, and the rights of individuals in the context of these new methods of conducting business. Another detrimental aspect of the Sharing Economy comes from the unintended consequences that may emerge. For example, the peer-to-peer accommodation model has increased real estate prices and rents as well as gentrification in large cities worldwide. Other challenges include sharing cyber-physical objects (devices that have both a physical and digital footprint), enabling a machine-to-machine Sharing Economy and enabling a Sharing Economy that involves both humans and machines. Finally, it is worth noting that the Sharing Economy, and the circular economy [81] have much in common. Developing these synergies is also likely to be a fruitful endeavor for the control engineering community.

Recommendations

For young researchers: Engineers are often at the forefront of making more efficient devices but can ignore the stimulus these provide to economic consumption. Future business models will increasingly seek to decouple economic growth from increased consumption. Control theory has much to offer in this context through the design of systems that govern where access to resources in an efficient and just manner. New generations of control engineers must expand the networked multi-agent system setting to capture agent interaction and cooperation in a dynamic resource-sharing environment. The Sharing Economy system poses a feedback control problem where controllers function as intelligent allocators of limited resources among agents that strive to turn what is presently a competitive dynamic game into a dynamic cooperative game, respecting all participants. In this context, new research is needed to explore Sharing Economy systems that address the problem of co-designing technology and human behavior.

For funding agencies: Funding agencies should recognize that a Sharing Economy with well-designed resource management systems is at least one promising way to address our societal sustainability problem of limited resources and the environmental impact of uncontrolled resource utilization. Overarching research goals should focus on turning the proven power of feedback control for dynamic systems into a tool for effective resource management and converting competition into cooperation over the use of resources. This requires a better understanding of how to model human behavior in a closed-loop system environment, study how equilibrium can be attained in such settings, and design robust strategies to optimize fairness and efficiency, which are often in conflict.

2.E Resilience of Societal-Scale Systems

Saurabh Amin, Henrik Sandberg, Jacquelien Scherpen, Dawn Tilbury, Andrew Alleyne, Giacomo Como, Johan Eker, Frederick Leve, Nikolai Matni, George J. Pappas, Thomas Parisini, Emma Tegling

> In this section, we explain how the control systems community is poised to play a key role in building resilience into societal-scale systems. We can contribute to this mission by developing monitoring and decision-making tools for better planning, management, and control of critical infrastructures, and facilitating coordination between various entities that can benefit from these tools.

Abstract Today's societal-scale systems face several risks. These include uneven economic recovery in the aftermath of the pandemic, environmental risks (climate action failure, extreme weather, natural resource crises), technological risks (attacks to strategic infrastructures, loss of trust in digital systems), and societal risks (livelihood crises and public health issues). The control systems science and engineering community can lead and collaborate with government and private sector entities to incorporate resilience in societal-scale systems. We are well-positioned to contribute tools for effective and timely management of disruptions in various sectors such as supply chains, manufacturing, transportation, and energy infrastructures. In doing so, we can leverage emerging regulatory levers and principled data-driven methodologies to limit the negative impact of disruptions. We can also account for the incentives of strategic decision-makers and interdependencies that impact the performance of control loops in networked systems. Specific recommendations include combining prediction methods in climate science and control-theoretic tools for climate change adaptation and mitigation; developing safe and secure technologies for automating and monitoring critical supply chains; leveraging advances in IoT, AI, and cloud computing to increase efficiency and agility of various industry sectors; and improving resilience to both random and adversarial disruptions by developing risk management tools that combine both technological defenses and institutional/regulatory requirements.

2.E.1 Introduction

As governments, the private sector, and communities strive to recover in the aftermath of the COVID-19 pandemic, we are reminded about the global risks that span the societal, economic, environmental, and technological dimensions of human civilization (see Figure 2.11 for examples of 30 global risks in five different categories). In the near term (two to three years), we expect to see uneven recovery trajectories in different sectors of the economy, as evidenced by fluctuating commodity prices, rising inflation, and changing debt levels. In the medium term (three to five years), this divergence is expected to coincide with extreme weather events, geopolitical tensions, and economic confrontations between trading nations. However, in the medium-to-long run (five to 10 years), the most significant risks are perceived to be environmental risks such as climate action failure, extreme weather, biodiversity loss, and natural resource crises and societal risks such as social cohesion erosion, involuntary migration, and livelihood crises. Additionally, technological risks such as cybersecurity failures and inequalities created by digitalization can be potentially significant.

On the one hand, the consequences of the pandemic have complicated the coordination needed to tackle long-term challenges such as climate action, digital safety, and resilience of critical systems. On the other hand, varied responses to public health crises have revealed several valuable insights on the resiliency-improving factors for managing such risks. Risk management should begin with governments balancing the key trade-offs (e.g., public health versus economy), demonstrating preparedness and responsiveness by making data-informed decisions, and facilitating effective communication between public and private sectors. Then, the private sector should contribute significantly to effective and timely responses at multiple levels (both local and national). This can be done by enhancing supply chain resilience, reconfiguring manufacturing to meet critical needs, and implementing a responsible code of conduct within industry. Communities would then play a key role by helping local governments at the grassroots level, raising awareness of systemic vulnerabilities, and encouraging the public to contribute to building resilience.

In this chapter, we explain how the control systems community is poised to play a key role in addressing global risks. These include developing resiliency-improving technological solutions for better planning and operational management of critical systems and facilitating the coordination and data-sharing arrangements needed to implement new policies and regulations.

In the last decade, control systems scientists and engineers have led significant efforts in designing and operating societal-scale CPHS. This has benefited the emerging service models and created new business opportunities in several areas, including data access and analytics, energy management (e.g., building systems control, demand response), automotive safety, metropolitan traffic management, environment and infrastructure monitoring, renewable energy integration, and resilient operation of power grids. However, the overarching goals of ubiquitous connectivity, scalability, reliability, security, adaptability, and flexibility of operation remain hard to achieve from a whole-of-society viewpoint.

To address these challenges, control systems modeling and analysis methods need to be grounded in quantifying the key service requirements. Our methodology for building and designing control must explicitly recognize the systematic vulnerabilities that our societal-scale networks face and effectively leverage emerging regulatory levers and principled data-sharing and analytics principles.

An important cross-cutting challenge is that traditionally, control system design has been considered largely disjointed from the economic incentives of strategic entities such as systems operators, managers, and users. In the context of legacy systems, control engineers have benefited from well-defined ownership structures and at best had a limited interface with regulators and policymakers. However, this loose coupling of control and incentive schemes cannot be sustained under the prevailing trends of digitalization, data-driven analytics, and human-machine interactions. Significantly, new control applications will require faster operational time scales, greater spatial interconnectedness (often crossing multiple borders and jurisdictions), many mixed initiative interactions, and heterogeneous components. It is likely that, without tight coupling of control and incentives, our critical systems will continue to face conflicts between key operational metrics such as performance efficiency versus robustness and faults versus intrusions. They will also face persistent misalignment between individually and socially optimal strategies. It is important to understand how strategic decision-makers' incentives and mechanisms influence the shaping of incentives that account for guarantees (limitations) of control loops and underlying network interdependencies. We believe that advancement in this area will contribute significantly to limiting societal losses in the face of global risks.

2.E.2 Challenges and Opportunities

We now discuss key challenges and opportunities for the control systems community. Focusing on four representative domains and associated risks, we discuss the following questions: How can control systems technology contribute to mitigating and adapting to major global risks? What specific modeling, analysis, and design advancements are needed? What targets should be set over a 5- to 10-year horizon? When appropriate, we also raise the question: Do control systems technology's current/ongoing deployments have unintended consequences that exacerbate these risks?

Infrastructure Failures Due to Climate Change and Extreme Events

Extreme events such as temperature fluctuations, hurricanes and tropical storms, and coastal flooding are increasingly threatening critical infrastructures such as energy and transportation systems. Due to global climate change, their destructive potential is expected to increase further in the coming decades (see Figure 2.12 for an example of synergistic risks that can lead to a global systemic crisis). For example, the risk of critical infrastructure damage due to hurricanes became evident when many energy utilities struggled to handle the aftermath of the 2017 hurricanes Harvey, Irma, and Maria. Although infrastructure systems are inherently sensitive to weather and climate variability, we currently have a very limited understanding of the climate risks faced by these critical systems. We need foundational modeling and control design tools to better plan, monitor, respond, and adapt to these risks.

It is especially important to understand the impact of multidecade climate variability on complex interdependent infrastructure systems. Integrating renewable energy and distributed energy resources (DERs) are increasing the supply-side variability in energy systems. The composition of energy sources is expected to change further as governments and industry chase ambitious decarbonization targets, even as the current transmission and distribution infrastructure remains overstressed (and often ill-managed). In addition, sustained deployment of new electromobility and low-carbon alternative fuels (e.g., biodiesel, hydrogen, non-fossil natural gas) is expected to increase the interdependence among energy, transportation, and manufacturing systems. Anthropogenic activities are also expected to change the distribution and variability of meteorological variables that directly impact both energy demand and supply. The spatiotemporal variability resulting from these trends is expected to evolve over seasons and decades, resulting in new tradeoffs for energy system planning and operations.

The control systems community can lead cross-disciplinary efforts in the climate-infrastructure nexus by combining advanced prediction methods developed by meteorologists and climate scientists with control-theoretic tools for decision-making under uncertainty. Such efforts can be built on the experience in designing operations of complex networks such as the electricity grid and help advance the integration of weather and climate data for assessing and managing climate risks. It is noted that the traditional approach to risk assessment only considers a relatively small set of (average) weather data. This limits our ability to develop risk-aware, long-term infrastructure renewal planning (e.g., transmission expansion and generation capacity expansion) and operational strategy (e.g., unit commitment and economic dispatch models). We need a system-theoretic foundation for infrastructure-climate risk assessment and management grounded in modeling and decision-making in the face of uncertainties.

The above goal can be addressed by developing a data-driven approach to evaluate weather-induced damage in large-scale networks, designing both pre-event resource allocation and post-event response/control strategies for rapid recovery. In particular, control-theoretic tools can be used to estimate spatiotemporal variability of extreme event forcings and predict their impact on infrastructures. We must also use heterogeneous, multimodal data to detect and identify failure events and active response via control tools. This would entail advancing modeling and computational tools for optimal resource allocation and response operations and developing systematic tools for resource allocation, network hardening, microgrid operation, and rapid recovery.

More work is also needed to systematically add high-dimensional time series data from climate models and simulations into coupled infrastructure-climate system dynamics and provide reasonable uncertainty bounds on key outputs and risk indicators. Furthermore, we need to develop useful tools for decision-makers that account for the probabilistic nature of climate risks and how they impact the various stakeholders (generators, transmission operators, utilities, and customers). This entails using ideas from stochastic modeling, dynamic optimization, and network control to limit the impact of climate risks on infrastructure systems. Finally, we need to recognize different ownership structures and strategic interactions between the government and private-sector entities that govern data sharing across sectors/organizations and the individual incentives needed to implement resiliency-improving actions.

Security Threats to Global Networks

Cyberattacks are consistently ranked among the top long-term global risks. They are of particular concern to the control systems community because of the digitalization and deployment of embedded IoT and computing devices in our power grid, vehicles, medical devices, buildings, and many other systems that we routinely interact with. It is well-recognized that such CPS have an enormous positive impact: they improve safety, efficiency, and adaptability and enhance the reliability, usability, and autonomy of emerging service models.

Quantifying and managing risks to CPHS remains a unique challenge, especially in the care of correlated failures caused by simultaneous attacks (security failures) and simultaneous faults (reliability failures). Due to tight cyber-physical interactions and interactions with multiple decision-makers, it is extremely difficult or at least prohibitively time-consuming to isolate the cause of any specific failure using available diagnostic information which, in general, is imperfect and incomplete. Essentially, security and reliability failures in CPHS are inherently intertwined. The control systems community is well-positioned to leverage technological defenses and institutional means to limit the risks due to interdependent reliability and security failures.

The technological means to reduce security and reliability risks include IT security tools like authentication and access control mechanisms, network intrusion detection systems, patch management, and security certification. It also includes control-theoretic tools like model-based and data-driven detection and isolation, robust control strategies that maintain closed-loop stability and performance guarantees under a class of attacks, and reconfigurable control strategies to limit the effect of correlated failures. The control community has recognized that the effectiveness of security tools may be limited by speed, cost, and usability considerations. For example, the frequency of security patch updates is limited by real-time constraints, and common criteria certification is limited by the resources for security. Control researchers have worked with government and industry organizations to develop security standards and recommendations that combine IT-specific security defenses with control-theoretic tools.

A significant challenge for implementing a security and reliability risk assessment framework is developing data-driven, stochastic CPHS models, which account for dynamics under interdependent reliability and security failures. Here the key challenge is to extend the basic models that focus on dynamics under a specific scenario to a composite model that captures various correlated failure scenarios, including simultaneous attacks, common-mode failures, and cascading failures. In principle, techniques from statistical estimation, model-based diagnosis, stochastic simulation, and predictive control can be used to generate new failure scenarios from real-time data, synthesize operational security strategies, and provide estimates of residual risks in environments with highly correlated failures and less-than-perfect information.

It is important to note that technology-based defenses alone are insufficient to ensure resilience to correlated failures. In particular, the lack of private parties' incentives for security improvements severely impedes achieving socially desirable improvements of CPHS security. This problem is compounded by information deficiencies arising from the conflicting interests of individual players, whose choices affect CPHS risks. In environments with incomplete and asymmetric information, the societal costs of a correlated failure typically exceed the losses of the individual players whose products and services affect systems operations and on whose actions the risks depend. The individual players are then likely to underinvest in security relative to a socially optimal benchmark, requiring the design of institutional means to realign the individual incentives to make adequate investments. Still, our limits to accessing risks of statistically rare events such as extremely correlated failures mean that our critical societal infrastructures face residual risks even under best technological defenses and institutional structures. The control system community will need to design and operate systems in the face of these residual risks. This is apparent from the emergence of organized cybercrime groups capable of conducting intrusions into critical control systems and cyber warfare actively employed by nations engaged in armed conflict.

Supply Chain Disruptions

Today's supply chains are complex, interconnected, and global networks. They include producers, transportation and logistics companies, and entities for the storage, assembly, and distribution of a large number of products and differentiated services. In recent years, supply chains have become globalized. This has improved productivity and income levels in producer countries (in particular, low-income countries). The ongoing digital transformation of supply chains has contributed significantly to increased levels of automation, cost control, and service quality. It has also increased the efficiency and timeliness of transactions while shifting the role of humans in manufacturing plants, warehouses, and logistics operations.

However, it is also clear that many supply chains have become more vulnerable to disruptions. The global pandemic has revealed issues such as recurring shortages of critical parts and raw materials (e.g., semiconductor chip shortages impacting manufacturing in many sectors) and rapidly evolving consumer buying patterns.

It is important to note that the movement toward fragile structures has been in the making for several decades. For example, many firms reduced capacity buffers and experienced personnel to meet consumer expectations of low prices. To cut costs and gain a competitive edge in tough markets, we have seen significant levels of outsourcing and offshoring. This has significantly increased reliance on developing economies, with China emerging as a major manufacturing hub. These are only some of the reasons behind well-known disruptions such as the 2011 Tsunami in Japan, the 2008 financial crisis, and the 2003 SARS epidemic.

The control systems community is well-positioned to improve the resilience of global supply chains in response to events such as natural disasters and weather emergencies (which are expected to rise with climate change), cyberattacks, labor shortages/strikes, and financial crises. Here are some of the opportunities:

End-to-End Visibility In Diverse Supply Chain Networks: Control engineers have already contributed to the deployment of IoT, robotics, and automation in various production and logistics processes. By building on the experience in designing large-scale monitoring systems, our community can provide solutions to ensure end-to-end visibility in diverse supply chain networks ranging from vertically integrated industries to ones with varying levels of offshoring, outsourcing, and collaborative relationships. The challenge here is whether we can design monitoring systems that provide fine-grained alerts about issues such as unforeseen delays or quality noncompliance.

Identify Disturbance Propagation and Weak Links: Control systems thinking can also provide new insights into how disturbances propagate in these complex networks and help stakeholders identify weak links in a timely manner. This is especially important given that both the magnitude and frequency of climate-related events continue to rise, and many of these events are causing supply-chain shocks. A significant step would be to leverage improved visibility of supply chain structures to identify if suppliers can meet the demand for their current production plans and inventory levels. For example, work by David Simchi-Levi has shown that metrics such as “time-to-survive” and “time-to-recover” can be very useful in evaluating resilience.

Building Redundancy Measures: Tools from the control of distributed systems, stochastic optimization, and recent developments in learning for control can provide useful insights on where and how to build redundancy measures like backup capacity and strategic holding of additional inventory. It is important to note that building redundancy is costly in general, but it may provide a competitive advantage and increase profits during periods of supply chain stress. The control systems community has long-standing experience designing and implementing redundancy measures in the power grid. There are interesting parallels and lessons to be learned that can be applied to supply chains.

Improve Supply Chain Agility: Maintaining agility in supply chains will enable them to respond quickly and substitute between inputs to respond to disruptions. Firms can improve agility in various ways including identifying bottlenecks and resolving them promptly, investing in flexible production processes and general-purpose manufacturing lines and maintaining long-term collaborative relationships with suppliers. All these solutions require decision-making under uncertainty and the ability to quantify key underlying tradeoffs.

Engage With the Government and Public Sector: The control systems community can potentially engage in the conversation around the role of the government and public sector in promoting resilience. Importantly, the public sector plays a crucial role in maintaining investment levels in sectors including energy production, climate security, and transportation of people, critical raw material, and supplies at national and international scale. During the COVID-19 pandemic, the diversity of government responses worldwide in aggregating and disseminating information and ensuring supplies of critical commodities underscored the importance of timely and coordinated interventions. Similar opportunities lie in providing crucial aggregate information to help effective planning and coordination between various entities running global supply chains.

2.E.3 Recommendations

Network analysis – structure and dynamics in the face of risks

- Model the dynamic propagation of disturbances in cyber-physical networks and use of multimodal data to learn network structure and dynamics (near-term)
- Develop quantitative methods for risk assessment, quantify risk metrics, and exploit data as well as models. Develop general methodology and infrastructure-specific analysis that accounts for common-mode failures and interactions between infrastructures (near-term)
- Learn about rare events, their risks, and ways to model and block cascading risks and failures (medium-term)
- Minimize risks to network structure design subject to budget, incentive, and infrastructure interaction constraints. This may involve the allocation of new devices and services and partial infrastructure upgrades (medium-term)
- Develop resilient control through coordination between multiple interacting subsystems potentially operated by decision-makers that have their own objectives and imperfect information (medium-term)
- Advance technology to improve the trustworthiness of critical supply chains, including integrity, traceability, and resilience to disruptions (medium-term)
- Develop foundational theory to integrate methods in analysis of large-scale complex networks with methods in strategic planning and operational response design in the face of global risks (long-term)
- Develop foundational knowledge of computational methods that address the interaction of global supply chains (including the energy system) with changing weather and climate extremes (long-term)

Understanding ownership structures and shaping incentives in networks

- Make connections between distributed control and dynamic mechanism design (near-term)
- Understand how to jointly design incentives and control to take advantage of opportunity at the intersection of game-theoretic modeling and learning in control with applications of networks (e.g., strategic network formation, learning and adaptation in networks with uncertain parameters, and network interdependencies) (near- to mid-term)
- Account for differences in regulatory regimes and drive them in a manner which facilitates the scale and robustness of deployed technology (mid-term)
- Learn how to build a network and control loops to mitigate unintended consequences. (mid-term)
- Understand how to control only some components of the network and how to integrate different components (mid-term)
- Define interfaces between components and understand what kind of regulation is needed for a trustworthy design (mid-term)
- Analyze security IT and control-theoretic defenses and policy mechanisms for risk management in CPHS under alternative informational and incentive structures (long-term)
- Establish formal channels of interaction between entities responsible for operating and managing critical systems, even when they are located in different jurisdictions to coordinate effective responses to disruptive events (long-term)

- Understand how institutions can enable new business cases for resilience and new ways of risk-sharing, like maintaining minimum levels of legal liability of secure CPHS, and incentivize vulnerability-finding, incident disclosure, and insurance instruments to improve information about risks (long-term)

Monitoring, early assessment of disruptions, and quantifying predictions

- Learn latent network state and rapidly identify correlated events based on heterogeneous data (near-term)
- Undertake data-driven approaches to leverage and create new datasets and intersections with model-driven frameworks (near-term)
- Understand the actuator, where control comes in, and opportunities for behavioral model analysis (near-term)
- Predict massive disruptions in global networks/supply chains and understand the limitations and impossibilities of those predictions. Leverage statistical machine learning and a mixture of worst-case deterministic modeling and detection of onset and failure-triggering events (medium-term)
- Successfully deploy computational tools for analytics-driven diagnostic models from heterogeneous multisource data on disruptions (medium-term)
- Understand ownership structures of data needed for diagnostics and sharing arrangements needed for ensuring the trustworthiness of data used for generating alerts (medium-term)
- Explore digital commodities, the flow of information and goods, security monitoring of computing and communication requirements, cooperation between unmanned aerial vehicle (UAV) fleets, and the risks of relying on cloud services hosted in a single country (medium-term)
- Integrate diagnostic tools that leverage data on past network disruptions and early warning indicators with the design of optimal resource allocation and network control strategies (long-term)
- Actively learn before and during disruptive events and accurately predict their impact on critical networks and services (long-term)

Illegal, criminal, adversarial actions on the networks

- Build on significant work in cyber secure and resilient control, and expand scope to interdependent systems (near-term)
- Use security allocation (game-theoretic) to disrupt illegal activities (near-term)
- Design honeypots in CPHS (near-term)
- Include critical supply chains to improve trustworthiness (medium-term)
- Disrupt illegal supply chains (human trafficking, natural resources, rare minerals, etc.) and understand the role of monitoring and inspection technology and supply chain network controls to eliminate illegal flows (medium-term)
- Apply foundational knowledge and tools to understand the role of humans in initiating illegal, criminal, and adversarial behaviors in CPHS (long-term)

- Evaluate the role of AI/ML techniques in creating new attack vectors and how regulatory and incentive frameworks can be created to limit their effectiveness. In addition, leverage national policies and international commitments on responsible use of digital technologies deployed in critical systems (long-term)
- Understand to what extent control systems designers and operators need to be trained to recognize such behaviors (long-term)

Reconfiguration and adaptation to build-in resilience

- Determine where to put buffers back into the system and memory, who pays and what the price is, and how to localize failures (e.g., If one bank goes bankrupt, the rest of the world does not follow) (near-term)
- Think about the tradeoff between efficiency and resilience: Is there a Pareto curve? Should providers, consumers, or governments pay for resilience? (near-term)
- Understand how recent examples impacted manufacturing and supply chains. For example, how the global chip shortage affected new and used car markets, how these markets adapted, and how demand shifted (near-term)
- Consider buffers or reconfiguration and the incentive for private businesses to have buffers. When a company provides buffers, lean or just-in-time (JIT) companies must pay a lot when a disruption occurs (medium-term)
- Define issues around adapting manufacturing and logistics to implement reuse and recycling (circular economy), such as tools, disposal costs and institutional structures, and incentives for manufacturers to send safe and secure products to market or recall/fix them when needed (medium-term)
- Evaluate and plan a global response to cyber conflicts involving nation-states, including situations in which cyber capabilities are deployed as a means of subversion and which create humanitarian distress (long-term)
- Understand why scale is a challenge. Determine how engineered systems like the power grid (e.g., deployment of storage in a renewable-heavy grid and decarbonization strategies in multi-vector energy systems planning) and be used in mechanism/incentive design and regulation to achieve scale (long-term)
- Orient control technology to create a business case for risk management (long-term)
- Understand when to convene a temporary strategic shortening of the global economy or wait for a crash (long-term)

Recommendations

For young researchers: To build resiliency in the next-generation societal infrastructure, we need foundational tools to understand network dynamics under disruptions and detect/identify failures; design control and incentive mechanisms that improve closed-loop performance in strategic environments; and implement optimal resource allocation and dynamic response strategies.

For funding agencies: Areas that are nascent or require more exploration include comprehensive and contextual design approaches to ensure the safety, trust, and security of critical systems in the face of emerging risks (specifically, technological, environmental, and societal risks) and robust, data-driven mechanisms to maintain functionality in a wide range of disruptions to networked systems.

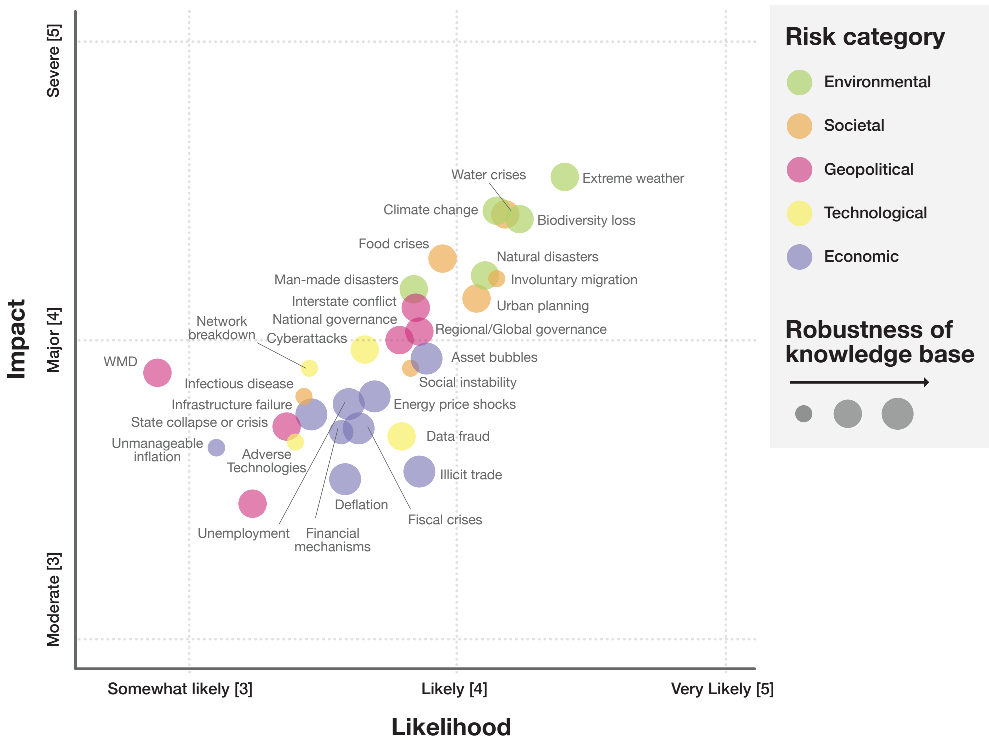


Figure 2.11: Mean ranked likelihood and impact of global risks, plus robustness of the knowledge base surrounding each risk (size of the circle), for the 30 global risks in 5 categories (risk types). Citation: Future Earth, 2020. Risks Perceptions Report 2020: First Edition. Future Earth [82] (reprinted with permission).

BIBLIOGRAPHY

[1] Intergovernmental Panel on Climate Change (IPCC), “Climate change 2021: The physical science basis, summary for policymakers,” https://www.ipcc.ch/report/ar6/wg1 (accessed August 26, 2021), August 2021.

[2] International Energy Agency (IEA), “Net zero by 2050: A road map for the global energy sector,” https://www.iea.org/reports/net-zero-by-2050 (accessed September 1, 2021), May 2021.

[3] Our World in Data, “Global greenhouse gas emissions by sector,” https://ourworldindata.org/emissions-by-sector (accessed August 26, 2021), 2020.

[4] Hydrogen Council, “Hydrogen for net zero,” https://hydrogencouncil.com/en/hydrogen-for-net-zero/ (accessed August 9, 2022), 2021.

[5] R. G. Gonzalez, “Towards 2030,” *Decarbonisation Technology*, 2021. https: //decarbonisationtechnology.com/article/34/towards-2030#.YzGxJHZBybg

[6] R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, “Green AI,” *Commun. ACM*, vol. 63, no. 12, pp. 54–63, nov 2020.

[7] OpenAI, “AI and Compute,” https://openai.com/blog/ai-and-compute/ (accessed August 9, 2022), 2019.

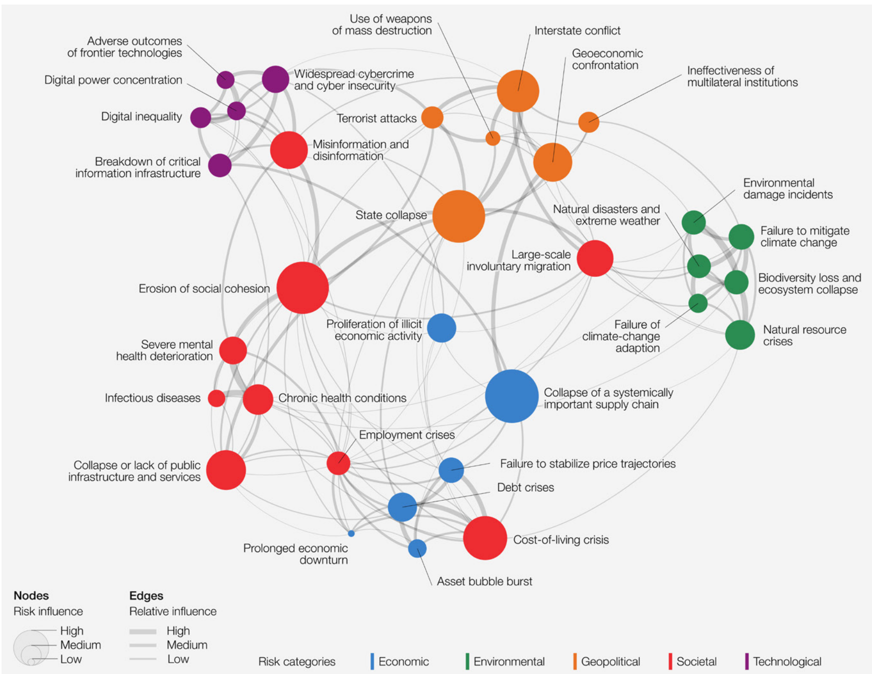


Figure 2.12: Interconnections between global risks. Source: The Global Risk Report 2023 from the World Economic Forum. The report identifies environment issues such as failure to mitigate and adapt to climate change, natural disasters, and extreme weather events among top 3 in terms of likely impact over a 10 year period. Risk interconnections imply that disparate failure events can interact such that their overall impact far exceeds the sum of each part. On a positive note, however, interconnections can also broaden the impact of risk mitigation efforts – increasing resilience in one sector can lead to improved preparedness for other risks [83] (reprinted with permission).

[8] Deloitte, "2022 Global Health Care Outlook," 2022.

[9] Population Reference Bureau, "Fact sheet: Aging in the United States," 2022. https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/

[10] The Commonwealth Fund, "U.S. Health Care from a Global Perspective, 2019: Higher Spending, Worse Outcomes?" 2022. https://www.commonwealthfund.org/publications/issue-briefs/2020/jan/us-health-care-global-perspective-2019

[11] Australian Medical Association Limited, "Climate change is a health emergency," September 2019. https://www.ama.com.au/media/climate-change-health-emergency

[12] Natural Resources Defense Council, "Report: Health costs from climate change and fossil fuel pollution tops $820 billion a year," May 2021. https://www.nrdc.org/media/2021/210520

[13] World Economic Forum, "Here are five priorities for better cooperation on mental health," Nov 2018. https://www.weforum.org/agenda/2018/11/five-ways-mental-health-care-better-depression/

[14] R. Bianucci, O. J. Benedictow, G. Fornaciari, and V. Giuffra, "Quinto Tiberio Angelerio and new measures for controlling plague in 16th-century Alghero, Sardinia," Emerging infectious diseases, vol. 19, no. 9, pp. 1478–1483, 2013. https://pubmed.ncbi.nlm.nih.gov/23968598

[15] Z. L. Grange *et al.*, "Ranking the risk of animal-to-human spillover for newly discovered viruses," Proceedings of the National Academy of Sciences, vol. 118, no. 15, p. e2002324118, 2021. https://www.pnas.org/doi/abs/10.1073/pnas.2002324118

[16] D. Vogel, Medical Device Software Verification, Validation and Compliance. Artech House, 2011. https://books.google.com/books?id=LYxH-zUSOTgC

[17] B. Sutton, "Overview of regulatory requirements: Medical devices-transcript," 2011. https://www.fda.gov/training-and-continuing-education/cdrhlearn/overview-regulatory-requirements-medical-devices-transcript

[18] FDA, "Software as a Medical Device (SAMD)," 2022. https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd

[19] FDA, "Software as a Medical Device (SAMD): Clinical Evaluation Guidance for Industry and Food and Drug Administration Staff," 2017. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/software-medical-device-samd-clinical-evaluation

[20] FDA, "The Content of Investigational Device Exemption (IDE) and Premarket Approval (PMA) Applications for Artificial Pancreas Device Systems Guidance for Industry and Food and Drug Administration Staff," 2012. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/content-investigational-device-exemption-ide-and-premarket-approval-pma-applications-artificial

[21] FDA, "Investigational Device Exemptions (IDEs) for Early Feasibility Medical Device Clinical Studies, Including Certain First in Human (FIH) Studies Guidance for Industry and Food and Drug Administration Staff," 2013. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/investigational-device-exemptions-ides-early-feasibility-medical-device-clinical-studies-including

[22] FDA, “Policy for Device Software Functions and Mobile Medical Applications Guidance for Industry and Food and Drug Administration Staff,” 2019. https://www.fda.gov/regulatory-information/search-fda-guidancedocuments/policy-device-software-functions-and-mobile-medical-applications

[23] A. Chakrabarty, S. Zavitsanou, T. Sowrirajan, F. J. Doyle III, and E. Dassau, Chapter 2 - Getting IoT-ready: The face of next generation artificial pancreas systems. Academic Press, 2019, pp. 29–57. https://www.sciencedirect.com/science/article/pii/ B9780128156551000119

[24] S. Zavitsanou, A. Chakrabarty, E. Dassau, and F. J. Doyle III, “Embedded control in wearable medical devices: Application to the artificial pancreas,” Processes, vol. 4, no. 4, p. 35, 2016. https://www.mdpi.com/2227-9717/4/4/35

[25] M. B. Gillingham *et al.*, “Assessing mealtime macronutrient content: Patient perceptions versus expert analyses via a novel phone app,” *Diabetes Technology & Therapeutics*, vol. 23, no. 2, pp. 85–94, 2020. https://doi.org/10.1089/dia.2020.0357

[26] N. C. Foster *et al.*, “State of Type 1 Diabetes Management and Outcomes from the T1D Exchange in 2016-2018,” *Diabetes Technol Ther*, vol. 21, no. 2, pp. 66–72, 2019. https://www.ncbi.nlm.nih.gov/pubmed/30657336

[27] R. Nimri, A. R. Ochs, J. E. Pinsker, M. Phillip, and E. Dassau, “Decision support systems and closed loop,” *Diabetes Technol Ther*, vol. 21, no. S1, pp. S42–S56, 2019. https://www.ncbi.nlm.nih.gov/pubmed/30785322

[28] E. Dassau *et al.*, “Twelve-Week 24/7 Ambulatory Artificial Pancreas With Weekly Adaptation of Insulin Delivery Settings: Effect on Hemoglobin A1c and Hypoglycemia,” *Diabetes Care*, vol. 40, no. 12, pp. 1719–1726, 2017. https://www.ncbi.nlm.nih.gov/pubmed/29030383

[29] A. O. Hebb *et al.*, “Creating the feedback loop: closed-loop neurostimulation,” *Neurosurg Clin N Am*, vol. 25, no. 1, pp. 187–204, 2014. https://www.ncbi.nlm.nih.gov/pubmed/24262909

[30] M. Parastarfeizabadi and A. Z. Kouzani, “Advances in closed-loop deep brain stimulation devices,” *J Neuroeng Rehabil*, vol. 14,no. 1, p. 79, 2017. https://www.ncbi.nlm.nih.gov/pubmed/28800738

[31] J. E. Fleming, J. Orlowski, M. M. Lowery, and A. Chaillet, “Self-tuning deep brain stimulation controller for suppression of beta oscillations: Analytical derivation and numerical validation,” *Front Neurosci*, vol. 14, p. 639, 2020. https://www.ncbi.nlm.nih.gov/pubmed/32694975

[32] S. Bose, “Early prediction of adverse clinical events and optimal intervention in intensive care units,” Ph.D. dissertation, Johns Hopkins University, 2021.

[33] G. A. Dumont and J. M. Ansermino, “Closed-loop control of anesthesia: a primer for anesthesiologists,” *Anesth Analg*, vol. 117, no. 5, pp. 1130–8, 2013. https://www.ncbi.nlm.nih.gov/pubmed/23835456

[34] A. Joosten *et al.*, “Anesthetic management using multiple closed-loop systems and delayed neurocognitive recovery: A randomized controlled trial,” *Anesthesiology*, vol. 132, no. 2, pp. 253–266, 2020. https://doi.org/10.1097/ALN.0000000000003014

[35] N. West *et al.*, “Design and evaluation of a closed-loop anesthesia system with robust control and safety system,” *Anesth Analg*, vol. 127, no. 4, pp. 883–894, 2018. https://www.ncbi.nlm.nih.gov/pubmed/29210791

[36] K. van Heusden *et al.*, “Optimizing robust pid control of propofol anesthesia for children: Design and clinical evaluation,” *IEEE Trans Biomed Eng*, vol. 66, no. 10, pp. 2918–2923, 2019. https://www.ncbi.nlm.nih.gov/pubmed/30763237

[37] J. Han and D. R. Trumble, “Cardiac assist devices: Early concepts, current technologies, and future innovations,” *Bioengineering (Basel)*, vol. 6, no. 1, 2019. https://www.ncbi.nlm.nih.gov/pubmed/30781387

[38] ABIOMED, “Impella© 5.5© with smartassist.” https://www.fda.gov/media/140766/download

[39] F. J. Doyle III, L. M. Huyett, J. B. Lee, H. C. Zisser, and E. Dassau, “Closed-loop artificial pancreas systems: engineering the algorithms,” *Diabetes Care*, vol. 37, no. 5, pp. 1191–7, 2014. https://www.ncbi.nlm.nih.gov/pubmed/24757226

[40] S. Tagougui *et al.*, “Artificial pancreas systems and physical activity in patients with type 1 diabetes: Challenges, adopted approaches, and future perspectives,” *J Diabetes Sci Technol*, vol. 13, no. 6, pp. 1077–1090, 2019. https://www.ncbi.nlm.nih.gov/pubmed/31409125

[41] E. Dassau *et al.*, “Intraperitoneal insulin delivery provides superior glycaemic regulation to subcutaneous insulin delivery in model predictive control-based fully-automated artificial pancreas in patients with type 1 diabetes: a pilot study,” *Diabetes Obes Metab*, vol. 19, no. 12, pp. 1698–1705, 2017. https://www.ncbi.nlm.nih.gov/pubmed/28474383

[42] K. L. Wolkowicz *et al.*, “A review of biomarkers in the context of type 1 diabetes: Biological sensing for enhanced glucose control,” *Bioeng Transl Med*, vol. 6, no. 2, p. e10201, 2021. https://www.ncbi.nlm.nih.gov/pubmed/34027090

[43] A. Haidar, “Insulin-and-glucagon artificial pancreas versus insulin-alone artificial pancreas: A short review,” *Diabetes Spectr*, vol. 32, no. 3, pp. 215–221, 2019. https://www.ncbi.nlm.nih.gov/pubmed/31462876

[44] A. Haidar *et al.*, “A novel dual-hormone insulin-and-pramlintide artificial pancreas for type 1 diabetes: A randomized controlled crossover trial,” *Diabetes Care*, vol. 43, no. 3, pp. 597–606, 2020. https://www.ncbi.nlm.nih.gov/pubmed/31974099

[45] UN General Assembly, “Transforming our world: the 2030 agenda for sustainable development,” A/RES/70/1, 2015

[46] E. Crisostomi *et al.*, Eds., *Analytics for the Sharing Economy: Mathematics, Engineering and Business Perspectives*. Springer, 2020.

[47] M. Hu, Ed., *Sharing Economy-Making Supply Meet Demand*. Springer, 2020.

[48] Y. Par and B. Sovacool, “Electricity market design for the prosumer era,” *Nature Energy*, 2016.

[49] R. Botsman and R. Rogers, “The rise of collaborative consumption,” *Harper Collin*s, vol. 2010.

[50] A. Sundararajan, *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. MIT Press, 2016.

[51] European Commission Press Release, “A european agenda for the collaborative economy,” June 2016. https://ec.europa.eu/commission/presscorner/detail/en/IP_16_2001

[52] R. Zafar *et al.*, “Prosumer based energy management and sharing in smart grid,” *Renewable and Sustainable Energy Reviews*, vol. 82, no. 1, pp. 1675–1684, 2018.

[53] D. Menniti, A. Pinnarelli, N. Sorrentino, and G. Belli, “A local market model involving prosumers taking into account distribution network congestions in smart cities,” *International Review of Electrical Engineering (IREE)*, vol. 9, no. 5, pp. 976–985, 2014.

[54] B. Alcott, “Jevons paradox,” *Ecological Economics*, vol. 54, no. 1, 2005.

[55] A. Schlote *et al.*, “Cooperative regulation and trading of emissions using plug-in hybrid vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, 2013.

[56] P. Katsikouli *et al.*, “Distributed ledger enabled control of tyre induced particulate matter in smart cities,” *Frontiers in Sustainable Cities*, vol. 2, 2019.

[57] A. Fioravanti, J. Marecek, R. Shorten, M. Souza, and F. Wirth, “On the ergodic control of ensembles,” *Automatica*, vol. 108, 2019.

[58] J. Zhang, S. Pourazarm, C. G. Cassandras, and I. C. Paschalidis, “The price of anarchy in transportation networks: Data-driven evaluation and reduction strategies,” *Proceedings of the IEEE*, vol. 106, no. 4, pp. 538–553, 2018.

[59] J. I. Nielsen, “A small community is pushing the agenda for a greener future.” https://www.cee.elektro.dtu.dk/news/2018/06/a-small-community-is-pushing-the-agenda-for-a-greener-future

[60] T. Sousa *et al.*, “Peer-to-peer and community-based markets: A comprehensive review,” *Renewable and Sustainable Energy Reviews*, vol. 104, 2019.

[61] A. Agarwal, M. Dahleh, and T. Sarkar, “Marketplace for data: An algorithmic solution,” in *Proceedings of the 2019 ACM Conference on Economics and Computation*, vol. 2019.

[62] M. Rasouli and M. Jordan, “Data sharing markets,” 2021. https://arxiv.org/abs/2107.08630

[63] R. Stanojevic, N. Laoutaris, and P. Rodriguez, “On economic heavy hitters: shapley value analysis of 95th-percentile pricing,” in I*MC ’10: Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 2010.

[64] A. M. Annaswamy and T. Nudell, “Transactive control–what’s in a name,” *IEEE Smart Grid Newsletter*, 2015.

[65] J. Hu, G. Yang, K. Kok, Y. Xue, and H. W. Bindner, “Transactive control: a framework for operating power systems characterized by high penetration of distributed energy resources,” *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 3, pp. 451– 464, 2017.

[66] A. Kiani Bejestani, A. M. Annaswamy, and T. Samad, “A hierarchical transactive control architecture for renewables integration in smart grids: Analytical modeling and stability,” *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2054–2065, 2014.

[67] A. M. Annaswamy *et al.*, "Transactive control in smart cities," *Proceedings of the IEEE*, vol. 106, no. 4, pp. 518–537, 2018.

[68] R. Sengupta, S. Amin, A. M. Annaswamy, S. Moura, and V. Bulusu, "Smart cities and control [technical activities]," *IEEE Control Systems Magazine*, vol. 35, no. 6, pp. 20– 21, 2015.

[69] C. King, W. Griggs, F. Wirth, K. Quinn, and R. Shorten, "Alleviating a form of electric vehicle range anxiety through on-demand vehicle access," *International Journal of Control*, vol. 88, no. 4, 2015.

[70] W. Griggs, J. Yu, F. Wirth, F. Häusler, and R. Shorten, "On the design of campus parking systems with qos guarantees," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, 2016.

[71] P. Ferraro, L. Zhao, C. King, and R. Shorten, "Personalised feedback control and social contracts and compliance strategies for ensembles," 2021. https://arxiv.org/abs/2103.07261

[72] J. Epperlein, S. Zhuk, and R. Shorten, "Recovering Markov models from closed-loop data," *Automatica*, vol. 103, 2019.

[73] K. J. Lancaster, "A new approach to consumer theory," *Journal of political econom*y, vol. 74, no. 2, pp. 132–157, 1966.

[74] A. Tversky and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.

[75] K. Jhala, B. Natarajan, and A. Pahwa, "Prospect theory-based active consumer behavior under variable electricity pricing," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2809–2819, 2018.

[76] Y. Guan, A. M. Annaswamy, and H. Eric Tseng, "Cumulative prospect theory based dynamic pricing for shared mobility on demand services," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 2239–2244.

[77] P. Ferraro, C. King, and R. Shorten, *Distributed ledger technology for smart cities, the Sharing Economy, and social compliance*. IEEE Access, 2018.

[78] P. Ferraro, C. King, and R. Shorten, "On the stability of unverified transactions in a dag-based distributed ledger," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, 2020.

[79] A. Cullen, P. Ferraro, W. Sanders, L. Vigneri, and R. Shorten, "Access control for distributed ledgers in the internet of things: A networking approach," *IEEE Internet of Things Journal*, 2022, accepted.

[80] R. Kwok, "Flexible working: Science in the gig economy," *Nature*, vol. 550, 2017.

[81] Ellen MacArthur Foundation, "What is a circular economy?" https://ellenmacarthurfoundation.org/topics/circular-economy-introduction/overview

[82] Future Earth, *Future Earth Risk Perceptions Report 2020,* 1st ed., 2020.

[83] World Economic Forum. *The Global Risks Report 2023*, 18th ed., 2023.

[84] M. Martcheva, *An Introduction to Mathematical Epidemiology*, ser. Texts in Applied Mathematics. Springer, Boston, MA, 2015, vol. 16.

[85] Y. Cui, S. Ni, and S. Shen, “A network-based model to explore the role of testing in the epidemiological control of the COVID-19 pandemic,” *BMC Infectious Diseases*, vol. 21, no. 58, 2021. https://doi.org/10.1186/s12879-020-05750-9

[86] R. Pastor-Satorras and A. Vespignani, “Epidemic dynamics and endemic states in complex networks,” *Phys. Rev. E*, vol. 63, p. 066117, May 2001. https://link.aps.org/doi/10.1103/PhysRevE.63.066117

[87] G. G. Katul, A. Mrad, S. Bonetti, G. Manoli, and A. J. Parolari, “Global convergence of COVID-19 basic reproduction number and estimation from early-time SIR dynamics,” *PLOS ONE*, vol. 15, no. 9, pp. 1–22, 2020. https://doi.org/10.1371/journal.pone.0239800

[88] E. H. Bussell, C. E. Dangerfield, C. A. Gilligan, and N. J. Cunniffe, “Applying optimal control theory to complex epidemiological models to inform real-world disease management,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 374, 2019.

[89] N. Ferguson *et al.*, “Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand,” 2020. http://hdl.handle.net/10044/1/77482

[90] J. Ashton, “COVID-19 and the ‘Spanish’ flu,” *Journal of the Royal Society of Medicine*, vol. 113, no. 5, pp. 197–198, 2020, pMID: 32407643. https://doi.org/10.1177/0141076820924241

[91] O. Sharomi and T. Malik, “Optimal control in epidemiology,” *Annals of Operations Research*, vol. 251, no. 1, pp. 55–71, April 2017.

[92] S. C. Fay, D. J. Jones, M. A. Dahleh, and A. E. Hosoi, “Simple control for complex pandemics,” 2020. https://arxiv.org/abs/2012.08755

[93] N. R. Initiative, “When to test,” 2020, accessed on December 15, 2020. https://whentotest.org/

[94] A. Sarker, “Mixtures matter,” 2020. https://idss.mit.edu/vignette/mixtures-matter/

[95] J. Page, M. P. Brenner, and R. R. Kerswell, “Revealing the state space of turbulence using machine learning,” *Phys. Rev. Fluids*, vol. 6, p. 034402, Mar 2021. https://link.aps.org/doi/10.1103/PhysRevFluids.6.034402

[96] X. F. Meng, D. Jones, R. Rigobon, and M. Dahleh, “Impacts of COVID-19 interventions: Health, economics, and inequality,” 2021. https://doi.org/10.21203/rs.3.rs-406480/v1

[97] Y. Liu and J. Rocklöv, “The reproductive number of the delta variant of SARS-COV-2 is far higher compared to the ancestral SARS-COV-2 virus,” *Journal of Travel Medicine*, vol. 28, no. 7, 2021.

[98] G. Chowell *et al.*, “Model parameters and outbreak control for sars,” *Emerging Infectious Diseases*, vol. 10, no. 7, p. 1258–1263, 2004.

[99] L. F. White *et al.*, “Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA,” *Influenza and Other Respiratory Viruses*, vol. 3, no. 6, p. 267–276, 2009.

[100] O. Moussa Tessa, “Mathematical model for control of measles by vaccination,” in *Proceedings of Mali Symposium on Applied Sciences*, 08 2006, pp. 31–36.

[101] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of Brain Network Analysis*. Academic Press, 2016.

[102] D. S. Bassett and O. Sporns, “Network neuroscience,” *Nature Neuroscience*, vol. 20, no. 3, pp. 353–364, 2017.

[103] H. Berger, “Über das elektrenkephalogramm des menschen,” *Archiv für psychiatrie und nervenkrankheiten*, vol. 87, no. 1, pp. 527–570, 1929.

[104] A. Zalesky, A. Fornito, and E. Bullmore, “On the use of correlation as a measure of network connectivity,” *Neuroimage*, vol. 60, no. 4, pp. 2096–2106, 2012.

[105] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Kötter, “Key role of coupling, delay, and noise in resting brain fluctuations,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 25, pp. 10302–10307, 2009.

[106] F. Váša *et al.*, “Effects of lesions on synchrony and metastability in cortical networks,” *Neuroimage*, vol. 118, pp. 456–467, 2015.

[107] J. Stiso et al., “White matter network architecture guides direct electrical stimulation through optimal state transitions,” *Cell Reports*, vol. 28, no. 10, pp. 2554 – 2566.e7, 2019.

[108] U. Braun et al., “Brain network dynamics during working memory are modulated by dopamine and diminished in schizophrenia,” *Nature Communications*, no. 1, p. 3478, 2021.

[109] T. M. Karrer et al., “A practical guide to methodological considerations in the controllability of structural brain networks,” *Journal of Neural Engineering*, vol. 17, no. 026031, 2020.

[110] E. Nozari et al., “Is the brain macroscopically linear? a system identification of resting state dynamics,” *bioRxiv*, 2020.

[111] M. P. Van Den Heuvel and H. E. Hulshoff Pol, “Exploring the brain network: a review on resting-state fmri functional connectivity,” *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519–534, 2010.

[112] A. Ponce-Alvarez et al., “Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity,” *PLoS Computational Biology*, vol. 11, no. 2, pp. 1–23, 02 2015.

[113] A. Politi and M. Rosenblum, “Equivalence of phase-oscillator and integrate-and-fire models,” *Physical Review E*, vol. 91, no. 4, p. 042916, 2015.

[114] J. Cabral, E. Hugues, O. Sporns, and G. Deco, “Role of local network oscillations in resting-state functional connectivity,” *NeuroImage*, vol. 57, no. 1, pp. 130–139, 2011.

[115] P. Hövel, A. Viol, P. Loske, L. Merfort, and V. Vuksanović, “Synchronization in functional networks of the human brain,” *Journal of Nonlinear Science*, pp. 1–24, 2018.

[116] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti, “A framework to control functional connectivity in the human brain,” in *IEEE Conf. on Decision and Control*, Nice, France, Dec. 2019, pp. 4697–4704.

[117] T. Menara, Y. Qin, D. S. Bassett, and F. Pasqualetti, “Relay interactions enable remote synchronization in networks of phase oscillators,” *IEEE Control Systems Letters*, vol. 6, pp. 500–505, 2021.

[118] F. C. Hoppensteadt and E. M. Izhikevich, *Weakly Connected Neural Networks*. Springer, 1997.

[119] A. Li et al., “Linear time-varying model characterizes invasive EEG signals generated from complex epileptic networks,” Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2802–2805, 2017.

[120] A. Daffertshofer and B. van Wijk, “On the influence of amplitude on the connectivity between phases,” *Frontiers in Neuroinformatics*, vol. 5, p. 6, 2011.

[121] A. Li et al., “Neural fragility as an eeg marker of the seizure onset zone,” *Nat Neurosci.*, vol. 10, pp. 1465–1474, Oct. 2021.

[122] T. Menara, D. S. Bassett, and F. Pasqualetti, “Structural controllability of symmetric networks,” *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3740–3747, 2019.

[123] S. Gu *et al.*, “Controllability of structural brain networks,” *Nature Communications*, vol. 6, 2015.

[124] T. Menara, V. Katewa, D. S. Bassett, and F. Pasqualetti, “The structured controllability radius of symmetric (brain) networks,” in *American Control Conference*, Milwaukee, WI, USA, Jun. 2018, pp. 2802–2807.

[125] E. Wu-Yan *et al.*, “Benchmarking measures of network controllability on canonical graph models,” *Journal of Nonlinear Science*, pp. 1–39, 2018.

[126] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti, “Conditions for feedback linearization of network systems,” *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 578–583, 2020.

[127] R. F. Betzel, S. Gu, J. D. Medaglia, F. Pasqualetti, and D. S. Bassett, “Optimally controlling the human connectome: the role of network topology,” *Scientific Reports*, vol. 6, p. 30770, 2016.

[128] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti, “Functional control of oscillator networks,” *arXiv:2012.04217*, 2021, submitted.

[129] E. Sorrell, M. E. Rule, and T. O’Leary, “Brain–machine interfaces: Closed-loop control in an adaptive system,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 167–189, 2021.

[130] A. J. Bullard et al., “Design and testing of a 96-channel neural interface module for the Networked Neuroprosthesis system,” *Bioelectronic Medicine*, vol. 5, no. 1, p. 3, Dec. 2019. https://bioelecmed.biomedcentral.com/articles/10.1186/s42234-019-0019-x

[131] G. Santhanam, S. I. Ryu, B. M. Yu, A. Afshar, and K. V. Shenoy, “A high-performance brain–computer interface,” *Nature*, vol. 442, no. 7099, pp. 195–198, Jul. 2006. http://www.nature.com/articles/nature04968

[132] D. M. Taylor, S. I. Helms Tillery, and A. B. Schwartz, “Direct cortical control of 3D neuroprosthetic devices,” *Science*, vol. 296, no. 5574, pp. 1829–1832, Jun. 2002. https://www.sciencemag.org/lookup/doi/10.1126/science.1070291

[133] C. E. Bouton et al., “Restoring cortical control of functional movement in a human with quadriplegia,” *Nature*, vol. 533, no. 7602, pp. 247–250, May 2016. http://www.nature.com/articles/nature17435

[134] L. R. Hochberg et al., “Neuronal ensemble control of prosthetic devices by a human with tetraplegia,” *Nature*, vol. 442, no. 7099, pp. 164–171, Jul. 2006. http://www.nature.com/articles/nature04970

[135] L. R. Hochberg et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012. http://www.nature.com/articles/nature11076

[136] C. Pandarinath et al., "High performance communication by people with paralysis using an intracortical brain-computer interface," *eLife*, vol. 6, p. e18554, Feb. 2017. https://elifesciences.org/articles/18554

[137] B. Wodlinger et al., "Ten-dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations," *Journal of Neural Engineering*, vol. 12, no. 1, p. 016011, Feb. 2015.

[138] P. Nuyujukian et al., "Cortical control of a tablet computer by people with paralysis," *PLoS ONE*, vol. 13, no. 11, p. e0204566, Nov. 2018. http://dx.plos.org/10.1371/journal.pone.0204566

[139] S. Gowda, A. L. Orsborn, S. A. Overduin, H. G. Moorman, and J. M. Carmena, "Designing dynamical properties of brain–machine interfaces to optimize task-specific performance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 5, pp. 911–920, Sep. 2014. http://ieeexplore.ieee.org/document/6756992/

[140] V. Gilja *et al.*, "A high-performance neural prosthesis enabled by control algorithm design," *Nature Neuroscience*, vol. 15, no. 12, pp. 1752–1757, Dec. 2012. http://www.nature.com/articles/nn.3265

[141] F. R. Willett *et al.*, "Feedback control policies employed by people using intracortical brain–computer interfaces," *Journal of Neural Engineering*, vol. 14, no. 1, p. 016001, Feb. 2017. https://iopscience.iop.org/article/10.1088/1741-2560/14/1/016001

[142] F. R. Willett *et al.*, "Principled BCI decoder design and parameter selection using a feedback control model," *Scientific Reports*, vol. 9, no. 1, p. 8881, Dec. 2019. http://www.nature.com/articles/s41598-019-44166-7

[143] M. M. Shanechi *et al.*, "Rapid control and feedback rates enhance neuroprosthetic control," *Nature Communications*, vol. 8, no. 1, p. 13825, Apr. 2017. http://www.nature.com/articles/ncomms13825

[144] C. Mead, *Analog VLSI and Neural Systems*. Reading: Addison-Wesley, 1989, vol. 1.

[145] S. P. DeWeerth, L. Nielsen, C. A. Mead, and K. J. Åström, "A simple neuron servo," *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 248–251, 1991.

[146] K. J. Åström, "Event Based Control," in *Analysis and Design of Nonlinear Control Systems*, A. Astolfi and L. Marconi, Eds. Springer Berlin Heidelberg, 2008, pp. 127–147.

[147] R. Sepulchre, "Spiking control systems," *IEEE Proceedings*, To appear. 2022.

[148] G. Gallego *et al*., "Event-based vision: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2020.

[149] L. Ribar and R. Sepulchre, "Neuromorphic control: Designing multiscale mixedfeedback systems," *IEEE Control Systems*, vol. 41, no. 6, pp. 34–63, 2021.

[150] J. C. Willems, "The behavioral approach to open and interconnected systems," *IEEE Control Systems Magazine*, vol. 27, no. 6, pp. 46–99, 2007.

Appendices

A Pandemics: Modeling and Control

Sarah Fay, Anette Hosoi, Dalton Jones

Many traditional epidemiological modeling approaches incorporate detailed descriptions of disease transmission and individual behavior [84, 85, 86, 87]. These models may be fitted to collected data and can be simulated to assess the viability of a range of control strategies. However, the complexity of these models precludes testing all but a small subset of potential interventions [88]. In particular, obtaining a closed-form control law associated with a complex simulation is difficult unless the dynamics can be simplified.

Simpler models allow scientists and leaders to design policies that are rigorous from a control theory perspective (in the sense that it is robust to variations in the model and from external disturbances) and is also more easily understood by policymakers and other nonexperts. The current emphasis on complex simulations and models may hinder the development of simple yet effective control laws. Simplicity is key to effective mitigation because the policies designed using these control-theoretic methods will necessarily require the participation of individuals.

To be practically implemented, policy decisions derived from a systems theory perspective must emphasize interpretability and simplicity. Individuals who must be relied upon to enact such policies modify their behavior only when it is in their self-interest. Making it clear why a particular control law exists makes it more likely that individuals will recognize that acting as per the recommendations leads to optimal personal outcomes.

In the case of the SARS-CoV-2 pandemic, a study from Imperial College London noted that the uncertainties around how the disease was transmitted led to uncertainty regarding policy effectiveness. This limited the extent to which the population adhered to policies that mitigated the risk of spreading the disease [89]. Throughout history, there are numerous examples of governments that refused to disseminate accurate information that led to unnecessarily risky behavior among the populace. From the U.S. government's failure to act or even publicize the extent of the AIDS problem in the 1980s to the scapegoating of Jews in the Middle Ages for the spread of the plague, disinformation is an insidious and misunderstood aspect of infectious disease in society. Similarly to the COVID-19 pandemic, the relationship between uncertainty and unwillingness to adopt socially responsible behavior was illustrated during the 1918 Spanish flu pandemic. The American government's failure to report the seriousness of the disease led to uncertainty surrounding its virulence and a piecemeal, disorganized response [90]. In essence, control policies developed to mitigate the spread of disease are only effective when a baseline level of trust depends on the quality of information available. It is an open question of how researchers can understand the notion of trust and try to build it for more effective policy implementation.

Early in the SARS-CoV-2 pandemic, random testing and quarantine of positive individuals were recognized as simple, relatively unobtrusive ways to limit the spread of the disease with minimal disruption to society. While existing network-based models can be used along with Monte Carlo simulation to estimate the effectiveness of specific testing strategies [88, 91], these models do not produce a straightforward answer to the question, "How much testing is necessary to stop the spread of the disease?" On the other hand, the authors in [92] used a simplified branching process model to model the spread of a disease so a closed-form solution could calculate sufficient levels of testing for disease mitigation.

The details of the aforementioned paper [92] can be summarized as follows: Using a branching model of disease transmission, one may derive a formula for the effective reproductive number of the disease R_{eff} (the expected number of new infections originating from a single infection over the course of the sickness) that incorporates the level of testing, as

$$f_{\text{test}} \geq \frac{\mathcal{R}_0 - 1}{d - 1}.$$

This formula expresses the essential relationship between testing and the rate of spread of a disease, showing that as long as levels of random testing exceed a certain bound, the number of infections must decrease over time (where f_{test} is the fraction of randomly tested individuals, R_0 is the measured rate of spread of the disease, and d is the average duration of the infection). This expression demonstrates the power of simple modeling and control.

Simplicity, in this case, means that the policy requires estimating and manipulating a single parameter, the effective reproduction number R_{eff}. The relationships between the measured rate of spread and the required testing rate are clear (one increases with the other), and policymakers and nonexperts can quickly understand the necessity for testing and be more likely to participate.

On a practical level, the policy also remains consistent so that, instead of quickly changing requirements leading to precaution fatigue, there exists a simple set of guidelines that depend only on the measured value of the reproduction number R_0. Additionally, the policy can easily adapt if a change is detected in the virulence of a disease (as happened with the delta variant of SARS-CoV-2). Interestingly, the modeling techniques used for these results stray from the goals of many other epidemiological methods in that, instead of trying to predict the trajectory of the disease, only stability, in the sense that the number of infections inevitably decays, was considered. This suggests that a systems perspective can help control, if not necessarily perfectly predict, the course of an epidemic.

Practical guidance produced using these methods is the foundation of the testing policy recommended by the Centers for Disease Control and Prevention (CDC) for organizations [93]. These policies were implemented at the Massachusetts Institute of Technology, which gradually resumed normal functioning over 2022 without any outbreaks. While this preliminary work shows promise in developing effective methodologies to mitigate the spread of infectious diseases, many new decision and control problems have arisen during the pandemic.

The successes of the research applying systems theory to epidemiology [92] illustrate exciting opportunities to apply the formalism of our field to large-scale societal problems. In particular, the areas of opportunity for the application of decision and control in epidemiology include:

Interactions Between Social and Physical Systems: In understanding the dynamics of processes like epidemics and designing control laws to steer such dynamics, we need to know how physical processes (like disease transmission) depend on human behavior (like going to a restaurant) and vice versa. Additionally, it is clear that to affect a policy, one must change an individual's actions, and as such, we must ask how best to influence behavior, whether through direct mandates and control or other incentives or mechanisms.

The application of control theory to slow the spread of disease lies in a broad class of problems at the intersection of systems theory and the social sciences. In particular, to control an epidemic's trajectory, instead of applying a force to a mechanical or electrical system, the engineer must rely on individuals behaviors to actuate control strategies. In this way, applying control to epidemiological problems is one example of a collection problem that seeks to understand the dynamics of systems driven by physical processes (like disease transmission) and human behavior. The structure in which these two processes interact may be described using a network or other spatiotemporal representations. Such representations act as an estuary between physical systems and social systems. A fundamental understanding of these interactions is crucial to effective control design in this class of problems.

Understanding and refining theories in the social sciences can help predict and take advantage of how people process and respond to information and incentives. Researchers are just beginning to understand how disseminated information can change individual behavior. Recent research [94] has shown that data on local infection rates has a self-regulating effect as people in places with high infection rates modify their behavior to reduce their risk. Therefore, the proper release of information to the public can serve as a signal for individuals to behave more safely. Similar ideas exist in game theory (such as the *coarse correlated equilibrium*), and these tools may be important in understanding how information can lead to action in a social system.

On the other hand, irregular dissemination of untrustworthy information can limit the extent to which a population adopts preventive behaviors. This is demonstrated by the continued erratic behavior of a large proportion of the American public during the pandemic. At a fundamental level, knowing what information to release and when requires understanding how individuals make decisions that they perceive to be good for them. This requires tools traditionally in the domain of economics and psychology. Effective models in this new domain of humans-in-the-loop control must partition inputs into those that can be directly specified (such as government mandates) and those which require indirect incentivization to shape human behavior. Such interplay between direct and indirect application of control is another novel avenue of research for future consideration.

Understanding Network Effects: Designing control systems in situations in which physical dynamics and social dynamics play a role require a thorough understanding of both types of systems. For social dynamics, many processes are driven through some underlying network. Networks also serve as a contact point where physical processes and social dynamics mix. For example, a social network for COVID-19 could be created by connecting any individuals who spent time in close proximity (which is the only way the disease could be spread.) In this way, the network is a product of the interaction of social processes and disease dynamics. Understanding the network sheds light on both of these aspects driving epidemiological problems. For example, the structure of the network may suggest something about transmission. If the degree distribution follows a power law, where many connections originate from highly central individuals, it may be that the bulk of transmission comes from "super-spreading" events. This information may be critical to implement policies limiting the spread of a certain disease. However, it should be noted that in certain general situations, as studied in [92], exact knowledge of the network is unnecessary to implement sufficient policies that stop the spread of disease. It remains an interesting question when network effects matter and when they do not.

Understanding the interplay between networks and dynamics is critical in effective controller design. Furthermore, the fidelity with which networks and interactions are modeled must be addressed. In the case of epidemiology, models range from very coarse population-level susceptible-infectious-recovered (SIR) models built upon a simple set of coupled differential equations to highly complex individual-based models that track the progression of a disease as it spreads through a simulated society. The scale at which these dynamics are studied depends on the data available to support the model and the scale at which control actions are implemented. For example, if we only have population-level infection data, it would be difficult to assert that disease transmits through one network or another. Additionally, if control is implemented on a weekly timescale, it is likely unnecessary to model disease progression hourly. Finally, to obtain tractable control solutions, simple, accurate models must be obtained that capture the essence of the underlying disease dynamics.

In the case of epidemics, [92] showed that for control, the full network structure need not be known to predict the asymptotic behavior of the system. However, the network impacts other aspects of a disease's progression. In particular, it can be seen in Figure 2.13 that different network structures result in different peak infection rates at different times when no control is applied, even when the reproduction rate R_0 remains the same across simulations.

Learning for Control: One of the difficult aspects of applying control theory to epidemiology is being able to quickly recognize the dynamics of the disease. Given the exponential nature of early spread, this recognition must lead to the rapid application of disease-slowing interventions before an epidemic gets out of control. However, as previously discussed, modeling disease dynamics is a complex problem.

The issue for control theorists is to simultaneously learn models with sometimes very limited noisy data and to design algorithms using those models to derive policies meant to slow the spread of disease. These algorithms ideally should adjust to changing parameters over time. One example is the work [92] in which a single parameter R_0 was estimated. This parameter was then used to derive the necessary amount of testing needed to drive the number of infections to zero over time. This algorithm additionally displayed a desired responsiveness to changing model parameters in the sense that, as the SARS-CoV-2 virus mutated, it became more infectious. As a result, the methodology recommended a higher level of testing in order to control its spread. These results are shown in Table 2.1 in which the testing required to stop the spread of various diseases is given in terms of their reproduction numbers.

Of course, our domain-specific knowledge in public health allows us to quickly implement policies limiting common pathways through which diseases spread (either from person to person or from some other common source). For example, social distancing and masks will slow the spread of the majority of human-borne infectious diseases like influenza, SARS-CoV-2, or the measles. Therefore, this prior knowledge can be used to immediately blunt the spread of a newly detected disease. The challenge then becomes how to learn and apply control quickly, leveraging previous knowledge in a provable, robust way.

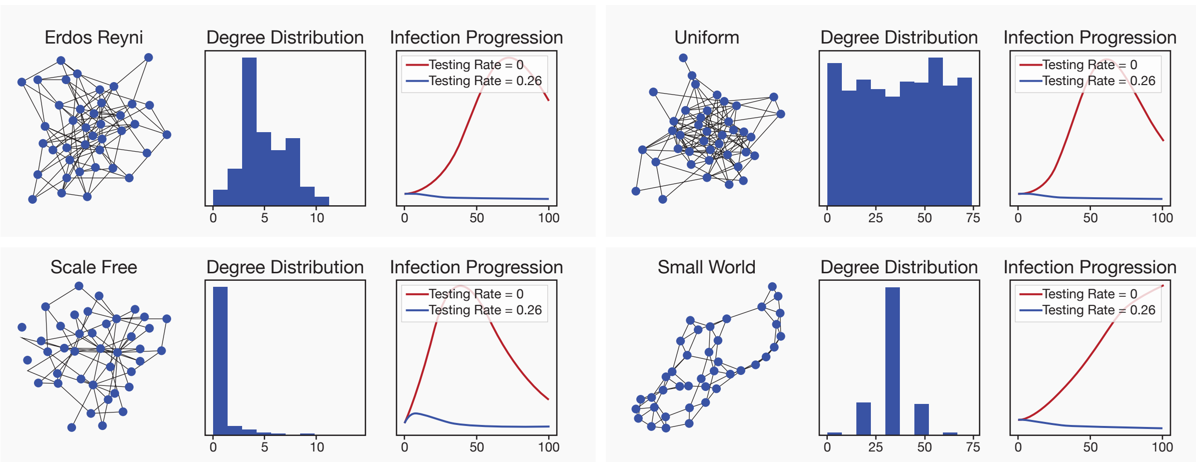


Figure 2.13: Sample realizations of four different network structures and their resulting degree distributions. Intervention-free infection was simulated on each network to make a measurement of R_0. The required testing to stabilize the system, $f_{T,\min,0}$, was determined and another infection simulation was implemented with $f_T = f_{T,\min,0} + \delta$, where $f_{T,\min}$ represents the minimum percentage of the population randomly tested at each time step required to move R_0 below 1. Both infection progressions are shown: intervention-free (red) and with stabilizing testing (blue). The two rightmost plots in each subplot show a histogram of measured degrees, giving an empirical estimate of the degree distribution, and the last shows the progression of the disease through percentages of a population (with a maximum of 1). All of the infection progression graphs are plotted on the same axes.

There are also several other nuances in learning that should be noted. In the case of SARS-CoV-2, uncertainty due to limited data regarding the dynamics of disease transmission and danger led to a delayed and inadequate response. Effective control must act quickly based on little information. A theoretical understanding of how quickly good control methods can be enacted in the case of limited historical system measurements must be developed. As different methodologies and interventions are enacted, it is crucial to establish the causal effects of system inputs to better refine future interventions. Data collected in 2020 indicated that the greatest changes in personal mobility (a proxy for the level of lockdown enacted) occurred in locations where the greatest number of deaths occurred. While the lockdown occurred *before* the deaths began to rise, this does not mean that the level of lockdown led to greater death rates. Instead, it showed that locations with higher infection rates had stricter lockdowns. A more thorough analysis would also consider the lag in the progression of the disease from infection to death. Untangling the web of causality and correlation is important to quantify the effectiveness of different policies for use in controller design.

As an aside, confusing the causality regarding lockdowns and death rates is a good example of how wrong inferences can occur if data is not handled carefully. In this case, an unobserved latent state drives the dynamics (the true infection rate). Identifying the effects of these latent unobserved states from a set of sparse measurements (potentially taken from different but similar systems) is a challenge for system identification. In the case of COVID-19, given information about lockdown timing and severity along with infection rates and death rates across a number of cities will allow us to make better inferences (provided the data can be properly combined). Using a model-based approach, doing a counterfactual analysis is straightforward enough. Fit the model using the observed data, then alter the inputs to determine the outputs change. However, assuming a particular model structure introduces bias to the system, it is advantageous to do such counterfactual analysis in a data-driven, not model-driven, way.

Learning how intervention A leads to outcome B is critical in designing effective epidemiology policy. Usually, such mathematical questions fall under the umbrella of *causal inference*, a set of statistical techniques used to determine the actual effects of particular inputs to a system. Interestingly, this theory is not so well developed for dynamical systems and control. However, the extension of this theory to our field would allow researchers to answer many important questions. In the case of epidemiology, we need a way to describe a situation in which an intervention was applied differently. For example, we might want to know how the rate of infections in Denver *would have* differed if a lockdown was not implemented. It is not possible to learn this through a control group because it would be unethical to place the individuals in such a group in a potentially more dangerous situation. Therefore, can we infer a change in outcome based on different inputs using observed data? One way to do this is by using “synthetic controls” which can eliminate underlying confounders to determine the causal effects of treatments on outcomes using only observed data. The application of such techniques to control systems is in its infancy but could provide new ways to solve difficult problems in system identification.

Uncertainty Quantification: Is it possible to quantify the uncertainty in model estimation to account for unknowns as part of the controller-design process? Moreover, can we demonstrate that there exist control methodologies that work with some degree of success in the presence of uncertainty and improve as more data about the process is collected? Such an understanding will help us determine which granularity control should be implemented. For example, we have good data on how SARS-CoV-2 spreads on a weekly basis. Based on this, does it make sense to develop a control method that changes minute by minute? (Probably not.) More broadly, how do the type and quality of data collected inform possible controller strategies and design?

Using Simulations to Design Policy: Can we use existing models and epidemic simulations to test our methods? In particular, can high-quality reinforcement learning methods be designed using complex simulations in the loop? Could such simulations be used with model reduction techniques to develop mathematical descriptions more appropriate for use in traditional controller-design algorithms? Of particular interest would be using existing simulations to learn simplified descriptions of the system, similar to the work done to understand complicated phenomena like turbulence [95].

In previous work, simulations have been used to motivate and corroborate the effectiveness of simpler models that were more tractable to use with control methods. In [92], a complex individual-based simulator was initially used to track the progression of several pandemic scenarios. Later, a branching model was used that upper-bounded the number of infections of the more complex simulator. The upshot of this was that any policy that controlled the upper-bounding branching model also necessarily controlled the epidemic in the simulator. This can be seen in Figure 2.14.

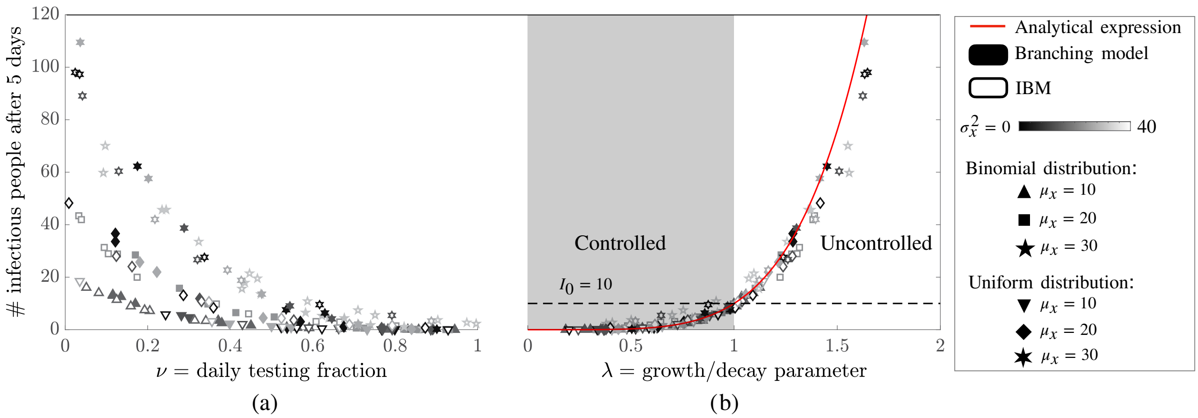


Figure 2.14: Number of infectious individuals at $t = 5$ for $I_0 = 10$ (dashed line) as a function of (a) the daily testing fraction and (b) the parameter λ, for various models of interactions in the population. The collapse in λ shown in (b) indicates that stability is independent of the variance and type of distribution of interactions. The shaded region represents where we expect to see stable dynamics, i.e., where the number of infectious individuals decreased from I_0. Each point represents the mean of 100 trials with the same parameters.

Equity in Decisions and Policy: How can we ensure that policies and decisions based on our methodologies treat individuals fairly and do not discriminate based on any underlying social condition (i.e., ethnicity, economic status, etc.)? Even better, how does one design control algorithms that promote social welfare?

Recent work [96] discussed the difficulty of balancing the economic impacts of COVID-era pandemic-easing policies like lockdowns on the upper and lower class. In particular, because lower income populations are more likely to live in larger family units (especially with older family members) and because their jobs entail a higher risk of contracting disease, the paper [96] argues that subsidies should be given to these at-risk groups to help them social distance and still pay their bills. Further research in this vein should consider ways to optimize the distribution of resources so that those without the means to sustain themselves during an emergency receive sufficient aid, while preventing wasteful allocation to those who do not need help.

Understanding Tradeoffs: In the case of epidemics, the most effective nonpharmaceutical policies to steer the course of a disease through a population is to shut everything down. This is also the worst possible thing to do in terms of economics. Other interventions may positively impact the course of an epidemic but might have unintended secondary effects on mental health.

When applying policies on a large scale, some unintended consequences naturally occur. In a sense, this phenomenon might be an extension of the waterbed effect from classical controls to social systems. It is important to quantify and respect such tradeoffs in tackling these problems as, in this case, totally neglecting economic or psychological well-being may potentially lead to worse outcomes than those that arise from the spread of disease.

Table 2.1: Levels of random testing required in the absence of a vaccine to control the spread of common diseases for two contact tracing scenarios. The rate of testing indicates the probability of an individual being tested on each day; 1 corresponds to daily testing for everyone. The required level of testing decreases as contact tracing increases (see fourth column where 50% of contacts are traced after a successful test). Estimates for COVID-19 are based on aggregate global data in early 2020 [87] and the delta variant in 2021 [97].

	(R_0)	(d)	Fraction tested	Fraction tested w/ contact tracing
COVID (Alpha)	3.5 [87]	14 [87]	0.23	0.2
COVID (Delta)	6 [97]	14 [97]	0.34	0.27
SARS	1.08 [98]	14 [98]	0.006	0.004
Flu	1.4 [99]	7 [99]	0.07	0.06
Measles	12.0 [100]	14 [100]	0.85	0.58

Future results in our field, when applied to systems that are intrinsically linked with society (such as energy generation and use or the economy), must take a holistic approach when considering which variables to control. Throughout the pandemic, public health and the economy seemed to be at odds as different pandemic-easing policies, such as lockdowns, placed stress on workers and businesses. As such, the aspiring societal control theorist must anticipate the consequences of different actions and incorporate robustness to minimize the effects of unexpected behaviors. While these methods are sometimes studied in the social sciences, finding ways to bridge the gap (particularly mathematically) between existing theory and the formalism of system theory is an important emerging problem.

Ultimately, systems and control theory is uniquely positioned to augment existing methodologies to model and control the course of epidemics. The problems discussed above also belong to an exciting area of research, coupling social behavior with the dynamics of complex systems. Solutions in any of these domains will naturally carry over to many interesting applications. The future of control theory is not to continue refining our analysis of essentially well-understood mechanistic systems but to begin applying proven and novel techniques in estimation and control to manipulate novel dynamics—in this case, the interplay between natural phenomena and society.

B Neuroengineering

Tommaso Menara, Timothy O'Leary, Fabio Pasqualetti, Sri Sarma, Rodolphe Sepulchre

Brain-machine interaction is a major scientific and technological challenge and opportunity of the early 21st century. This is compellingly demonstrated by the BRAIN Initiative in the U.S. and the Human Brain Project in Europe. New imaging techniques have advanced our understanding of brain function by providing data at an unprecedented temporal and spatial resolution. Novel implant technologies have given us the ability to stimulate brain activity and offer new medical treatments for brain disorders such as depression and Parkinson's disease. Brain-machine interfaces provide unprecedented ways to close the loop between biology and engineering and create new rehabilitation protocols for injuries of the spinal cord or the nervous system. Every development in neuroscience inspires progress in artificial intelligence, presenting new opportunities for embodied robotics, neuromorphic sensing and actuation, and brain-inspired computing.

Models for Brain Diagnosis

Reverse-engineering the structure-function relationship in the brain remains an open challenge in modern neuroscience. However, there have been tremendous advances in methods to measure and quantify the structural backbone (DTI, MRI) and the functional evolution (fMRI, EEG, iEEG, MEG) of brain dynamics (see Figure 2.15 for an example). In this context, the maturation of the fields of connectomics and network neuroscience [101, 102] has given rise to the study of the brain as a dynamical system whose relevant dynamics are governed by network-wide rules. In this setup, the communication between neuronal populations is constrained by axonal bundles (white matter tracts) and is hierarchically regulated across many scales of spatial and temporal organization. That is, the interconnection scheme between neural components constrains the evolution of brain dynamics, and the integrative nature of brain function can be addressed from a complex network perspective.

Since the discovery of brain rhythms in 1929 [103], a multitude of brain functions have been associated with synchronization phenomena. During a cognitively demanding task or at rest, the brain exhibits a rich repertoire of large-scale synchronization patterns supported by its static interconnection scheme. These patterns are a measure of the coherence among the neural activities in different brain areas and are typically referred to as functional connectivity. Functional connectivity quantifies the level of correlation between time series of neural activity measured from distinct brain regions. Given a specific time window, taking the correlation across all brain regions yields a pattern that characterizes the synchronization levels between distinct regions in that window and is affected, for instance, by the tasks being performed and by external stimuli [104]. Importantly, functional connectivity can be used as a biomarker in multiple psychiatric and neurological disorders, and abnormal or decreased synchronized activity has been linked to neurological damage or cognitive decline due to aging [105, 106].

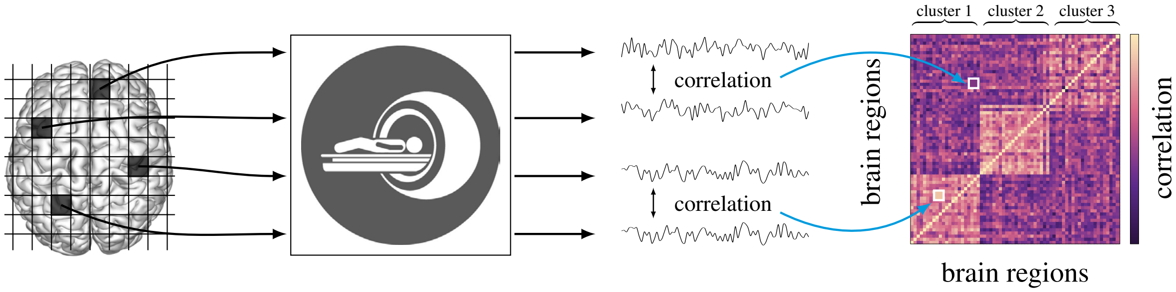


Figure 2.15: The human brain can be partitioned into a number of distinct brain regions, whose neuron activity can be measured through, for instance, an fMRI machine. Computing the correlation between fMRI time series in some time window yields a pattern in which each entry quantifies the level of correlation between the time series of two regions, and where clusters of synchronized regions can be identified. Such a pattern identifies what is also known as functional connectivity.

Dynamical brain networks can be used to model and analyze the spontaneous evolution of functional connectivity over time and in response to endogenous or exogenous stimuli. In these models, the brain structure is estimated by DTI or MRI technologies, and determines a graph that constrains the evolution of brain activity. Mathematically, the activity x obeys a differential equation of the form $\dot{x} = f(x)$, where the vector field $f(\cdot)$ is based on the underlying anatomical network structure. A seminal line of work employs noise-free, linear, continuous-time, and time-invariant models of network dynamics [107, 108]. While the choice of linear dynamics may overly approximate the intrinsic nonlinear nature of large-scale neural communication, it provides a starting point for the analysis of brain network dynamics, and enables the utilization of a battery of powerful tools from linear algebra [109]. Furthermore, at a coarse scale, brain activity seems to effectively fit linear models [110]. Nevertheless, more complex (nonlinear) dynamical models have been proposed.

The interaction between static large-scale structural architecture of the human brain and local oscillations of neural communities is a key factor in determining the functional connectivity patterns that are empirically observed through, for example, fMRI, especially when the brain is in a resting-state condition [111]. Motivated by this observation, extensive literature resorts to Kuramoto phase oscillators to model fMRI data [105, 112, 113, 114, 106, 115, 116]. Many studies focus on the analysis of the oscillatory behaviors of neural populations that lead the emergence of functionally connected networks by modeling fMRI data as the output of networks of Kuramoto oscillators [105, 112, 114, 117]. The main working assumption is that at each node of the structural brain network exists a community of excitatory and inhibitory neurons whose dynamical state is in a regime of self-sustained oscillations. From a modeling standpoint, this assumption is equivalent to employing a network of weakly coupled Wilson-Cowan oscillators [118, 120]. In this setting, the neurons' firing rates describe a closed periodic trajectory in phase space. That is, the firing rates delineate a limit cycle. Thus, the dynamics can be approximated by a single variable, which is the angle (or phase) on this cycle. This regime is then modeled by a network of coupled heterogeneous Kuramoto oscillators that are connected to each other according to the architecture of the human brain.

Another approach to constructing brain network models is to estimate dynamics directly from EEG recordings with no structural constraints [119]. One clinical application involves identifying where seizures start in an epileptic cortical network, called the seizure onset zone (SOZ). Drug-resistant epilepsy patients often undergo surgical treatment of the SOZ since medications are ineffective, but surgical success rates average 50%.

Localizing the SOZ is challenging and involves analyzing intracranial EEG recordings captured during seizures. A recent study estimates a sequence of linear time-invariant (LTI) models of the form $\dot{x} = Ax$ for 500 millisecond windows from iEEG snapshots. Then the "fragility" of an iEEG node j, defined as the minimum perturbation $||\Delta||$ that is applied to the jth column of A as follows: $\dot{x} = (A+\Delta)x$, is computed from the model A. The hypothesis tested is that the most fragile nodes in the iEEG network are the SOZ, and it was shown that fragility outperformed standard spectral features and graph-theoretic measures in a cohort of 91 patients treated across five epilepsy centers [121]. The processing steps from iEEG to a fragility heatmap is shown in Figure 2.16.

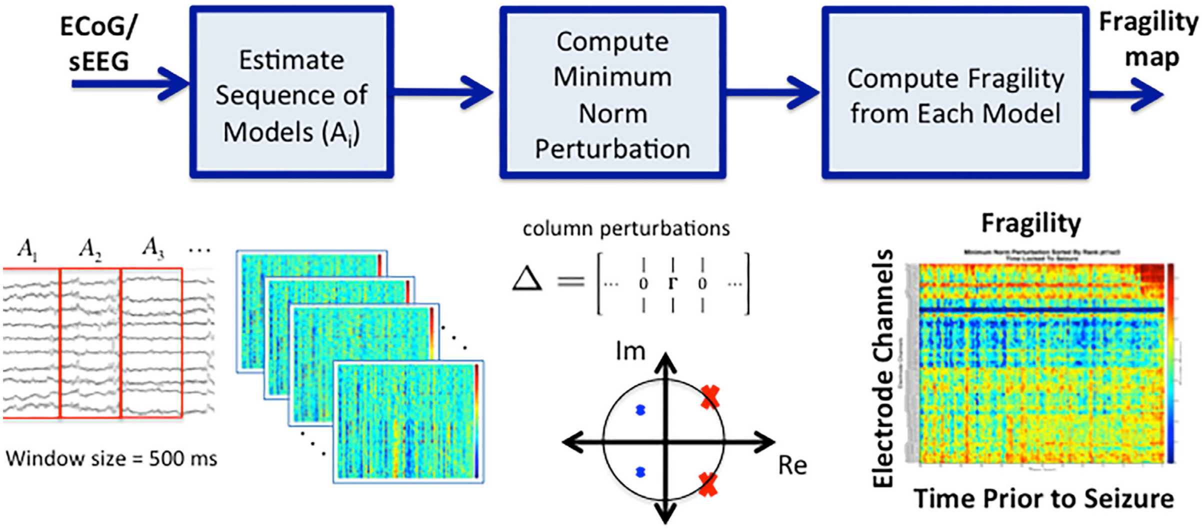


Figure 2.16: Processing steps to compute fragility heta map from intracranial EEG data.

Control for Brain Medicine

The brain is constantly juggling between multiple cognitive states as the result of external stimuli or spontaneous activity. The topological and algebraic features of network systems that enable exogenous/endogenous inputs to steer the system's state to desired values can shed light on how distinctive stimuli affect the system-wide state of the brain. In engineering terms, this analysis can be performed by investigating controllability of neural network systems. The main contribution of this line of research is to lay the groundwork for the extremely challenging task of synthesizing control signals to safely and precisely alter brain activity. Seminal work has derived prescriptive conditions for structural brain networks to be theoretically controllable [122]. The investigation of controllability properties has also shown that i) some regions are easier to control [123], ii) brain networks possess distinct controllability profiles with respect to randomly generated topologies [124, 125], and iii) it is possible to characterize the organizational features that allow the exact linearization of nonlinear (neural) circuits through feedback loops [126].

Going beyond the study of controllability, network control theory provides tools such as optimal control methods that have been used, for instance, to quantify the energy needed to switch between different cognitive states [127]. Optimal control principles have also been used to integrate structural and functional brain data to reveal the relationship between different memory states and the underlying anatomical maps. In fact, it is possible to estimate how regional activity would deviate from its initial state in the presence of external current stimulation and to test which brain regions would efficiently improve memory encoding when stimulated [107].

Other contributions of optimal control theory reveal that, compared to healthy controls, subjects affected by schizophrenia show altered network control properties, including decreased stability of working memory states [108].

Because synchronization (and, thus, functional connectivity) can also be used as a biomarker, some work pursues the ambitious idea of analyzing, predicting, and treating neurological disorders by enforcing the formation of desirable patterns of neural activity to recover from abnormal or undesired ones. Such a goal can be achieved by developing a framework that allows to control functional connectivity by means of minimally invasive and localized interventions [116, 128]. The notion of cluster synchronization—which describes synchronized groups of regions that coexist in the brain network—enables the blending of mathematically rigorous methods with physiological models of brain activity, with the goal of steering whole-brain synchronization dynamics. In this framework, network parameters such as natural frequencies and oscillators' connectivity can be used as control knobs to achieve different configurations of synchrony. Minimally invasive control actions are then computed by casting the pattern assignment problem as optimization methods.

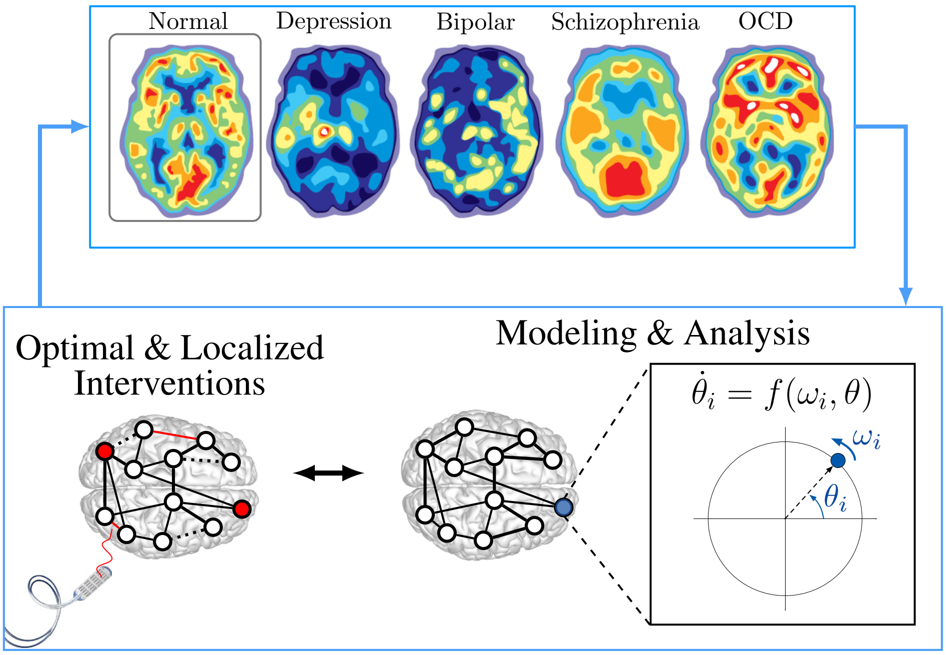


Figure 2.17: Modeling, analysis, and control of abnormal neural activity. Networks of nonlinear oscillators can be used to inform the design of sparse and localized interventions.

Brain-Machine Interfaces

Physiological signals can be measured directly from the brain, opening the possibility of decoding these signals in real time to control devices and to report internally represented information in the brain. The field of brain-machine interfaces aims to deliver technology that will allow individuals to directly control and interact with external hardware using brain signals alone. Despite huge progress in this field, significant challenges remain, some of which are fundamental. The human brain is a large and fragile organ enclosed in a skull several millimeters thick. Electromagnetic signals obtainable noninvasively (for example, via scalp EEG) are thus subject to significant attenuation and interference from external sources, as well as larger muscle-related potentials from the face and neck. Even when placed inside the skull, (ECoG, intercranial EEG), signals lack spatial resolution and represent pooled measurements from many thousands of neurons, mostly those located in superficial layers of the cortex. Other modalities such as fMRI are not practical for most BMI applications as they lack temporal as well as spatial resolution and require a subject to lie immobile in a large and expensive piece of medical equipment.

Together, these considerations explain why high-resolution, high-bandwidth BMI applications require invasively measured brain signals [129]. Such signals are usually obtained from microelectrode arrays that are implanted in the brain that can gather extracellular potentials from nearby cells and nerve fibers. Array technology is rapidly advancing in its channel capacity (number of electrodes), biocompatibility, and signal resolution [130]. It is now possible to measure activity at single cell resolution from populations of thousands of cells. However, resolving individual cells requires source separation, as each recording channel will typically detect signals from many tens of individual cells. So called single-unit resolution can be obtained by offline source separation, colloquially known as spike sorting, but current state-of-the-art techniques preclude source separation in real time. As a consequence, signals used for controlling devices (such as prosthetic arms, speech synthesizers, or computer cursors) operate the superposed field potentials obtained from each electrode in an array [131, 132]. Thus, the measured signal is essentially a low-dimensional projection of the state-space representing neural population activity. Remarkably, it is often possible to control devices with several degrees of freedom in real time despite this coarsening of the spatial signal [133, 134, 135, 136, 137]. State-of-the-art BMI devices have enabled human control of robotic arms [135] and text transcription speeds of 30 characters per minute on a tablet computer [138].

Cutting-edge BMI technology presents many open control and signal processing problems. Neural signals obtained from intracranial probes are subject to nonstationary disturbances due to electrode and tissue movement and changes in electrode impedance due to gradual changes in the wet electrochemical interface of an extracellular probe. Consequently, it is necessary to devise methods for tracking and recalibrating BMI decoders. Just as mechanical computer and machine interfaces present nontrivial design considerations for control with a human-in-the-loop, there are means of optimizing the stability and responsiveness of BMI controllers, taking the dynamics of neural signals into account [139, 140, 141, 142]. Finally, the brain itself is an adaptive controller. This means that the long-term dynamics of the human-BMI system are subject to ongoing learning and reorganization of neural dynamics. This means that BMI users can often compensate for imperfections in BMI performance [143, 141]. Conversely, it makes the long-term performance of BMI challenging to guarantee, likely requiring deeper knowledge of the dynamics of neural activity on long timescales as humans learn, adapt, and experience daily life [129].

Neuromorphic Machines

To date, the brain remains our model of an intelligent machine. But despite the tremendous progress of machine intelligence over the last 50 years, the gap between machine intelligence and brain intelligence remains phenomenal. Wiener's cybernetics vision of a unified theory of control and communication in the animal and in machine remains to date a distant dream. One could even argue that the distance has never been greater in the digital age, where bits and automata reign supreme in artificial intelligence while they are completely absent from animal intelligence. At least in the pre-digital age of Wiener, analog electrical circuits provided a common language to model signals and systems in the animal and in the machine.

In the midst of the golden age of silicon technology, Carver Mead realized that digital technology would eventually face limitations that would call for a post-digital age [144]. He became fascinated by the efficiency gap between the digital technology of artificial intelligence and the spiking technology of animal intelligence and created the new field of neuromorphic engineering. Karl Åström immediately recognized the significance of a spiking technology for control, in a seminal article that has remained largely unnoticed to the present time [145]. Ten years later, he wrote the first article about event-based control [146]. Event-based control is now flourishing, but still in its infancy.

It takes more than a spike to make a machine intelligent, but the question is whether spiky rather than analog or digital communication is essential to intelligence. Control is nowadays divided into automation and feedback control. We praise automation for its reliability, which is owed to the discrete nature of bits and automata. We praise feedback control for its adaptation, which is owed to the continuous nature of analog signals and physical systems. But automata are inefficient in their adaptation capabilities, and feedback systems are unreliable in their decision-making. Spiking control systems [147] will perhaps be essential in mimicking the ability of nervous systems to combine adaptation and reliability in a way that defeats any current form of artificial intelligence.

The forthcoming decade will see the emergence of sensing and actuator devices that are spiky and event-based rather than analog or digital. Event-based cameras revolutionize the technology of vision, but it is still largely unknown whether and how they will revolutionize the intelligence of artificial vision [148]. Interconnecting event-based sensors and event-based actuators will lead to a novel generation of machines. Whether the intelligence of those machines will be one step closer to the intelligence of animals is a grand challenge for control theory. Building reliable functions from uncertain components is at the core of feedback control. Neuromorphic control will aim at imitating the animal world in it success to achieve both adaptation and reliability across a broad range of temporal and spatial scales [149].

Control theory has championed the concept of control as variable sharing, and of variable sharing as being facilitated between systems that share external properties [150]. Whether we think about designing artificial systems that will interact with animal brains or artificial systems that emulate animal brain functions, we will need to integrate some essential properties of animal brains that are currently absent in AI. Spikes, excitability, and neuromodulation are some of those key elements. The dramatic progress of neuroscience, the growing promise of brain medicine, and the booming technology of neuromorphic engineering provide a clear call for this road map document.

CHAPTER 3

Technological Trends

Technological Trends

Innovations in computation have dramatically impacted the field of control systems. As computing has become faster, smaller, and more ubiquitous, these trends have transformed control systems design and implementation. As networking has grown over the past three decades, new visions of networked and distributed control have emerged. While developments in computation and networking will continue to impact control systems, this chapter discusses new scientific and technological trends that are likely to shape the evolution of control systems over the next decade.

First, we discuss the dramatic rise of artificial intelligence and big data. While the objectives of AI and control systems are conceptually aligned, recent advances in AI infrastructure and emerging AI models and architectures present new opportunities for our field. In a totally different technological space, we are witnessing a second electrification revolution in the transportation sector. Control will be affected by this trend and new opportunities will emerge. As a more disruptive trend, biology may serve as a new substrate for building new biological systems for engineering, agricultural, energy, and healthcare functions. Control can be a critical design methodology in modeling, robustness, and the design of biological engineering systems.

Finally, the focus of robotics and autonomy is transitioning from indoor industrial robotics to robotics in the wild, bringing new challenges for control systems. Each section of this chapter describes these trends, opportunities for control systems, and recommendations for going forward.

3.A AI and Big Data

John Baras, Mykel J. Kochenderfer, James Anderson, Luca Schenato

> An agent that is capable of sensing its environment and taking actions to maximize its utility is said to possess artificial intelligence (AI). Half a century of control systems has produced a rigorous theoretical framework for addressing this issue in highly structured environments. Over the next 30 years, we expect control systems to merge with AI, enabling big data analytics to inform robust autonomous systems that can be deployed in real-world, human-inhabited environments.

Abstract The emergence of AI, big data, and machine learning has played a major part in what is now referred to as the Fourth Industrial Revolution or Industry 4.0. Breakthroughs in these fields have provided us with enhanced autonomy, nearly real-time decision-making, and smart infrastructure. Despite numerous success stories, there have been many severe failures and issues of reliability and robustness, especially in the context of human-machine interaction. Control engineering has a long history of providing both rigorous theory and implementable algorithms for providing guaranteed performance and robustness bounds in dynamic and uncertain environments. In this chapter, we address how control systems and AI can synergistically evolve to solve societal-level engineering problems.

3.A.1 Introduction

AI is a field concerned with designing and building intelligent agents that sense the environment and take actions that affect that environment [1]. There are deep connections with systems and control theory, though to some extent the AI community has developed independently. The AI community formed with the 1956 Dartmouth Summer Research Project on Artificial Intelligence, where academics such as John McCarthy, Claude Shannon, Marvin Minsky, Herb Simon, and Alan Newel discussed ways to build machines that could exhibit intelligence [2].

An important component of AI is machine learning (ML) [3], where data is used to build models [4]. Major breakthroughs in AI came with the "big data" revolution [5]. Large datasets [6] enabled breakthroughs [7] in many domains, notably in computer vision [8] and natural language processing [9]. Machine learning methods have influenced approaches in the control community and vice versa [10]. There are other areas of AI outside of ML that may become ripe for synergies with control theory, including knowledge representation and reasoning (KRR) [11].

AI is continuing to evolve in many different directions. This section will summarize some important technical and scientific trends in AI and big data that may be beneficial to the controls systems community. In addition, this section will sketch some of the important challenges and opportunities for the control community to contribute.

3.A.2 Technical Trends

There are several technical trends in AI and big data that may be helpful in the engineering of control systems (see Figure 3.1 for a schematic and Figure 3.2 for technology trends).

Hardware: The increased availability of inexpensive hardware for storing and processing large datasets has helped enable major advances in AI. Many ML applications use neural network representations of models and decision strategies, often with many layers and many neurons. The standard method for optimizing the parameters in these neural networks to achieve a given task, such as classification or control, is stochastic gradient descent [12]. Computing architectures like GPUs and TPUs [13] can help make computing gradients efficient enough to train extremely large networks [14]. There have also been recent innovations in chip design. For example, so-called "neuromorphic" chips use a biologically inspired design modeled after neurons and synapses [15]. Such designs may provide greater parallelism with significantly reduced energy costs.

Software: Training neural networks is key in the computation of gradients. There have been software advances that make it easy for users to compute gradients of not only neural architectures but also general software programs. Automatic differentiation has been around for decades [16], but there have been several software packages that have emerged that enable practitioners to efficiently define computational models and optimize parameters [17]. Some, such as PyTorch [18] and TensorFlow [19], have been central to large machine learning systems.

Simulation Frameworks: The AI community has spurred the development of many high-quality simulation frameworks, many of them open source. The OpenAI Gym [20] has emerged as a standard toolkit for reinforcement learning. It contains a collection of interactive benchmark problems, providing a common interface for evaluating different reinforcement learning algorithms. The physics engine MuJoCo [21] has been used extensively in robotics research. ROS [22], a robot operating system, has greatly enhanced the pace of academic robotics research and has seen broad deployments in industry.

Cloud Computing: The AI community has greatly benefited from the availability of cloud computing infrastructure. Training on large datasets often requires massive computing resources that may only be needed sporadically. As an alternative to numerous research labs and companies purchasing and maintaining their own infrastructures that are only used in bursts, there has been a strong movement toward the cloud, where resources are pooled and users only need to pay for the computing they use. Although major advances in AI are still being led by industrial research laboratories with enormous computing resources of their own (like Google and Microsoft), the availability of cloud computing resources from AWS, Azure, and GCP has enabled academic institutions to contribute. Cloud computing also allows AI applications to easily scale up from initial research prototypes to deployed systems.

3.A.3 Scientific Trends

There are several scientific trends in AI and big data that may also be helpful in the engineering of control systems.

Reinforcement Learning: A subfield of AI known as reinforcement learning is concerned with the problem of learning how to choose actions in a dynamic environment to maximize a reinforcement signal [23, 24]. Although this field dates back decades, it has attracted tremendous attention recently, especially with the ability to train neural networks to play Atari games at a human-competitive level based on pixel data [25]. Reinforcement learning has also led to machines beating the best human experts at Go [26], which for many years was suspected to be beyond the reach of AI. Major advances have also been made in applying reinforcement learning to cooperative decision making [27].

Neural Architectures: The AI community has explored a wide variety of neural network models to address fundamental issues like scalability and gradient estimation in recurrent networks. Examples of such architectures include long short-term memory (LSTM) [28], Gated-Recurrent Units (GRUs) [29], transformers [30], and Generative Adversarial Network (GANs) [31]. Many of these architectures can be used for control systems.

Trustworthy AI: Many of the early successes of AI have been in domains that are not safety-critical. However, there is growing recognition that AI techniques can be helpful in building high-stakes systems, such as automated driving and aircraft collision avoidance [32]. Compared to traditional techniques, AI has the potential to provide greater robustness to the various sources of uncertainty inherent in the real world. Major research efforts are underway in explainability [33], validation [34], and verification [35].

3.A.4 Challenges and Opportunities

The previous sections illustrate many exciting new hardware and software technologies in which control systems has the potential to advance fundamental progress in the broad area of AI and big data. The potential downstream applications where AI is poised to make a substantial contribution are so numerous that we do not attempt to enumerate them all. However, each societal area described in Chapter 2 of this road map is well within this scope. In particular, control engineering is uniquely placed to help analyze, design, and provide a better understanding of neural network architectures; provide a principled framework for uncovering and incorporating robustness into AI algorithms and pipelines; and provide verifiable guarantees of safety for complex human-in-the-loop, AI-driven applications. The remainder of this section surveys the challenges and opportunities that control engineers face at the rapidly evolving intersection of control, AI, and big data.

Designing Neural Networks: As described in the previous section, numerous different neural network architectures have been designed and promoted. A key question that arises is, how do we choose the appropriate architecture for a given task? Or, once an architecture has been chosen, how do we optimize it for the task at hand? For example, how many hidden layers should the network have and what activation function should be used?

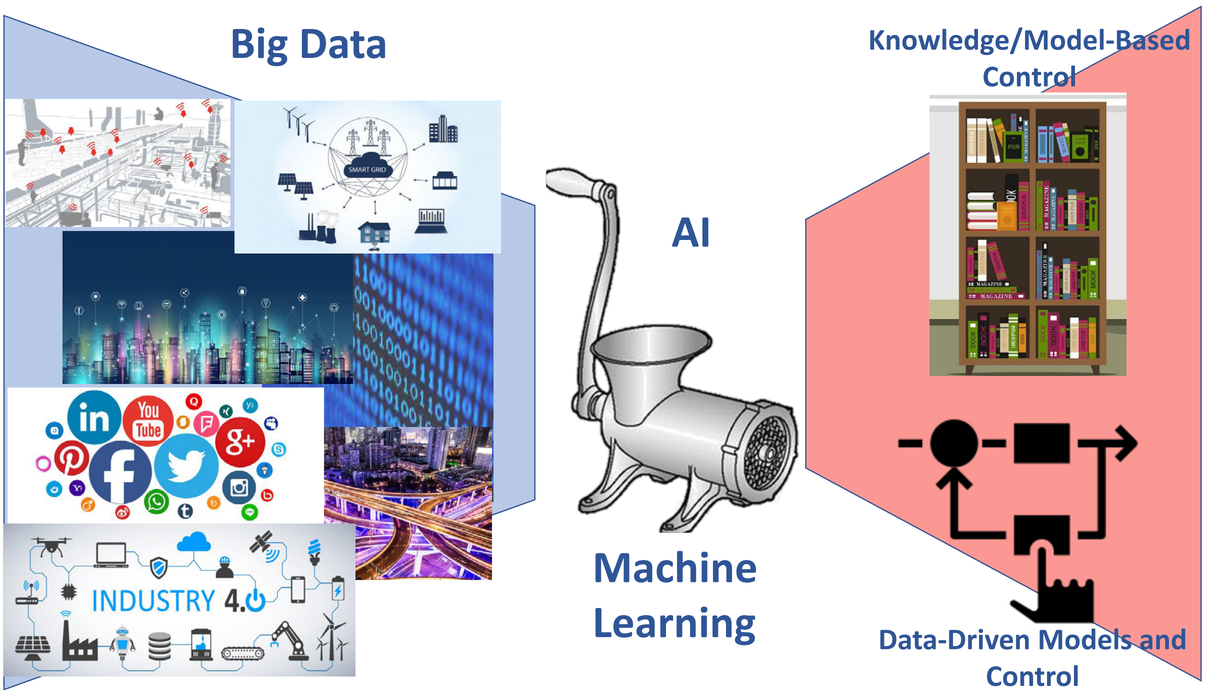


Figure 3.1: Many large-scale applications generate Big Data that has to be distilled by AI and ML into interpretable data-driven models and structured Knowledge Representation and Reasoning (KRR). Integration of these methods for control and decision is a major challenge.

Going a step further, can we design the neural network architecture in a principled manner so that the verification process is simplified? At present, there is no systematic theory for solving such problems as the design space is discrete and unstructured [36]. The control community is well-positioned to contribute to solving such open problems. In particular, co-design and model reduction techniques may offer a path to progress. Co-design aims to optimize some subset of sensor and actuator placement and communication topology [37, 38, 39], while model reduction searches for more simple network structures given an allowable mismatch error [40]. More generally, approaches to control system architecture design (see Section 4.E) may offer deeper insight into designing neural networks.

Physics Inspired Learning: Many of the target applications of AI and ML (such as computational fluid dynamics [41] and climate modeling [42]) involve mapping input/output sample data collected from physical processes or experiments and building a map from the input space to the output space. This must be performed in a way that allows the network to correctly classify unseen data or predict future outputs. One of the biggest challenges in this area is that the models produced (for example, as encoded by the weights and architecture of a neural network) are unable to directly take into account the underlying physics of the process. For example, it is not possible to force the network to model conservation, Newtonian, or thermodynamic laws. As a special case, consider a simple mechanical system that can be "accurately" modeled by Newton's second law, $F = ma$. The same tension between data-driven modelling and system identification arises in control engineering. Surprisingly, one of the most promising methods for achieving physically realistic models shares common ground with the most basic control strategies: work in the frequency domain [43].

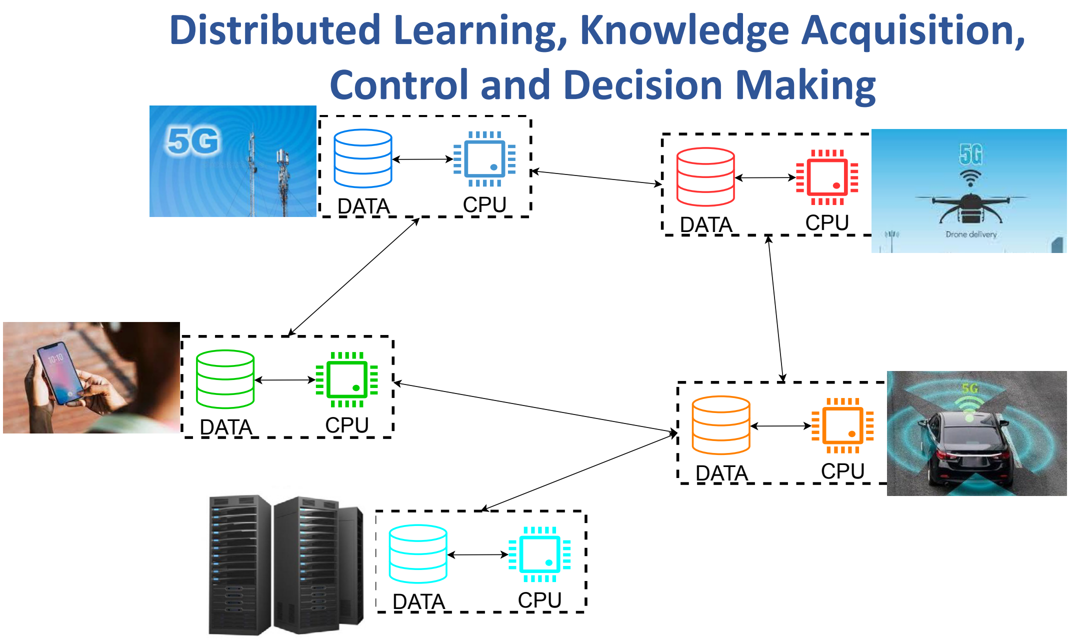


Figure 3.2: Current technology trends like Edge-Computing and Distributed Databases as well as the need for energy and carbon-footprint reduction is pushing for distributed learning, distributed knowledge acquisition, and control/decision architectures beyond federated learning.

Federated Learning: Federated learning (FL) [44] was developed as a means to endow distributed optimization algorithms with properties desirable for situations in which training data is stored on individual devices in order to ensure data privacy [45]. Learning is carried out on each device and, based on a client-server model of communication, only summary statistics from devices are sent to the server. The server then computes an average and broadcasts this average to all clients. Thus, participating devices do not have to share their data, but they are able to benefit from the aggregated model [46, 47]. The enabling features of FL are that it allows for:

i. **Data and device heterogeneity:** Data held across clients need not be independent and identically distributed (IID). Moreover, participating devices will have different computational and storage capacity profiles.

ii. **Reduced communication overhead:** Minimal information is sent between client and server, and there is no client-to-client data exchange.

iii. **Partial participation:** FL algorithms will be resistant to client dropout and failure. This is important as client devices such as cell phones frequently and unpredictably lose connection. There is a clear synergistic relationship between control and FL. Distributed optimization algorithms [48, 49] with the properties described above are necessary for synthesizing distributed controllers for networked systems [50, 51]. Moreover, the communication efficiency and robustness that FL exhibits is something that distributed controllers for large-scale CPS strive for [52]. Conversely, theory in the area of network control can be used to develop and analyze the communication robustness aspects of FL algorithms. As data privacy becomes more important, and the internet of things (IoT) becomes more widespread, the interplay between FL and control will rapidly strengthen.

Simulation to Real-World Robustness: As mentioned in Section 3.A.2, one of the byproducts of AI research, and reinforcement learning in particular, is the wealth of high-fidelity simulation engines with quality application programming interfaces (APIs). These environments allow AI-based algorithms developed from a particular data set to be tested in physically realistic environments. More often than not, when simulated in a more complex environment, algorithms fail spectacularly. Utilizing feedback to provide robustness to model environmental uncertainty is the bread and butter of control engineering. The last few decades have produced a rich framework of robust control frameworks for linear, nonlinear, hybrid, and stochastic dynamical systems. More recently, these developments have come to the model-free paradigm. Incorporating and expanding these techniques into AI pipelines is a clear example of how control engineering can make a direct and substantial impact on AI.

Distribution Shift: ML has been shown to perform exceptionally well in many complex application areas such as computer vision, autonomous driving, and bipedal robotics. However, the core assumption—one that is often broken in practice—is that the data used for training and testing follows the same distribution. There are no guarantees for how a system will perform when faced with an out-of-distribution input [53, 54]. Indeed, many ML pipelines produce systems that fail catastrophically when faced with unseen test data or are deployed to real-world systems. Distribution shift is one of the main obstacles that prevents AI systems from being deployed in the wild. Feedback controllers designed from learning-based approaches have been shown to outperform traditional model-based controllers in many applications [55, 56]. Such outcomes typically occur in scenarios in which the task environments and/or the systems to be controlled are difficult to model accurately.

However, closing the loop around the plant and a learned controller can produce a distribution shift, since the behavior of the controlled system will differ from that of the open-loop system. Control theory has a wealth of tools that can contribute to mitigating the effects of distribution shift, including model validation and experiment design, robust optimization, and safety-based control.

Neurosymbolic Approaches: Neurosymbolic AI can improve how a neural network arrives at a decision by adding classical rules-based (symbolic) AI to the process. Motivated by neuroscience and studies of human perception and cognition, understanding how the human brain's massive neural networks create, store, and use concepts has been a long-standing challenge in AI research [57]. The ability to use symbols is a defining feature of human intelligence. However, neuroscience has yet to explain the fundamental neural circuit mechanisms for flexibly representing and manipulating abstract concepts. Results in fundamental learning theory suggest that the process of learning evolves between steps of learning new concepts and then compacting existing knowledge to develop meaningful knowledge enhancements. Knowledge compaction results in reduction of memory, fast recovery of knowledge and information, and fast comparison of new features and concepts with known (stored) ones.

This is particularly important for systems and control problems, as data and models involve large sets of trajectories, models, and heterogeneities in form and symbolism. Neurosymbolic programming [58] is an emerging area that bridges the areas of deep learning and program synthesis. As in classical machine learning, the goal is to learn functions from data. However, these functions are represented as programs that can use neural modules in addition to symbolic primitives and are induced using a combination of symbolic search and gradient-based optimization. Neurosymbolic programming can offer multiple advantages over end-to-end deep learning. Programs can sometimes naturally represent long-horizon, procedural tasks that are difficult to perform using deep networks. Neurosymbolic representations are often easier to interpret and formally verify than neural networks. The restrictions of a programming language can serve as a form of regularization and lead to more generalizable and data-efficient learning. Compositional programming abstractions can also be a natural way of reusing learned modules across learning tasks.

Distributed Learning and Control: Recent technological advances have led to the widespread use of devices with sensing and computational capabilities, including mobile phones, medical devices, vehicles, and smart sensors. These become sources of abundant data and enable the application of machine learning in a range of previously inaccessible scenarios. In a traditional centralized approach, the data collected by these devices would be transmitted to a central location and there used to fuel a learning algorithm [59]. Such a centralized setup, however, requires the transmission of large volumes of data (potentially overtaxing the network) and may expose the users to privacy risks, especially with data from personal devices.

On the other hand, the distributed learning paradigm presents a promising alternative, in which these computationally enabled devices can cooperate in the learning of a model by processing locally sourced data and communicating with each other. The adoption of a decentralized perspective would allow the control community to build onto the wide literature on multi-agent systems—from consensus algorithms to distributed optimization and distributed control [60]—to design novel distributed learning methods and analyze their performance. In turn, these distributed learning algorithms can complement traditional control techniques by enabling data-based control in distributed, cooperative scenarios [61].

Recommendations

For young researchers: It is now commonplace for students specializing in control to take classes in machine learning, data science, reinforcement learning, and AI. Likewise, these fields can greatly benefit from control systems, both centralized and decentralized. Initial attempts to consider iterative algorithms as dynamical systems have indicated the promise of such a cross-disciplinary approach. Systems thinking and concepts have even more to offer. In the short-term, relevant topics include:

- The development and application of systematic complexity measures
- Rigorous characterizations of hardware and learning architectures and the resulting trade-offs with performance and robustness
- Effects of decision-making across multiple timescales

Longer term goals will need to address the fundamental problems like:

- The distribution shift problem
- The mitigation of the sim-to-real gap
- The incorporation of symbolic (rule-based) AI and simple physics-based models into data-driven ML pipelines

It will be essential to overcome these obstacles as human–machine systems and human-centered AI are investigated from the perspective of control and decision-making. Such efforts hold promise for creating a more integrated and quantifiable theory and methodology for the use of ML and AI in control. This will shape and support the emerging field of trusted autonomy and human–machine systems.

For funding agencies: As humans and machines increasingly commingle in sociotechnical systems (e.g., the metaverse), increased funding can have tremendous impact on all aspects of life, work, society, economy, climate, and health. Short-term needs include:

- Expansion of research portfolios in human-machine systems and human-centered AI for control and decision systems involving both single and multiple agents
- Support for experimental testbeds for human-machine and human-centered AI systems
- The promotion of joint programs with human behavior and neuroscience experts

In addition to the technological and theoretical hurdles that need to be addressed, we need to make a long-term commitment to assessing the benefits *and* risks of AI (and control) technology. Long-term needs include:

- Support cross-disciplinary programs that can address the negative effect of prolonged exposure to social media on teenagers and develop methods to mitigate risks
- Integrated systems science and engineering approaches, methodologies, and software suites to synthesize these complex systems with some assurance of expected and robust behavior

Support in these areas will be necessary to bridge the sim-to-real robustness issues that prevent promising AI systems from being fully deployed.

3.B Electrify Everything

Andrew Alleyne, Jakob Stoustrup, Aranya Chakrabortty, Thomas Parisini, Lucy Y. Pao

> The increased use of electrical energy to support human needs is a long-term megatrend that affects all aspects of society, including transportation, manufacturing, and housing. Control scientists and engineers will play a critical role in accommodating continued growth by optimally managing key energy assets in an efficient and reliable manner.

Abstract Global trends foretell an increased demand for electrified infrastructure to meet the needs of society. Some of this demand is driven by governmental and intergovernmental mandates to reduce our collective carbon footprint, while some is driven by market forces. The percentage of energy needs serviced by electrified systems is increasing steadily with no signs of leveling off, and projected needs cannot be met by managing existing infrastructure. This section advocates for a smart coupling of power generation, storage, consumption, and reuse within key industries like transportation and manufacturing as well as across societal sectors. For example, sectors focused on heating, such as district heating systems, could couple with sectors providing electricity to improve the movement of energy among domains. If appropriately controlled, the resulting system of systems could optimize the utilization of precious energy resources. While the connection and control of these complex systems will not, by itself, accommodate the growing needs of an energy-hungry world, this proposed approach can ensure that the resources consumed will be minimized. Additionally, this approach can empower a well-planned deployment of new resources, thereby reducing the need for investments in new infrastructure. Specific control-relevant topics are suggested for future investment to enable a safe, resilient, and reliable transition to greater electrification.

3.B.1 Introduction

Drivers for Electrification

Electrification means powering a system or a process by electricity, especially when electrical power replaces another energy source. Currently, electrification is taking place in almost every sector. There are several drivers of this trend which vary in importance between sectors. According to the World Economic Forum [62], electrification is critical to decarbonization, which in turn is crucial for reducing the impact of human activity on the global climate. Decarbonization primarily involves substituting fossil fuels with renewable energy such as wind power, PVs, hydro power, etc. Currently, most energy use is *not* electrical. As can be seen in Figure 3.3, only about 15% of global energy use is electrical. Its proportion, however, is clearly and steadily increasing. The drive towards further electrification is driven in part by governmental regulations, especially by governments that have signed the Paris Agreement [63].

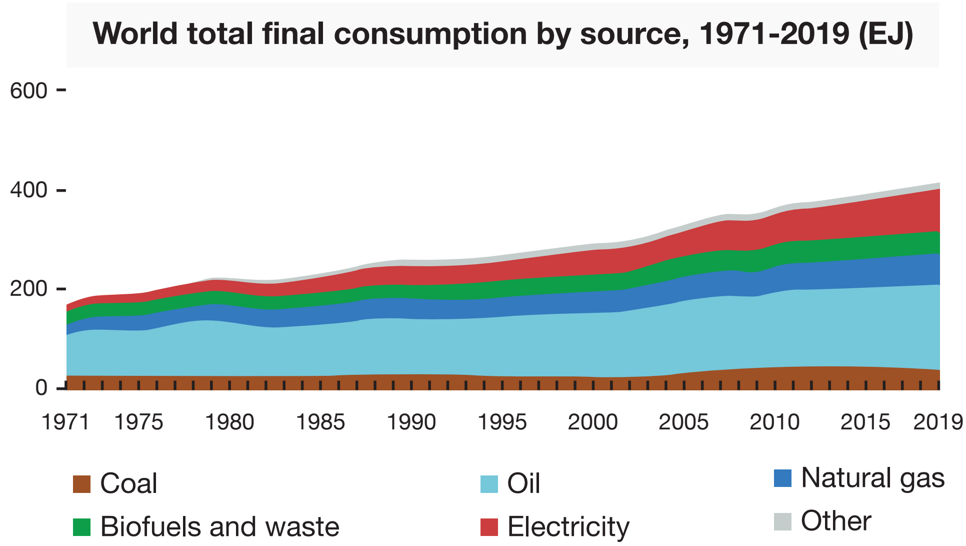


Figure 3.3: World total energy consumption by source. Source: Key World Energy Statistics 2021, International Energy Agency, 2021 https://www.iea.org/reports/key-world-energy-statistics-2021 (reprinted with permission).

In many sectors, the drive towards electrification is also motivated by considerations of performance and productivity [64]. One example is in agriculture, where it has been realized that electrical machinery can outperform conventional machinery in terms of performance, productivity, and safety, as well as environmental impact. Similar advantages accrue across a range of market sectors.

Trends in Electrification

Electrification is already a megatrend in many sectors, including transportation, manufacturing, construction, agriculture, and the energy sector itself. Electrification is increasingly a large part of all areas of life and will continue to grow for the foreseeable future.

Electrification in the Energy Sector

Currently, only a small percentage (~15%) of the world's energy is generated by electricity. This proportion, however, is rapidly growing across the world.

Wind power accounted for 6.59% of global electricity generation in 2021, which is nearly exponential growth from its 1.66% share in 2010 [65]. Wind power is estimated to have a potential for delivering at least 40 times the current worldwide consumption of electricity, which corresponds to at least five times total energy use across all sources. *Photovoltaic (PV)* power also grew exponentially, from 0.15% of global energy production in 2010 to 3.7% in 2021 [66]. In 2020, the global energy consumption was 557 EJ. PV could theoretically supply three to 90 times the world's current energy consumption [67].

As the levelized cost of energy for wind and PV comes down, their growth rates will continue. While the steady-state percentage contribution is unknown, it is clear that their role will continue to increase as a proportion of our global energy mix. Notably, other renewables and low-carbon energy sources such as hydropower and nuclear make up the remaining contribution to electrical generation. However, their percentage of the market over the past 10+ years has not grown at the same rate as wind and PV.

Electrification in the Transportation Sector

Ground vehicle transportation has seen an exponential growth in electrification, from under 0.01% of global car sales in 2010 to over 8.6% of global sales in 2021. This is a vital trend, given that transportation accounts for around one-fifth of global CO_2 emissions. Increased levels of electrification have cut across all modes of transportation, including marine, aviation, rail, and road. Even if full electric propulsion is not the mode of mobility, the increased use of electrical systems onboard these systems has increased tremendously. In marine systems, long-range shipping is still largely reliant on fossil fuel, but that fuel now runs small microgrids to power electric propulsion systems as well as ancillary loads. The result is a more efficient mobility system.

Realizing that energy for transportation is very effectively stored in chemical form, one important trend will be the continued development of technologies that can use electricity (such as wind and solar) to produce liquid fuels (through, for example, electrolysis and subsequent catalysis of the produced hydrogen). One emerging term for this pathway is Power-to-X (P2X). Current investments are in the hundreds of millions of U.S. dollars and this is likely to increase by at least an order of magnitude over the next decade. If successful, this will accelerate the spread of electrification across a wider swath of transportation applications, albeit with some remaining CO_2 emissions cost.

Electrification in the Industrial Sector

Industry employs a wide variety of processes that are currently supplied by energy forms other than electricity. In 2017, electricity constituted 21% of total estimated industry energy consumption [64].

For industrial robotics and machinery, the use of electricity is approaching 100%. The penetration of electrification in other process industries is mainly related to the temperature range of the industrial processes involved. For low temperature processes (like washing, rinsing, and food preparation), electrification has already a high penetration. The same applies to medium temperature processes (like drying, evaporation, distillation, and activation). For high temperature processes (like steam reforming and cracking in the petrochemical industry), there has been a moderate amount of electrification. Low, medium, and high temperature processes that have mature electrical solutions account for roughly 50% of all industrial processes [64]. For very high temperature processes, i.e., above 1,000°C (like melting of glass in furnaces, reheating of slab in hot strip mills, and calcination of limestone for cement production), electrical processes are currently in research or pilot phases but are evolving rapidly as shown in the following example.

One noteworthy instance of electrification making industrial inroads is in green steel production. Iron and steel necessitate massive energy inputs: the iron and steel sector used 33.57 EJ of energy in 2018, and this represents a significant portion of steel manufacturing costs, ranging from 20% to 40% [68]. From raw material preparation to casting, rolling, and finishing, all processes require very high temperatures. For example, blast furnaces and electrical arc furnaces (EAFs) operate around 1500 to 1600°C. When outlining its 2020 technology road map toward more sustainable steelmaking, the International Energy Agency (IEA) suggested direct electrification as one of the core technologies. Figure 3.4 depicts a state-of-the-art steelmaking plant in which two green power-electronics-enabled solutions are implemented, illustrating the significant impact of electrification in the iron and steel industry. Two elements are highlighted in Figure 3.4: i) induction-based reheating furnaces (RHFs) and ii) high-performance real-time control of EAFs. Both solutions are enabled by high-speed nonlinear control of power electronics devices.

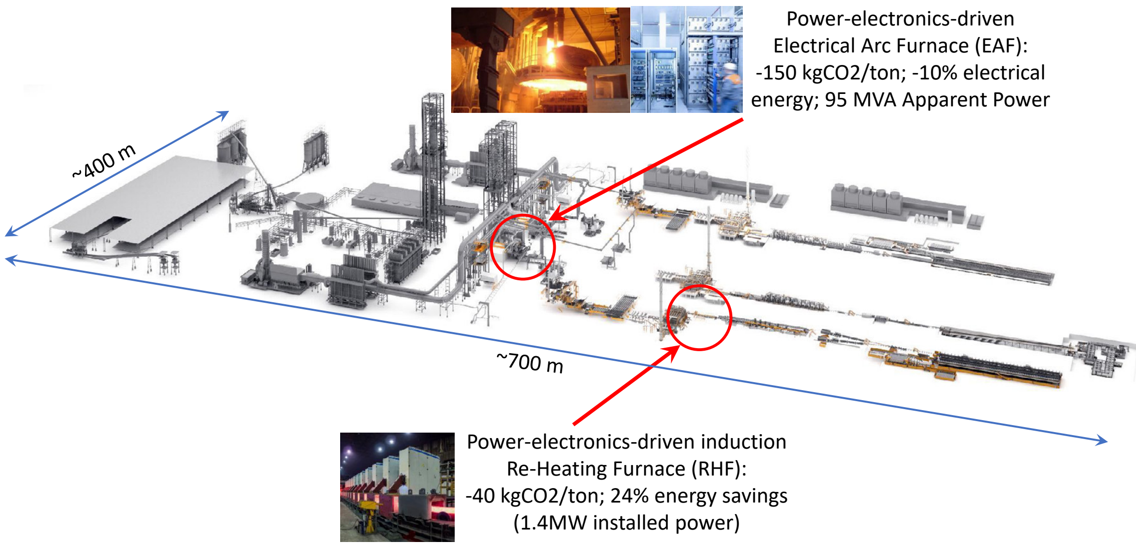


Figure 3.4: Layout of green steel production plant (Courtesy Danieli & C. S.p.A., modified with permission). CO_2 reduction is expressed in kilograms of CO_2 per ton of liquid steel. Two power-electronics-enabled electrification solutions are highlighted.

For the RHFs, an induction slab heating approach is used with an operating discharge temperature of 1,250°C, a power consumption of 320 kWh per ton, and 15 MW of typical installed power. Electric heating by induction makes it possible to directly inject heat in the core of the slab instead of only the surface, thereby increasing efficiency. For the system of Figure B, the electrified RHFs yield a 25% overall power savings and prevent the emission of ~74 kg of CO_2 per ton of steel. The EAFs operate by replacing series reactors, special furnace transformers, and static volt ampere reactive (VAR) compensators with real-time control of the arc current behavior through power electronics current generators. This solution saves approximately 50 kWh per ton of equivalent energy, 10% of electrical energy, and reduces about 150 kg of CO_2 per ton of steel.

The scale of the numbers justifying electrification are impressive. Even more impressive is the introduction and acceptance of electrified systems into industrial processes that were previously considered outside the scope of feasibility for electrification. As the barriers to sector penetration fall with improved technology, electrification will inevitably be everywhere in our future.

3.B.2 Current Status

As discussed, electrification is an ongoing trend in all sectors and is gaining speed in most areas. The number and variety of technologies involved in this transition is mind-numbing. For most of these technologies, control engineering plays an important role, as accounted for in several of the other chapters in this road map. Electrified systems have the obvious advantage of being far easier to integrate with information-based systems than their mechanical or chemical counterparts. Computer-based sensing, communication, and control systems are at the core of our control systems technology infrastructure. They can interface with electrified systems better than other systems since both sensing and command are already in its domain. This is unlike mechanical systems where behaviors must be transduced into electrical systems for sensing or electrical systems must be converted for actuating mechanical components.

One major challenge posed by the global electrification trend is the immense capacity required for new and enhanced electrical infrastructure. Scaling up from 15% of energy use to close to full electrical coverage would require an enormous, and possibly financially prohibitive, expansion of the current electrical infrastructure. To avoid a sudden explosion in required investments and resources, control systems can play a crucial role in facilitating *smart operations and sector coupling*.

Smart Operation of Infrastructure

Smart operations offer the opportunity to reduce needs for excess generation and storage capacity. The principle is illustrated in Figure 3.5. In order to be able to meet consumption needs at all times, a design without smart operations (i.e., control) would lead to capacities for generation and transmission governed by peak consumption values. Any generation source with a high degree of variability (e.g., wind and solar power) would require enormous capacities for excess generation and/or storage. Intelligently controlling the components of the infrastructure can reduce the needed capacities.

Controlling consumption in such a way that peak values are reduced ("peak shaving") would lead to reduced requirements for generation and transport. Similarly, controlling storage could reduce the need for excess generation by storing energy when generation exceeds consumption (and draining energy in the opposite situation). The need for energy transmission can also be reduced by temporal smoothing facilitated by appropriate control of storage capacity. Finally, the controllable part of generation capacity can be used in a similar way to reduce transmission capacity by controlling storage levels appropriately.

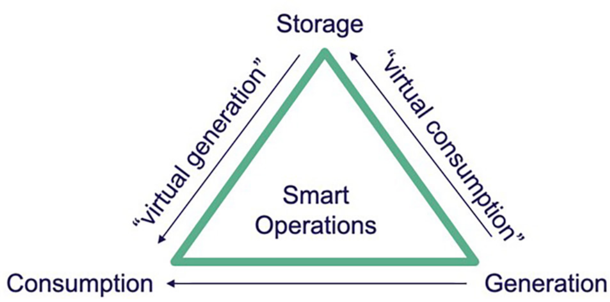


Figure 3.5: Smart operations of an infrastructure. By coordinating operation of generation, consumption, and storage, the total need for generation, storage, and transmission capacity can be reduced.

Sector Coupling Between Infrastructures

Sector coupling has great potential for ensuring that capacity is utilized better across all energy-related infrastructures. To put this into action, an entirely new level of control in the hierarchy might be required. The principle is illustrated in Figure 3.6. Sector coupling is primarily based on energy conversion between two energy-related infrastructures. Converting energy from Infrastructure A to Infrastructure B at times when A has excess generation and B has a deficit can reduce the need for overall excess generation, storage, and energy transmission capacity in the combined infrastructure system. By controlling conversion between infrastructures:

- The capacity needed for generation can be reduced by applying conversion in peak demand situations
- The capacity needed for storage can be reduced by applying storage and consumption capacity in downstream infrastructure
- The capacity needed for transmission can be reduced by applying conversion as "virtual storage" and thereby obtaining temporal smoothing

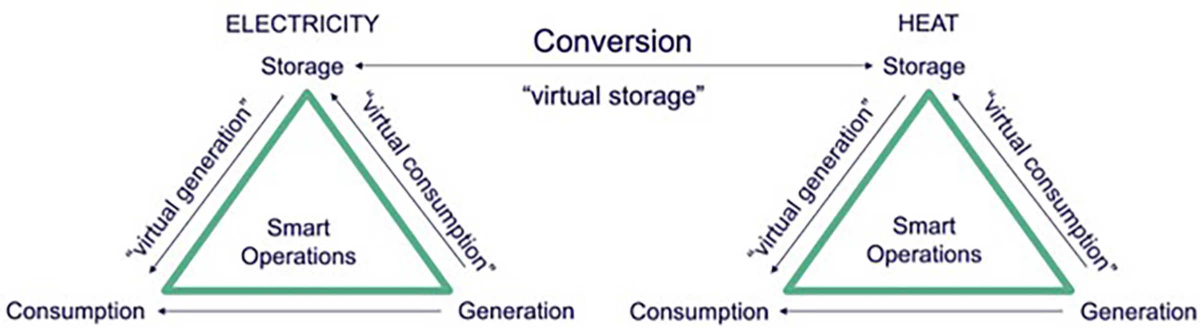


Figure 3.6: Sector coupling. By controlling the conversion of energy appropriately between two infrastructures (here illustrated by electrical and heating infrastructures), the capacity needed for excess generation, storage, and energy transmission can be reduced.

3.B.3 Looking Ahead

Near-Medium Term (5–10 years)

Over the next several years, the current trends in electrification are expected to continue. Many of the individually engineered systems we design, build, and operate will increasingly rely on electricity as a primary mode of energy storage, transmission, and conversion at the point of application.

As a result, the near-medium-term outlook for electrification is strong. Much of the impact will increase the functionality of individual systems within key sectors. In the transportation sector, electrified vehicles (EVs) will become more prevalent and more capable with roughly one-third to one-half of all new car sales being electric cars within the next 10 years. Given the decade-long lifespan of a typical automobile, there will be a balance of both EV and fossil-fuel vehicles well into the following decade. Even without electrification replacing the internal combustion engine for primary propulsion, many of the ancillary loads in vehicles will become electrified, such as steering assist, braking, and HVAC [69]. Similar trends can be seen with other mobility systems such as aircraft as engine bleed air is gradually replaced by electrified actuation for subsystems [70]. The same will hold for the other sectors highlighted at the beginning of the chapter.

Within each sector, we foresee a focus on making electrified systems work well. As market penetration increases, new opportunities will arise. Using the EV example again, there will be an ability to network charging systems, power delivery systems, and vehicles to create a more efficient overall system. There have been many studies on various ways to accomplish this integration of sources, sinks, and storage for EVs and local grids [71, 72]. Considering current consumer preferences and legislation, there have not been enough EVs deployed to provide large-scale validation. Many prior studies rely on assumptions to perform simulation-based optimization and decision-making. The next five to 10 years will see some of these assumptions validated, many of them overturned, and some new pathways emerge that were nearly impossible for the controls field to identify *a priori*.

The improvement of individual sectors due to increased electrification will expand beyond electrified mobility and include manufacturing (such as the prior steel mill example), as well as building systems, cloud computing infrastructure, and other stationary sectors. These elements will have increased ability to coordinate with primary production and storage systems such as PV, nuclear, wind, hydro, and fossil fuels in the near to mid-term. One key to this will be the continuous evolution of the electric grid structure, including micro-grids, as the connective tissue enabling greater collaboration across sectors.

In its simplest form, an electric grid is a network. The use of the term “grid” refers to a complete infrastructure that encompasses power generation, transmission, and distribution. Distribution networks are larger in size in terms of the number of electrical points or buses as well as loads and metering devices, while transmission networks have more complicated topologies as they are often interconnected with other regional transmission networks to provide greater redundancy in power flows. Such interconnections also provide greater reliability in feeding distribution networks. Transmission networks must effectively manage both power generation and consumption as a power failure or spike in consumption in one area may result in adverse effects in another area of the network.

The current electric grid does not optimally accommodate renewable energy sources. Over the near-to-medium term, power grids could resolve these shortcomings by utilizing advanced communication technology and information to optimally transmit and distribute electricity from suppliers to consumers. Examples of such advancements include: i) integrated two-way communication, ii) advanced sensing and actuation, iii) advanced power electronics and their control methods, iv) electricity markets, and v) improved interfaces and decision support to monitor and control the grid in the face of critical contingencies [73]. These functionalities will continue to evolve as existing technologies evolve and new technologies are developed. The type, configuration, and implementation of these technologies and the use of relevant information will also necessitate privacy and cybersecurity of the grid in every operational layer.

Clearly, the role of control scientists and engineers will be critical during this vital transition period to greater electrification. This is an excellent opportunity for the controls field to adopt a “big tent” philosophy incorporating theory, large-scale computation, and experimental implementation. Through this inclusive approach, our field will be able to impact the state of the world as we become more electrified.

Long Term (> 10 years)

Control is poised to play a critical role in furthering electrification and driving meaningful change within and across many sectors of society, as shown in Figure 3.6. Being able to communicate across these various sectors and share energy as sources, sinks, and storage elements will enable a tremendous increase in efficiency and resiliency through appropriate management of our systems. The concept of interconnected networks of networks across different markets will need continued attention to bring to fruition. Examples include using buildings as virtual storage elements through electrified space conditioning to balance out variability in source generation due to renewables. Similarly, automotive systems could be used en masse as actual storage elements aggregating multiple individual vehicles to balance grid loads, stabilize grid frequency, and even shave peaks in generation.

This large-scale interconnectedness will require scalable network-based control design and verification tools that are resilient to anomalies, whether benign or malicious. Each sector network will have its own characteristics. For example, manufacturing systems have a very different set of energy and power profiles than buildings, and EV systems in China may be very different in usage patterns than those in Scandinavia. The power of the control engineer is the ability to understand each sector, formulate decision-making tools to manage the sector subsystems, and connect with other decision-making tools from related sectors. While there are several efforts underway that look at these large-scale networks in concept, there remains much to be done in the long term to deploy guarantees of performance and safety in practice. Meeting, and possibly influencing, regulations and policy will also be an important aspect to cross-sector operation.

3.B.4 Targets of Opportunity for Control Systems Scientists and Engineers

There are tremendous opportunities for control systems scientists and engineers to have an impact on our future electrified society. More electrified and smarter agricultural systems will feed a planet of 10 billion with increased efficiency. This includes smarter in-field agriculture as well as burgeoning vertical farming near urban areas. Similar cases can be made for construction with increased electrification and autonomy. The steelmaking example illustrates that even the most formidable frontiers of manufacturing are becoming accessible to electrical power. The entire domain of mobility, including air, land, and marine modes, will be transformed by electrification over the next decade [74]. The scope of electrified opportunities is sufficiently vast as to preclude detailed enumeration and restrict discussion to an illustrative example.

Here, we utilize two target opportunities, transportation and buildings, as examples of how our field is poised to lead impactful efforts. Buildings are a sink for a large amount of energy in society—up to 40% by many estimates [75]. Within building systems, the modes of energy use various functions: from space conditioning to lighting to computational systems and beyond. Similarly, transportation is a major consumer of energy at nearly 30%. Building systems, if aggregated, can be considered a network of energy nodes with a fixed topology. Likewise, vehicles are a network of energy nodes with a changing topology based on mobility. There are opportunities at scale for aggregating these two networks in a given region into a system of systems to provide greater flexibility and efficiency than individual building or vehicle usage could offer. Building systems can be used to create virtual energy sources through control of their electrical functions. By reducing usage from a baseline, we achieve the equivalent of adding a source.

EV networks can also be used as aggregated storage, source, and sink networks to balance production by renewables or demand from other societal sectors. Indeed, energy could potentially be shuttled among buildings, vehicles, and the grid seamlessly. What are needed are modular and scalable decision-making strategies, including resiliency and contingency strategies, to enable the optimized utilization of these coordinated networks. In concept, this can be done with assumptions on building behavior, grid usage, production cycles, and vehicle operations. Indeed, many within our community have addressed these issues with varying degrees of practicality. Control systems is needed for implementation that firmly grounded in analytical understanding. Of course, this leadership must occur in concert with the market and policy considerations which can often overshadow the technological aspects.

3.B.5 Methods

To date, there are many different tools and techniques that have been used to control electrified systems. Here, we will focus on examining the methods that seem likely to make a large impact in the future.

First and foremost is the idea of learning and data-driven methods to be introduced in Section 4.A. Since energy system infrastructure has long time constants and life cycles, many of the systems that will be used in the future already exist and there is ample opportunity to gather a large amount of historical data. Coupling vast amounts of data with increases in communication and computing capability allows many of the advanced algorithms in machine learning to be performed centrally as well as at the edge. Current ML algorithms, including multilayer deep learning algorithms, are capable of interpolating within their learned dataset to quickly make decisions such as distributed resource allocation within the network of networks, as described in Section 3.B.4. Moreover, emergent "foundational models" for ML offer greater potential to extrapolate outside of the parameters of learned datasets and handle anomalous edge cases better than previous algorithms. Learning and data-driven control, across networks of networks, will be key to successfully capturing the potential of electrification.

Along with the increase in connectivity across electrified sectors of society comes opportunities but also risks. The risk of information misuse and privacy concerns are key. Therefore, deploying the methods that will be introduced in Section 4.C—Resiliency, Security and Privacy—will be vital to achieving greater efficacy and efficiency with electricity. Cross-sector network-based approaches can exploit some of the resiliency-focused methods by explicitly trading off capabilities (e.g., sources) in one sector such as transportation for those (e.g., sinks) in another sector such as manufacturing. This ability to capture and exploit resiliency through the physical assets of a network requires appropriate algorithms to enable benefits. An interesting tradeoff that meshes controls with economics and public policy is the valuation of information from businesses versus that of individuals when integrating these networks of networks. Again, control should be a leader in this space and explore the possibilities alongside policymakers.

Finally, the methods of Architecture and Control introduced in Section 4.E will be vital. As described, a multi-physics network of physical networks will emerge, all connected by communication and a resilient grid. Some type of multilevel hierarchical structure is likely to emerge either by an agglomeration of multiple smaller players or by some type of top-down set of standards. Within this structure, large-scale optimization—likely with data-driven learning—will play a role at the higher aggregated levels. Lower levels will likely use a combination of existing model-based or data-driven methods, with communication up to higher levels. The co-design of communication, sensing, and control architecture along with the design of the physical networks (e.g., charging station placement and grid storage sizing and placement) will be an excellent opportunity for control to lead positive change.

3.B.6 Conclusions

Electrification is currently accelerating across almost every sector of society. This trend is driven not only by the need for decarbonization but also by the promise of improved performance and productivity. Most industries have the potential to shift a large proportion of energy consumption to electricity. However, wholesale conversion to electrified systems presents many significant challenges, including the management of power capacity within realistic targets for infrastructure investment. Control technology will be vital in managing the rapid increase in demand on electricity infrastructure as various sectors progressively pursue electrification. Control technology will be needed in applying smart operation to electrical power grid components to minimize capacity needs for power generation, storage, and transmission. This increased performance can be realized over the mid-term for certain sectors, some faster than others. In the longer term, control technology will play a crucial role in sector coupling, which will further these same goals. We expect that in the more distant future (> 10 years), a smart coupled network of networks will provide services and benefits to the global community that cannot currently be enumerated. Because of these vast possibilities, there has never been a better time to be a controls engineer focused on electrification.

Recommendations

For young researchers: The continued electrification of society is a long-term megatrend that will influence all market sectors. One major technical challenge is the coordination of disparate energy sources, sinks, and storage systems that operate over multiple spatial and temporal scales. Research in control systems is needed to optimize the large-scale utilization and coordination of existing and future capabilities. Additionally, research is needed to co-design controlled operation of these systems with their construction and deployment. This is a broad and deep multidisciplinary control challenge that will push the limits of distributed decision-making, optimization, and computing over disparate networks in the face of uncertainty and potentially adversarial exogenous actors. The systems under consideration will have states and decision variables that number in the millions, all with tight operating constraints. Researchers just entering the control systems field are perfectly situated to have a long-term impact.

For funding agencies: The development and deployment of control systems that support societal electrification will be critical in meeting many of our long-term planetary decarbonization goals. The past several decades have seen significant investment in distributed decision-making over networks, of which electrified societal sectors can be a key application. To further our decarbonization endeavors, we will need a translational effort to move many of the algorithms and approaches that have been developed, with varying degrees of realism, into real-world practice. In addition to the necessary mathematical analysis, this translation should include incentives and policies that promote deployment. We should pursue both incremental and high-risk/high-reward strategies in parallel to rapidly determine what will work in practice. These funding efforts should be maintained over a significant period of time in order to address new challenges that emerge during deployment. Governmental investments in translational research will need to be coordinated with commercial investments to optimize the deployment timescale.

3.C Engineering Biology

Neda Bagheri, Mustafa Khammash, Richard Murray

> Over the last 20 years, engineering biology has emerged as an exciting domain that leverages engineering tools to optimize the design of beneficial biological systems. Control theory has the potential to advance applications of engineering biology in a wide range of areas, including agriculture, healthcare, and energy.

Abstract The field of engineering biology (also referred to as biological engineering or synthetic biology) seeks to arrive at a better understanding of biological systems and to design components into biological circuits that can carry out useful functions. This section of the road map describes some of the opportunities for control theory to benefit engineering biology and articulates some unique challenges posed by this application area. We focus on engineering at the cellular and molecular level, highlighting potential areas of research at the intersections of the biological sciences and various fields of engineering. These areas of interest include industrial and environmental biotechnology, food and agriculture, health and medicine, and energy. Key challenges include multiscale, stochastic, and spatiotemporal modeling of biological systems; modeling and managing robustness and uncertainty; designing tools that better support composability, modularity, scalability, and interconnectivity; and designing more complex, multicellular systems. Emerging applications include metabolic engineering and industrial biotechnology, engineered living materials, personalized medicine, control of living neuronal cells, and environmental remediation and sustainable agriculture.

3.C.1 What Is Engineering Biology?

Engineering biology refers to the use of engineered genetic elements to implement a variety of operations. These operations range from environmental sensing and remediation to biochemical synthesis to understanding, detecting, and preventing disease. A typical workflow for engineering a biological system involves the design of a biological circuit (or metabolic pathway) to carry out a specific function, which can then be "compiled" into a set of DNA sequences that represent the various proteins and other molecules required for implementation (see, e.g., [76, 77]). DNA can be commercially synthesized and either transformed into living cells where the genetic program is executed or be implemented in cell-free environments that carry out the core processes of biology (e.g., transcription and translation) in a nonliving environment. One example of this workflow is shown in Figure 3.7.

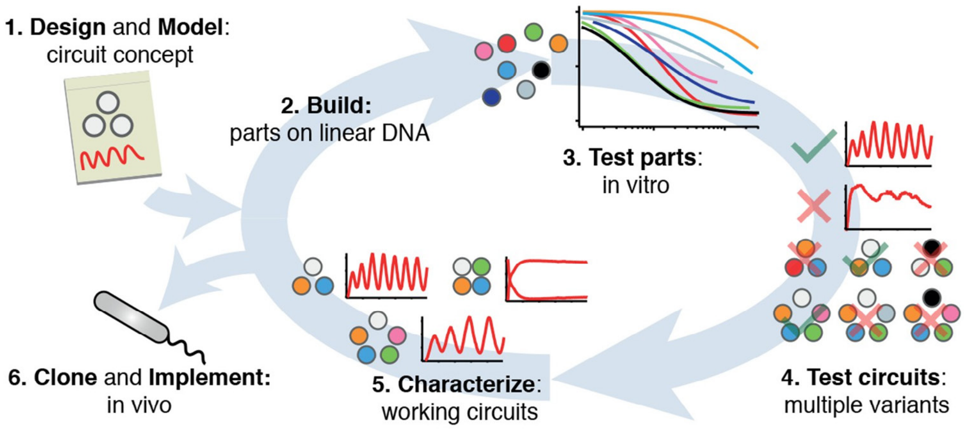


Figure 3.7: Design flow for engineering biological circuits (from Niederholtmeyer and Sun et al, 2015 [77]).

As in other engineering disciplines, engineering biology involves iterations of design, implementation, and testing (referred to as the design-build-test-learn, or DBTL, cycle), including the use of modeling and analysis tools familiar to those in the field of control. Feedback plays a critical role in the operation of natural biological systems, but its use is still nascent in engineering biological systems. Engineering of biological systems can occur at a variety of different scales (from molecules to ecosystems) and can also be combined with non-biological components, such as silicon-based computing. While much of the field of engineering biology is focused on providing useful devices and systems, tools from engineering biology can also be used to provide a better understanding of natural systems and predict emergent dynamics across scales.

The 2019 Engineering Biology Research Consortium (EBRC) Research Roadmap [78] (Figure 3.8) summarizes some of the applications of engineering biology that are being actively explored:

- **Industrial Biotechnology** focuses on the industrial use of engineering biology, including manufacturing of bio-based products, making sustainable manufacturing processes cost-competitive, accelerating innovation and discovery of new products and technologies, and generating products at scales necessary for economic viability.
- **Environmental Biotechnology** focuses on technologies to enable deployment of bioengineered systems on land, in air, in water, and across human landscapes for the purposes of remediation, natural resource management, environmental monitoring, and species management.
- **Food and Agriculture** focuses on tools that impact how we feed the Earth's people and animals. Engineering biology provides unique means for supporting growing populations with more and different types of food, addressing changes to food security and demand, and reducing the impact of climate change and urban growth.
- **Health and Medicine** focuses on applications of engineering biology relevant to the well-being of human and animal populations. Applications of engineering biology in this sector focus on preventing and eradicating disease and supporting longevity and quality of life.
- **Energy** focuses on the application of engineering biology technologies to advance clean and affordable energy sources and to reduce overall energy consumption. Biology can be a source for renewable energy by providing biomass for electricity generation and for the production of highly energy-dense transportation biofuels. It can also be used to optimize processes to use less energy.

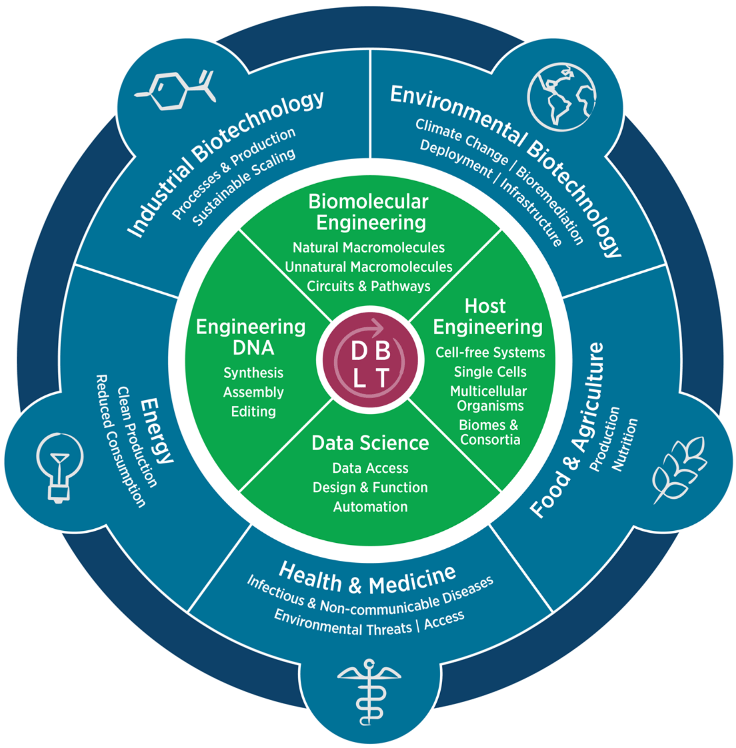


Figure 3.8: EBRC Engineering Biology Road Map [78].

Advances in these application areas have realized in recent years through many new techniques, including the CRISPR/Cas9 system that facilitates genetic manipulations in a wide variety of organisms [79], the development of gene drives that allow the propagation of genetic changes through a population [80], and the rapid increases in capacity for gene sequencing and synthesis [81].

Engineering biology has many potential ethical, legal, and safety considerations, and these are active areas of discussion in the engineering biology community. The EBRC has developed a set of guiding ethical principles for the field to [82]: i) use engineering biology to benefit the world, ii) weigh benefits of research against potential harms, iii) incorporate justice into all aspects of engineering biology, iv) share research, v) protect the freedoms of individuals and researchers, and vi) support open communication between researchers and other stakeholders. These principles build on a well-established set of guidelines and regulations in various jurisdictions that have been established to allow the safe manipulation of living organisms in a manner that protects human health and the environment (see Chapter 2 in [83] for a review of the U.S. Coordinated Framework for the Regulation of Biotechnology).

3.C.2 Contributions From Control to Engineering Biology

Several recent articles have surveyed the contributions of control theory to the field of engineering biology, including Del Vecchio et al. [84] (see Figure 3.9), Hsiao et al [85], and Khammash et al. [86]. As with other engineering domains, high-level feedback serves as an invaluable tool for providing robustness to uncertainty and designing closed-loop dynamics to match performance needs. It also serves as a mechanism for enhancing modularity by providing desired dynamics of a module or subsystem (which are relatively independent of the detailed dynamics of other modules it is connected with) and by mitigating the negative effects of host-cell resource sharing. Control-theoretic perspectives provide useful modeling, analysis, and design tools for a variety of applications.

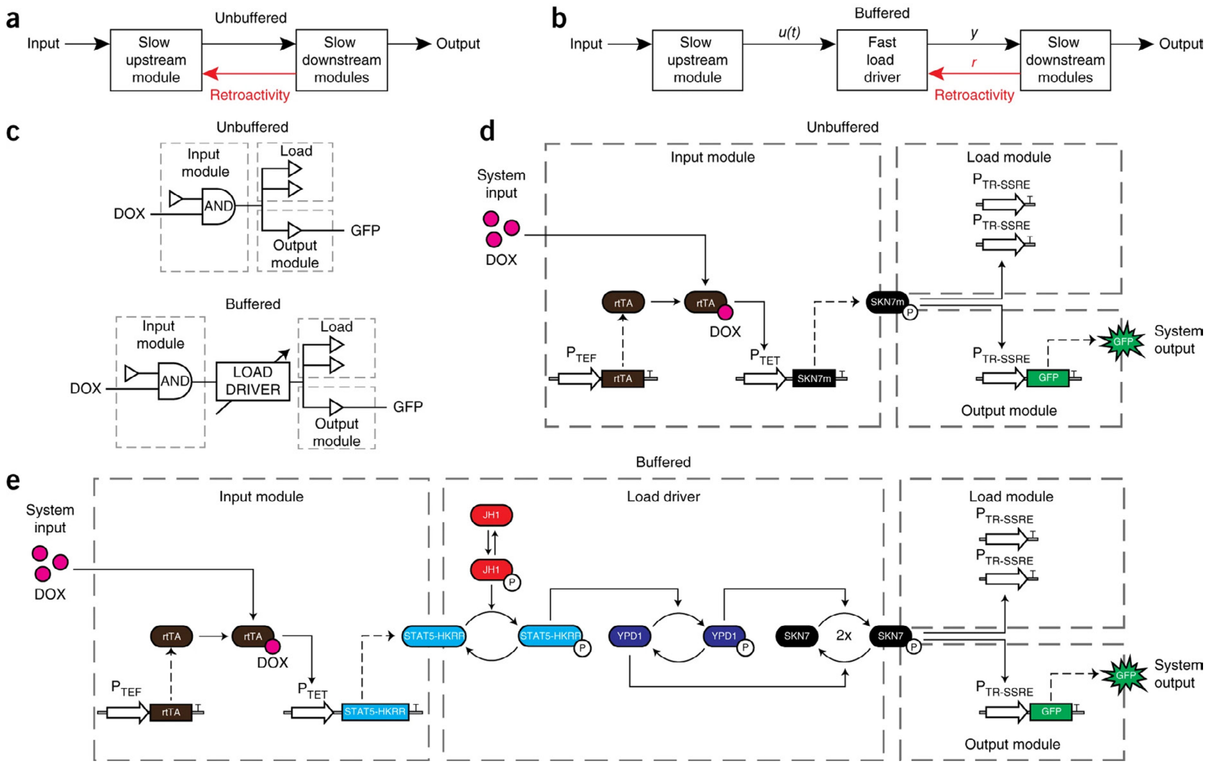
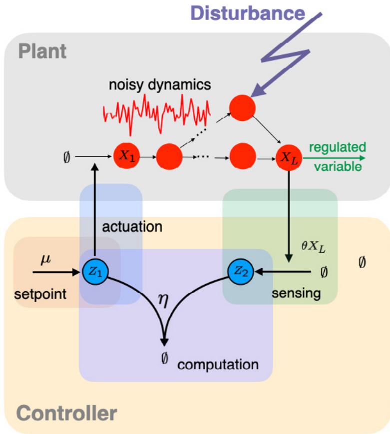


Figure 3.9: A load driver device for engineering modularity in biological networks [87] (reprinted with permission).

Biological systems are inherently noisy and highly variable from cell to cell and organism to organism. This presents a challenge in implementing biological controllers but also an opportunity to use these controllers to mitigate the effects of noise and variability. Stochastic modeling of the dynamics of biomolecular circuits makes use of a rich set of tools from random processes and stochastic simulation. An important difference from other disciplines is that Gaussian processes are often not a good fit for the underlying stochastic mechanisms. Stochastic simulations and reduced-order stochastic models have been developed that build on stochastic differential equations (chemical Langevin equation) and Kolmogorov-based techniques (chemical master equation). Feedback control mechanisms in cells can provide a path to attenuating and even exploiting noise, but the fact that the control components themselves are highly variable (and that accurate sensing mechanisms are in short supply) must be taken into account.

One especially important use of feedback and control concepts is in understanding and minimizing the burden and loading ("retroactivity") effects of engineered biological circuits in cells. By making use of feedback, it is possible to isolate upstream components from downstream loads, manage the use of cellular resources, and provide adaptive designs to modulate the expression of genetic elements depending on the resources available.

Another important concept in the construction of feedback control systems is the use of "antithetical" mechanisms for providing feedback signaling [88] (see Figure 3.10). These mechanisms have been shown to implement molecular integral feedback control that is "resistant" to cellular noise. The incorporation of this integral feedback endows the biomolecular system with the important property of robust perfect adaptation, enabling setpoint tracking and constant disturbance rejection that is robust to parameter and dynamic uncertainty.



Antithetic integral control architecture

Figure 3.10: Antithetic feedback [88].

A more recent approach to engineering biology is the combination of genetically controlled systems with computer-controlled systems. These systems combine the advantages of biomolecular sensing and actuation with the advantages of modern computation. Computer-controlled genetic systems can be used for rapid prototyping of designs, integrated "silicon-cell" devices, or various applications in personalized medicine (e.g., using biomolecular sensing and actuation coupled to MPC using a physiological model of the overall system).

Control techniques have also been useful in diagnostics and therapeutics for diseases and other physiological conditions, particularly in the context of personalized medicine (see also the related material in Section 2.B). Models of the pancreas, coupled with MPC have enabled accurate real-time prediction and management of glucose in patients with diabetes [89]. Optimal control has been applied to patient-specific pharmacokinetics and pharmacodynamics models to design strategies for the administration of drugs to manage chronic myeloid leukemia [90]. This complements the pioneering work that engineered biology facilitated in the design of chimeric antigen receptor (CAR) T-cell therapy for managing soluble tumors. Similarly, nascent spatial models of the solid tumor microenvironment are being developed to predict in vivo tumor growth and coupled vascular dynamics, providing an in-silico testbed to help uncover the rules that drive emergent dynamics and to evaluate the efficacy of therapeutic interventions before treating patients [91]. Large, high-resolution models (such as those used to characterize the microenvironment) can prove computationally expensive for optimization and control. Model reduction tools from control theory, in combination with machine learning, are being integrated to alleviate this burden. Despite great advances in prediction, there remains a rich space for traditional nonlinear ordinary differential equation (ODE) models to guide our understanding of the human body and serve as tools for hypothesis generation. Muldoon et al. trained such a model to uncover how immune cells activate and drive their own population dynamics [92].

3.C.3 Challenges and Opportunities at the Intersection of Control and Biology

As illustrated in Figure 3.8, there are many opportunities for control to have a substantial impact on biological systems. These opportunities include achieving a better understanding of biological systems in nature and designing and implementing new circuits, pathways, and systems from biological components. They also include integrating engineered biological systems with other technologies, such as electronics, materials, and mechanical systems. These opportunities are enhanced by recent technological advances, especially the development of CRISPR-based technologies [79] and the increased ability to synthesize custom DNA [81].

Looking across these different prospects, there are a number of common challenges that must be overcome and that can lead to new insights and methods for control. In this section, we summarize some of these methodological challenges and some of the opportunities for using control in emerging engineering biology applications.

Modeling Stochasticity and Uncertainty In Biological Networks: A key feature of biological systems is the large amount of stochasticity and variability that is present over a wide range of scales [93, 94, 95]. Many of the tools developed in control theory for understanding stochasticity may need to be further developed to deal with highly non-Gaussian stochasticity (e.g., Poisson processes for molecular binding and multimodal distributional responses). Feedback provides a mechanism for reducing variability when desired, but it can also be used to tune the distributional response to exploit the ability of biology to maintain "bet-hedging" strategies that protect against large changes in the environment. So far, reasoning about the evolution of distributions of responses in these cases has been tackled using numerical approaches that are not necessarily tuned to design needs. Additional research is needed to provide useful mathematical frameworks and tools for analysis and design.

Multiscale and Multi-Physics Modeling of Biological Systems: Biological systems operate on a wide variety of spatial scales (from molecules to ecosystems) and time scales (from microseconds to centuries), and the interaction between these scales is both intricate and essential [96, 97]. Effective modeling, analysis, and design must consider biological, chemical, informational, and physical properties across several scales, with strong interactions between them. To engineer complex biological systems that rival the designs we see in other domains (aerospace, chemical processing, electronics, etc.), we may need mathematical frameworks that allow for different layers of abstraction in which these different aspects can be partially separated. While our experience from these other fields will be relevant, natural biological systems are integrated to such an extent that it is difficult to separate out these effects, which in turn limits our ability to effectively combine components and subsystems into large-scale, complex networks.

Spatial and Spatiotemporal Properties: A simplistic view of a biomolecular feedback system is that it is a set of chemical reactions that take place in a well-mixed environment. Many existing engineered biological systems make use of this simplifying assumption as a basis of operation. The distribution of biological functions across the spatial domain, and the coupling of spatial and temporal responses, are important avenues for future advances. Early applications might include decentralized control of well-mixed cell populations, but more advanced applications will involve spatiotemporal strategies that include transport, signaling, and vastly heterogeneous properties among engineered elements of a complex system.

Robustness and Predictability: A key property of feedback systems is their ability to provide robustness by design as well as performance (fast response, small errors, etc.). A critical factor in biological systems is context dependence (resource sharing, noise, growth conditions, genetic context) and new methods are needed to overcome this for engineered biological circuits. Additional issues include new methods to avoid mutational escape and minimize the effects of burden (e.g., low copy number control). Many of these issues are unique to biological systems and hence new control paradigms may be required.

Better and Faster Biological Design Workflows: One of the major drivers of current results in engineering biology is the improvement in the DBTL cycle. Tools for rapid prototyping and testing have enabled increasingly fast design cycles as a means to iterate and explore the design space. As our systems get more complex, we must develop more sophisticated workflows that allow rapid development and deployment of working systems, enabled by modeling, simulation, and testing. The need for more automated and repeatable design workflows can be enabled by the use of tools from control and the perspectives provided by the control community.

Composability, Modularity, Scalability, Interconnectivity: In order to build biological systems that can operate in real-world environments, it will be necessary to create components and subsystems that can be combined and function as expected. Current biological components do not have this property: It is very difficult to get a circuit or pathway produced in one laboratory to function in another laboratory, and even harder to integrate a circuit or pathway produced by one group into a system being built by a different group without a substantial amount of rework. In other disciplines, feedback is an enabler for providing modular systems that provide desired dynamics in the face of uncertainty, including uncertainty in the other systems to which they are connected. It is likely that feedback will similarly provide one of the key tools to enable composition, modularity, scalability, and interconnection in engineering biology.

Design of More Complex Functionality: Current use of feedback and control in engineering biology is relatively rudimentary compared to the advanced control techniques that we see in other disciplines. Many engineered biological control systems consist of simple feedback loops with only a few elements. As we explore more complex biological applications of control, it will be necessary to develop more complex functionality that allows for filtering, estimation, prediction, adaptation, optimization, and other advanced functions. This will require substantively new approaches that utilize modular and programmable architectures for implementing complex functions.

Design and Control of Multi-Cellular Systems: To date, most of the work in engineering biology has looked at relatively simple systems in which a single species is engineered to perform a desired function. Nature makes it clear that multicellular systems are capable of a vast repertoire of functions. Tools from control theory can provide key insights into the interactions between different types of cells in a large system, as well as design methods that allow for the separation of functions between different cell types [98, 99]. In particular, principles from decentralized control can elucidate how a subpopulation of cells without a master decision-maker can coordinate and drive higher-level responses. The "systems engineering" of biological systems will eventually have to grapple with the types of problems that control practitioners in other disciplines face every day.

In addition to these broad challenges in control-oriented techniques, there are a number of emerging applications where biological control engineers can play a critical function.

Metabolic Engineering and Industrial Biotechnology: Feedback control can provide new methods for regulating gene expression to optimize metabolic pathways that enable increased yield, are robust to substrate variations, and are capable of handling toxic intermediates. Such regulation of gene expression can be accomplished with external feedback loops on a bioreactor using process control techniques, but there are also opportunities to couple bioreactors with circuits at the genetic level. Applications range from high-value chemicals to advanced biologics (e.g., using Chinese hamster ovary [CHO] cells to make biologics with feedback control to enhance yield).

Engineered Living Materials: Recent investments in developing engineered living materials have enabled materials to respond in diverse ways, including reacting to environmental stimuli and damage. Control will also be useful for regulating the function of soft devices that integrate biological materials, like muscle and nerve cells, with non-biological material. By carefully regulating developmental processes, control methods can be used to guide the engineering of new tissues and organs.

Personalized Medicine: There is increasing interest in using "designer cells" to produce or deliver therapeutics. For example, recent results by Krawczyk et al. [100] describe a bioelectronic interface that uses wireless-powered electrical stimulation of cells to promote the release of insulin. Other examples included engineered microbes that can be used to sense and/or modulate conditions in the gut microbiology as well as in broader applications, such as tissue engineering. Better understanding of the microbiome and how it impacts the immune system and brain function is needed, and that will require a systems perspective. Similarly, new methods in stem cell differentiation may be able to make use of control systems to regulate differentiation.

Control of Living Neuronal Cells: Recent efforts in developing invasive and noninvasive technological solutions for the direct control of living brain cells provide a promising path for the contributions of control engineers. One recent revolutionary advance is the development of optogenetics [101], a technique that uses genetic modifications to enable the use of light (via fiberoptics) to both sense and actuate specifically targeted neurons. By genetically engineering organisms such as insects and mice, neurons can be modified to express light-sensitive ion-channels [101], allowing for the control of neuronal activities (e.g., turning a selected neuron on or off for the purpose of neuromodulation [102]). This approach allows for an increased understanding of neuroscience and potential therapeutic applications as well as pathways for controlling the behavior of (biological) neural networks.

Environmental Remediation and Sustainable Agriculture: Similar to the use of engineered microbes in the gut microbiome, engineered microbes in soils can be used for a variety of purposes. These include the detection and remediation of toxins as well as the enhancement of growth conditions for crops through the detection of soil conditions and the triggering of plant responses by the release of small molecules. The use of engineered microbes in soils will require circuits and pathways that carry out a desired function while being robust to potentially large environmental fluctuations (temperature, water content, pH, etc.). It will also require new means of ensuring the safety of open release (versus contained use) products [83]. These efforts are closely related to climate change mitigation objectives, as described in detail in Section 2.A.

Engineering biology presents many key challenges and opportunities for the control community as we continue to work in conjunction with biologists, bioengineers, and data scientists.

Recommendations

For young researchers: With historic roots in aerospace, chemical, electrical, information, and mechanical engineering, control theory integrates skills across STEM disciplines to enable broad and impactful applications. This places control in an excellent position to make significant contributions to emerging areas in biology. More students and faculty need to get involved in engineering biology to ensure that the control community can inform this evolving field as it expands in the coming decade. New courses, collaborations, and ideas are needed to connect existing results and methods in control theory with new applications, as well as to expand control theory to encompass specialized methods for biological systems.

For funding agencies: Investments in engineering biology should be increased over the next decade to accelerate innovation and expand the scope of the bioeconomy, which is already recognized as an important area of focus for the global economy. The relevance of engineering biology to health, wealth, and the environment—and the critical role of control theory in designing large-scale systems that can operate in uncertain conditions—justify more investment in research that can engender meaningful advances.

3.D Robots in the Real World

Dimos V. Dimarogonas, Dawn Tilbury, Karl H. Johansson, Tor Arne Johansen

> With major advancements in sensing and computing technologies, robots have an increased ability to perceive and reason about their environments. Control challenges include promoting resilience in unstructured and uncertain domains, proving correctness of closed-loop operations under large uncertainties, and ensuring safe operations around humans.

Abstract Recent advances in mechanisms and batteries have enabled the development of robots that can manipulate objects as well as drive, walk, and fly around their environments. Mobile manipulators, with arms and legs or wheels, can pick up packages and move them across the room or across town. Robots have the ability to operate in dangerous environments like construction sites, disaster areas, and even outer space. However, three overarching control challenges must be addressed for this intrepid new generation of robots to successfully accomplish useful tasks in the real world. First, real-world environments are dynamic and unstructured, unlike traditional manufacturing environments where operations are preplanned and repetitive. Thus, control strategies must promote resilience to unforeseen conditions and disturbances. Second, traditional robust control addresses structured uncertainties and is overly conservative in realistic domains. Given the large uncertainties that exist in real-world environments, closed loop-controlled systems must be able to operate in a provably correct manner. Third, robots must operate safely around humans and engender their trust. Robots made out of soft materials that move slowly are inherently safer, although potentially less accurate and productive than their hard and fast counterparts. The future will see multirobot systems deployed in complex environments, with distributed computation and communication to achieve their ultimate goals.

3.D.1 Introduction

Robots began to be introduced into manufacturing plants in the 1960s. They were able to perform repetitive tasks with increasing accuracy and precision, relieving human workers from dull and sometimes dangerous labor. With high power, robots could lift entire vehicles, and with high speed, they could pick and place small parts. However, these powerful and fast robots could not sense and react to their surroundings, so for most of the 20th century, robots were placed in cages to protect workers from unexpected encounters that could result in injury or death.

In the 21st century, robots are emerging from their cages and operating in the wider world. With the ability to sense their environment, they are mobile and are able to work with and alongside people. Advances in computing, communication, sensing, and power storage are accelerating this transformation.

Service robots, autonomous vehicles, UAVs, surgical robots, and field robots are operating in less structured environments while taking on more complex and safety-critical tasks ranging from everyday chores and societal applications to warfare to extraterrestrial exploration.

For the purposes of this section, a robot is understood to have sensors to observe the surroundings and measure its position, as well as perception to attain an understanding of its environment. It is able to reason about its options, formulate decisions, then take actions to interact with its environment, including other robots. A simple robot would be a robotic vacuum cleaner, while a self-driving car is a much more complex example.

3.D.2 Enabling Technologies

Perception

Improved sensing and computing technologies have enabled great advances in perception and opened up new perspectives for control. Multimodal sensing, which integrates audio and video signals with laser scanning and haptic feedback, shows promise for further enhancing robot perception.

In relatively open outdoor environments, robots and vehicles can often localize themselves with the aid of global navigation satellite systems (GNSS). Nowadays, this can often provide centimeter-level absolute accuracy with real-time kinematics and differential correction services. However, GNSS services may be denied or degraded as a consequence of multipath signal obstruction, electromagnetic interference, or cyberattacks. While inertial navigation is commonly used to provide high-rate motion estimates and resilience to short-term GNSS degradation and outages, additional aiding sensors are needed when GNSS is denied for long periods of time. Local navigation aiding based on radio and acoustic waves is helpful, but it depends on infrastructure that must be installed.

The use of visual odometry and simultaneous localization and mapping (SLAM) methods based on imaging sensors (visual and infrared cameras, lidar, etc.) is a very active area of research for robotics and autonomous vehicles. These sensors can be used in unknown and unexplored environments that have sufficient visible features, without any supporting infrastructure. They also provide situational awareness through the detection, recognition, and tracking of features and objects in the environment, such as roads, landmarks, unmapped static, and dynamic obstacles. Such information is used for relative and absolute navigation aiding, collision avoidance, and autonomous control that can direct the robot toward features or objects of interest while exploring its environment. Perception of the environment can be incorporated into various spatiotemporal models, maps, grids, graphs, and 3D world representations—often called “digital twins.” This is the basis for perception-based control, in which sensors explore the unknown environment and simultaneously collect data for end-user needs. One key application is information-driven control, with adaptive sampling of spatiotemporal fields to detect hotspots with high concentrations of contamination or valuable resources.

While significant progress has been made in perception-based control over the last decade (e.g., [103, 104, 105, 106]) and robots have been proven to perform better than humans in certain tasks and environments, it is also clear that humans have a holistic understanding of the world around us that we are not yet able to integrate into robotic world models. The human brain far exceeds a robot’s capability to integrate its perception with its model of the world. This leads to a limited understanding of risk, incomplete situational awareness, and less accurate predictions.

Thus, it is still rare to see full autonomy: Some level of human supervisory control, monitoring, or assistance is often needed. Machine learning techniques such as reinforcement learning and deep learning are increasing our ability to implement better perception-based control for robots and vehicles. One key remaining concern is the limited transparency of such methods and models, whereas important questions related to explainable AI and trusted AI remain relatively open.

Manipulation

Over the last decade, the mainstream focus on robot manipulation has shifted toward soft robotics [107]. This concept seeks to overcome the inherent limitations of traditional rigid manipulators equipped with application-oriented control algorithms (such as impedance controllers or inverse kinematics control laws) as well as the more recent flexible/compliant manipulators that employ variations of these algorithms to cope with extra-passive degrees of freedom. Soft hardware promotes safety in interactions, but it comes with its own limitations [108]. Advanced considerations, such as disturbance rejection and motion imitation, are directly impeded by hardware design.

Nevertheless, this has led to an increasing demand for advances in the limited interaction capabilities of current controllers for non-soft manipulators. Different control techniques focused on efficiency and robustness have been explored. These include transient controllers (such as barrier-based techniques and prescribed performance control), model-based controllers with enhanced disturbance rejection capabilities (such as robust nonlinear model predictive controllers [NMPCs] and adaptive controllers) and learning techniques to mimic the human motion for prescribed tasks (such as in learning from demonstration). These diverse research lines will result in controllers that are capable of performing increasingly delicate tasks with basic perception capabilities [109]. Efficient interaction with the environment, and with other human and/or robot agents, is expected to shake up our perspective on which applications are viable for technology maturity [110]. Instead of controllers for single robots substituting for human operators in mass production facilities or collaborating in limited tasks with conservative safety guarantees, control solutions for efficient human-robot collaboration should make robot manipulation become more accessible to the general public. This also applies to mobile manipulation, which introduces the additional challenge of dealing with the coupling between the mobile base and the robot manipulator, as well as the computation for extra degrees of freedom of the base [111].

Legged Robots

Legged robots have recently transitioned from the laboratory to real-world use cases. This is due in large part to the large number of commercial platforms for both quadrupedal and bipedal robots. With the increased availability and deployment of these robots comes multiple opportunities for the control community. From the modeling side, this includes hybrid systems models of walking. From a control perspective, the real-world deployment of legged robots requires robustness of these controllers to uncertain environments (e.g., different terrain types). Finally, perception must be added to the feedback loop to enable legged robots to reason about the world they move through.

Unmanned Aerial Vehicles

Unmanned aerial systems (UASs) have matured significantly during the last decade and are now extensively used for a wide range of applications. While miniaturized rotary-wing and fixed-wing aircraft were dominant in the early days, unmanned multi-rotor aerial vehicles are most common today. Their popularity is due to their low cost and the simplicity of their operation for many applications such as aerial photography and infrastructure inspection. Unmanned fixed-wing aircraft still have much greater speed, range and endurance, and hybrid designs that combine vertically and horizontally mounted propellers are emerging in order to enable user-friendly vertical-take-off-and-landing (VTOL) capabilities while retaining the advantages of fixed-wing aircraft. UASs are still mostly controlled using a mix of manual remote control and automatic control with preplanned routes. The need for expert pilots is being gradually reduced, which makes UASs available to an increasing number of users.

Going beyond visual-line-of-sight (BVLOS) requires safe systems and procedures that limit the risk for collision and damage. This has impacts for regulation, air traffic control, certification of airworthiness, pilot competence, and the concepts of operation. While high levels of autonomy may be perceived as a risk by stakeholders today due to the technology's relatively immature state, it also has the potential to be an efficient solution for ensuring a high level of safety. There is a need for extensive research on autonomy and related technologies to enable BVLOS flight in challenging and unfriendly environments, e.g., without satellite navigation dependency [112], in strong winds and turbulence, icy conditions, and darkness, while providing the necessary resilience to internal faults and unexpected external disturbances.

The UAS industry also wants to extend beyond flying cameras and sensors (e.g., for mapping, surveillance, and inspection) to more heavy-duty tasks such as transportation (e.g., persons, medical supplies, and cargo) and manipulation (e.g., maintenance, repair, and construction) [113, 114]. Consequences of faults will increase, and research on accurate and resilient sensing, control, and autonomy will be more important than ever.

Digital Twins

For decades, simulations have been used to evaluate potential robotic solutions, and data has been gathered for use in control systems to make real-time decisions about following actuations. Advances in high-performance computing, high-speed networking, and high-volume and cloud data storage allow these simulations and data to be used in new ways to improve control of robotic and automation systems. Digital twins is a recent concept that encompasses multiple aspects of physical counterparts to help make decisions to improve operation outcomes.

The term digital twin has been used in many domains. Here we adopt the definition of [115] that a digital twin is some level of replica of a real thing, exists in the cyber world, has the purpose of impacting an aspect of the environment in which its real counterpart exists, incorporates models and subject-matter expertise, and relies on both models and data (historical and real-time) to achieve its purpose.

Massive amounts data collected from robots, whether in the real world or on the plant floor, can be stored instead of overwritten every sample time. This data can be used to build models of expected behavior. Real-time streaming can be compared to past data. In a manufacturing plant, comparisons can be made to the same process from the same machine, the same process from a different machine in the same plant, or even a similar process at a different plant. These multiple dimensions of potential comparisons enable drifts and trends to be detected and potentially corrected.

Simulation models can be run in real time and compared to real-time data streaming from the physical system [116]. These simulation models can be built from first principles, using physics and chemistry, or they can be constructed using historical data. The output of a simulation model can be compared to real-time streaming data and be used to detect anomalies or faults that may occur. Real-time data can also be used to update the simulation models as drifts or small changes occur in the operating system.

Future challenges in the use of these data-driven and digital twin approaches include:

i. Automation of the development and tuning of models to reduce the amount of time required by subject-matter experts to set up these systems

ii. Development of methods to separate natural and expected drifts in system processes (which should result in model updates) from anomalous or faulty behaviors (which should trigger an alert to the system operator)

iii. Improved methods for including confidence intervals in the predictions of simulation models, in order to incorporate the natural uncertainty inherent in physical systems

iv. Numerically tractable and accurate (i.e., with respect to preserved quantities and context) reduced-order modeling and methods of composing the multi-physics and multiscale component models prevalent in digital twins

3.D.3 Challenges for Control

The high-level control challenges for robots in the real world are:

- Handling real-world data with unknown disturbances (perception and multimodal sensing)
- Provable correctness under uncertainty
- Trust and safety

Resilience, Reactivity, and Reconfigurability in Real-World Operations

As robots operate in novel environments, increased levels of resilience, reactivity, and reconfigurability are required at the control and planning levels and in the fusion thereof. Beyond the traditional robust control approaches, there needs to be approaches to model potential disturbances including discrete events as well as drastic environmental changes (e.g., sunshine to snowstorm or asphalt to mud). The uncertainty associated with these unknown environments and the propagation of that uncertainty through the robot control must be able to be quantified and manipulated so that appropriate control and planning decisions can be made.

Construction sites are semi-structured environments, and “smart construction” aims to automate them to the highest possible degree. This concept originates in flexible automation, in which a large variety of different goods are produced at a single manufacturing site through the on-demand reconfiguration of machines and robots. This idea is carried into the construction vertical where products typically have lot size of one (one-off production). While construction sites still rely on manual labor, the vision for the future is to leverage automation to speed up the production cycle and enhance quality while also reducing human risks, carbon emissions, and costs. Control challenges involve the integration of high-level planning and flexible, low-level control toward distributed solutions applicable to the domain in hand. Solutions will also need to address the heterogeneity of different construction operators in the overall system architecture.

Outer space, subterrain, and underwater domains are even less structured and uncertain. These environments are often remote and unfriendly for humans, so the use of robotic and autonomous systems is of great interest. Robots must be designed to survive and perform with limited power, communication, and navigation capabilities. In the underwater and subterrain domains, there may be a complete lack of

Figure 3.11: Heterogeneous multirobot coordination with humans-in-the-loop in extreme environments. Figure credited to NTNU AMOS/Stenberg.

infrastructure and services, and high levels of autonomy may be the only option. Research into perception, situational awareness, AI, and fault-tolerant control is therefore essential to achieve the resilient autonomy that is required by these environments, where the intervention of a human operator or assistant is impossible [117, 118].

Furthermore, all these domains call for hardened robotic systems with robust hardware and software, whether it is for planetary exploration, ocean science, subsea mining, or in-orbit satellite maintenance. Robotics makes these environments more available to us at a lower cost, and contribute to democratization of space and the oceans. In addition to research on the robotic systems themselves, research is also needed on their payloads, instrumentation, sampling systems, mechatronics, and interaction with robotic platforms (see Figure 3.11).

Provable Correctness Under Uncertainty

Given the large uncertainties associated with these diverse environments, as well as the critical nature of these operations, robots must be certified to be behave correctly. Traditional robust control methods can handle structured uncertainties but can result in overly conservative control approaches. Robotic systems must be designed and built to react and adapt appropriately under changing conditions. The uncertainties associated with the robot itself, its capabilities (including sensing and mobility), and constraints (including computation and power) must all be considered.

The limitations of the environment, including sensing, computation, and power constraints, must be taken into account. The time needed to compute a solution to the overall control problem is important to consider, as robots in the real world need to operate in real time. Power- and time-efficient algorithms and computation methods need to be developed.

Recent advances in sensing, computation, and data management have enabled the development of smart systems, providing greater autonomy and flexibility. These emerging control systems motivate sophisticated control objectives which go beyond the standard goals) pursued in classical control theory (i.e., stability, invariance, and optimality). For instance, one sophisticated objective is to control connected autonomous vehicles merging at a traffic intersection while ensuring safety and fuel economy constraints. Safety is the primary objective here, and at the same time, there is a need for flexible cooperation between these systems and the users. The complexity of such control objectives call for automated and provably correct techniques to verify or synthesize controllers for the emerging applications of control systems.

A promising methodology for addressing the above issues may be achieved through a careful integration of concepts from control theory and those of computer science, e.g. formal methods and assume-guarantee rules [119, 120]. Formal method approaches allow defining the tasks and objectives in an effective, real-time, and user-friendly way for dynamical systems under multiple complex constraints. For the past decade, formal languages such as linear temporal logic (LTL) [121] have been shown to be expressive enough to capture many important properties in various robotic scenarios. More recently, signal temporal logic (STL) [122], which induces quantitative features, has also been considered. Control barrier functions have been introduced as an extension of barrier certificates in order to provide probable safety [123].

Optimization can also be a powerful approach to achieve specified objectives and constraints in a systematic way. While progress has been made, current methods are restricted to tasks in fairly structured environments. More efforts are needed to develop control algorithms that can handle realistic problems in unstructured and dynamic environments. Further challenges to be addressed are scalability (increasing in numbers, heterogeneous, dynamic objects, and agents) and uncertainty (complex or unpredictable dynamics), which may require replanning and adaptation by robot systems in real time.

Trust and Safety With Humans Involved

The social applicability of robot operations—especially when it comes to daily life implementation—brings several new human considerations into the picture. These include hard constraints such as safety and security, as well as soft constraints such as comfort, trust, and acceptability.

Advances in sensing, computing, and actuation have enabled robots to operate more safely around humans, with appropriate control methods. Robots must be able to perceive their environments, including humans, and make reasonable predictions so that they can plan their own operations and paths without causing harm. Although much progress has been made in these areas, humans can be unpredictable and environments are dynamic. It remains a challenge to develop control strategies that ensure safe operations for human-robot interaction [124].

Humans and robots are expected to cooperate in many domains and modalities in the future. Some robots may be taught by humans how to perform specific tasks, such as assembly or cooking. Other robots may teach or help humans in such contexts as physical recovery from a stroke or even in an educational setting. Humans and robots may be teammates, working together to accomplish a common goal of searching a disaster area for victims or caring for patients in a hospital.

As robot hardware increases in functionality and perception and computing continues to improve, robots will be able to work together with humans and have the potential to help make our lives safer and more enriching.

Following the maturity of perception, localization, planning, and control functions as building blocks of autonomous driving, the focus of autonomous driving research in the 2020s has shifted toward system reliability and safety [125]. Situational understanding involving prediction of human behavior and risk-aware planning to incorporate traffic behavior uncertainties are the two main aspects that enable safe and efficient human-machine interaction [126]. The multimodal nature of uncertainty in traffic has been an extensive area of research in risk-aware planning [127]. Approaches such as partially observable Markov decision processes (POMDPs), probabilistic reachable set computation, and robust invariant set computations are employed to address the safety in the presence of uncertainty problem [128]. These research directions ultimately contribute to a highly autonomous transportation system that promotes increased safety, improves urban access for people with disabilities, and reduces carbon emissions and the need for personal vehicle ownership through planned shared usage. At the same time, the need to expand these approaches toward larger-scale and distributed implementations remains an open challenge.

The design and deployment of multirobot systems (MRS) is a fundamental challenge, as individual robots typically have access to limited resources. In centralized approaches, the robot team is typically considered to be a composite robot system to which a single robot motion planning algorithm can be applied. Due to the high dimensionality of the multirobot state-space, centralized approaches tend to be computationally impractical when the number of robots increases. While significant efforts have been devoted in recent decades to the distributed control of multirobot systems [129, 130] there are a number of directions that need improvement. Distributed approaches, in which all the robots are homogeneous with respect to control and are completely autonomous in the decision-making process, are more promising for large-scale multirobot systems. Multirobot systems that work in dynamically changing environments are safety- and mission-critical. Failures are expensive, as repair of both hardware and actions in place is often difficult or even impossible.

The importance of safety is even more evident when explicit or implicit interaction with humans takes place. To enable multirobot systems to safely handle more complex tasks, an effective framework is needed to induce what the team should be accomplishing as a global specification, and to produce low-level distributed control algorithms. Thus, heterogeneous multirobot system control in the presence of complex specifications will be a key research area in the future [131].

To realize the full potential of a MRS, a fundamental understanding of the interplay between sensing, communication, and execution in these systems is needed. In complex and dynamic environments, it may also be that some tasks are conflicting or even unachievable. In such cases, it is expected that the overall MRS will be able to reconfigure its plans according to that real-time environmental information in a way that the objectives for the whole group are maximally satisfied. Moreover, when the tasks are still too complex for the robots, coordination between humans and robots needs to be further explored.

Manufacturing

Automation has been used in manufacturing for more than a century, with mass production, assembly lines and conveyors. Robots were introduced a half-century ago, greatly expanding the capabilities of machines. Automated manufacturing, including robotics, led to high volumes and low costs for high-quality production. In particular, automation can lead to highly repeatable operations with low variances. These highly automated systems often incorporated cages around robots and safety sensors to protect humans who worked in the factories from the powerful, high-speed motions of the machines. Although many of the most repetitive and dangerous operations were automated, human workers were still required for setup, maintenance, troubleshooting, and dexterous assembly. The vision of a fully autonomous “lights-out” factory remains futuristic.

Recent trends in manufacturing are leading away from mass production and toward mass customization, in which many different models are produced in a single factory—or even personalized production, where each product coming off of the assembly line has been made for an individual consumer ("lot size of one"). These trends require advances in both the physical machines that can make different types of products as well as control and data methods that can coordinate production efficiently and maintain the low-cost advantages found in mass production [132].

There are opportunities to improve the efficiency, quality, and cost of production at every level in a traditional manufacturing control hierarchy [133]. The lowest level of the hierarchy includes controllers for individual machines, robots, conveyors, etc. in the plant. The next level up includes the coordination of all of these elements through a PLC (programmable logic controller). Although changing the G-code of a computer numerical control (CNC) machine or uploading a new robot program to a manipulator arm is straightforward, new control approaches are needed to enable flexibility at the next level up. It is currently difficult to make any changes in the control programs that synchronize the system via a PLC program and ensure that these changes result in a satisfactory solution. Currently, most high-volume production lines that produce multiple products are designed from the outset to produce a set of products, and reconfiguration to add a new product remains a challenge.

At the higher levels in the hierarchy, there are opportunities for control to address the monitoring and supervision of the production system. Predictive maintenance has the potential to increase the uptime of a factory and significantly reduce the costs of unscheduled downtime. The large amounts of data that are required to realize these solutions are becoming available, and methods that can leverage real-time and historical data from the same or multiple machines while including contextual factors and both human and artificial intelligence are beginning to show promise.

The move toward mass customization and personalized production will require rapid adaptation to shifting demands, as well as better coordination with customers, distributors, and suppliers. At all levels of the supply chain, there are major opportunities for estimation and control strategies, including predictive models, to enable more responsive, resilient, and reliable manufacturing systems. These will enable customers to get the products they need, when they need them, at a reasonable cost.

Recommendations

For young researchers: Among the biggest challenges in real-world robotics are making robots effective, resilient, and safe. This is essential for us to accept robots among us and to realize the full potential of robotics in society. We recommend focusing on the higher levels of the control hierarchy, including perception, situational awareness, autonomy, AI, and the use of digital twins. We will also need interdisciplinary collaboration between scientists in robotics, computer science, other engineering disciplines, and social sciences such as ethics and safety.

For funding agencies: To realize the promise of robots operating in the real world's unstructured, dynamic, and uncertain environments, fundamental research is needed to advance the frontiers of control. While traditional robust control guarantees performance within structured uncertainties, new frameworks of resilient and reactive control must be developed for unstructured and dynamic uncertainties. We need to create methods for real-time characterization and perception of the environment, along with control frameworks that can be proven to operate correctly under large, dynamic unknowns. Another critical thrust for research is the collaboration of humans and robots in these dynamic environments while ensuring safety and engendering the trust of human operators. Positive results will lead to improved productivity and efficiency. Short-term research can focus on semi-structured environments such as construction sites, manufacturing plants, and warehouses, while long-term efforts can address open environments alongside humans.

BIBLIOGRAPHY

[1] S. Russel and P. Norvig, *Artificial intelligence—a modern approach*, 4th ed., 2020.

[2] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the Dartmouth summer research project on Artificial Intelligence, August 31, 1955," *AI Magazine*, vol. 27, no. 4, pp. 12–14, Dec. 2006. https://ojs.aaai.org/index.php/aimagazine/article/view/1904

[3] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.

[4] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning.* Springer New York, Apr. 2017. https://www.ebook.de/de/product/8023140/trevor hastie robert tibshirani jerome friedman the elements of statistical learning.html

[5] R. Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences.* SAGE, 2014.

[6] J. Deng et al., "Imagenet: A large-scale hierarchical image database," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE Computer Society, 2009, pp. 248–255.

[7] A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," *IEEE Intelligent Systems*, vol. 24, 2009.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM,* vol. 60, no. 6, pp. 84–90, 2017.

[9] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," *CoRR*, vol. abs/1810.04805, 2018. http://arxiv.org/abs/1810.04805

[10] A. Bensoussan et al., "Machine learning and control theory," in *Handbook of Numerical Analysis*. Elsevier, 2022, pp. 531–558.

[11] R. J. Brachman and H. J. Levesque, *Knowledge Representation and Reasoning*. Elsevier, 2004. https://doi.org/10.1016/B978-1-55860-932-7.X5083-3

[12] I. Goodfellow, J. Bengio, and A. Courville, *Deep Learning*. MIT Press Ltd, Nov. 2016. https://www.ebook.de/de/product/26337726/ian goodfellow joshua bengio aaron courville deep learning.html

[13] N. P. Jouppi et al., "In-datacenter performance analysis of a tensor processing unit," in *Proceedings of the 44th Annual International Symposium on Computer Architecture*. ACM, Jun. 2017.

[14] T. B. Brown et al., "Language models are few-shot learners," in *Advances in Neural Information Processing Systems (NeurIPS)*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., 2020. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

[15] S. Greengard, "Neuromorphic chips take shape," *Communications of the ACM*, vol. 63, no. 8, pp. 9–11, Jul. 2020.

[16] A. W. Andreas Griewank, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. CAMBRIDGE, Nov. 2008. https://www.ebook.de/de/product/23743815/andreas griewank andrea walther evaluating derivatives principles and techniques of algorithmic differentiation.html

[17] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: a survey," *Journal of Machine Learning Research*, vol. 18, pp. 153:1–153:43, 2017. http://jmlr.org/papers/v18/17-468.html

[18] A. Paszke et al., "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems (NeurIPS)*, H. M. Wallach et al., Eds., 2019, pp. 8024–8035. https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

[19] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *USENIX Symposium on Operating Systems Design and Implementation,* K. Keeton and T. Roscoe, Eds. USENIX Association, 2016, pp. 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

[20] G. Brockman et al., "OpenAI gym," *CoRR*, vol. abs/1606.01540, 2016. http://arxiv.org/abs/1606.01540

[21] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2012, Vilamoura, Algarve, Portugal, October 7-12, 2012*. IEEE, 2012, pp. 5026–5033.

[22] M. Quigley et al., "ROS: an open-source robot operating system," in *ICRA Workshop on Open Source Software*, 2009.

[23] R. S. Sutton and A. G. Barto, *Reinforcement Learning*. MIT Press, Nov. 2018. https://www.ebook.de/de/product/32966850/richard s university of alberta sutton andrew g co director autonomous learning laboratory barto reinforcement learning. html

[24] M. J. Kochenderfer, T. A. Wheeler, and K. H. Wray, *Algorithms for Decision Making*. MIT Press, 2022.

[25] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[26] D. Silver et al., "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[27] O. Vinyals et al., "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, Oct. 2019.

[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[29] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," *CoRR*, vol. abs/1409.1259, 2014. http://arxiv.org/abs/1409.1259

[30] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, I. Guyon et al., Eds., 2017, pp. 5998–6008. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[31] I. J. Goodfellow et al., "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec*, Canada, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds., 2014, pp. 2672–2680.

[32] M. J. Kochenderfer, J. E. Holland, and J. P. Chryssanthacopoulos, "Next generation airborne collision avoidance system," *Lincoln Laboratory Journal*, vol. 19, no. 1, pp. 17–33, 2012. https://www.ll.mit.edu/sites/default/files/page/doc/2018-05/19_1_1_Kochenderfer.pdf

[33] V. Arya et al., "AI explainability 360: An extensible toolkit for understanding data and machine learning models," *Journal of Machine Learning Research*, vol. 21, pp. 130:1–130:6, 2020. http://jmlr.org/papers/v21/19-1035.html

[34] A. Corso, R. J. Moss, M. Koren, R. Lee, and M. J. Kochenderfer, "A survey of algorithms for black-box safety validation of cyber-physical systems," *Journal of Artificial Intelligence Research*, vol. 72, no. 2005.02979, pp. 377–428, 2021.

[35] C. Liu et al., "Algorithms for verifying deep neural networks," *Foundations and Trends in Optimization*, vol. 4, no. 3–4, pp. 244–404, 2021. https://arxiv.org/abs/1903.06758

[36] R. Luo, F. Tian, T. Qin, E. Chen, and T.-Y. Liu, "Neural architecture optimization," A*dvances in neural information processing systems*, vol. 31, 2018.

[37] N. Matni and V. Chandrasekaran, "Regularization for design," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3991–4006, 2016.

[38] K. Manohar, J. N. Kutz, and S. L. Brunton, "Optimal sensor and actuator selection using balanced model reduction," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 2108–2115, 2021.

[39] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies." *Journal of Machine Learning Research*, vol. 9, no. 2, 2008.

[40] R. Drummond, M. C. Turner, and S. R. Duncan, "Reduced-order neural network synthesis with robustness guarantees," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[41] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," *Annual Review of Fluid Mechanics*, vol. 52, pp. 477–508, 2020.

[42] K. Kashinath *et al.*, "Physics-informed machine learning: case studies for weather and climate modelling," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20200093, 2021.

[43] Z. Li *et al.*, "Fourier neural operator for parametric partial differential equations," *arXiv preprint arXiv:2010.08895*, 2020.

[44] J. Konecny, B. McMahan, and D. Ramage, "Federated optimization: Distributed optimization beyond the datacenter," *arXiv preprint arXiv:1511.03575*, 2015.

[45] K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning," in *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191.

[46] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communicationefficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics. PMLR*, 2017, pp. 1273–1282.

[47] K. Bonawitz et al., "Towards federated learning at scale: System design," *Proceedings of Machine Learning and Systems*, vol. 1, pp. 374–388, 2019.

[48] T. Yang et al., "A survey of distributed optimization," *Annual Reviews in Control*, vol. 47, pp. 278–305, 2019.

[49] A. Nedic and J. Liu, "Distributed optimization for control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 77–103, 2018.

[50] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, "System level synthesis," *Annual Reviews in Control*, vol. 47, pp. 364–393, 2019.

[51] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," I*EEE transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.

[52] R. Baheti and H. Gill, "Cyber-physical systems," *The impact of control technology*, vol. 12, no. 1, pp. 161–166, 2011.

[53] Y. Ovadia et al., "Can you trust your model's uncertainty? evaluating predictive uncertainty under dataset shift," *Advances in neural information processing systems*, vol. 32, 2019.

[54] P. W. Koh et al., "Wilds: A benchmark of in-the-wild distribution shifts," in *International Conference on Machine Learning*. PMLR, 2021, pp. 5637–5664.

[55] F. Codevilla, M. Müller, A. López, V. Koltun, and A. Dosovitskiy, "End-to-end driving via conditional imitation learning," in *2018 IEEE international conference on robotics and automation (ICRA*). IEEE, 2018, pp. 4693–4700.

[56] Z. Li et al., "Reinforcement learning for robust parameterized locomotion control of bipedal robots," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 2811–2817.

[57] Q. Do and M. E. Hasselmo, "Neural circuits and symbolic processing," *Neurobiology of Learning and Memory*, vol. 186, p. 107552, 2021. https://doi.org/10.1016/j.nlm.2021.107552

[58] S. Chaudhuri *et al.*, 2021.

[59] T. Gafni, N. Shlezinger, K. Cohen, Y. C. Eldar, and H. V. Poor, "Federated Learning: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 39, no. 3, pp. 14–41, May 2022.

[60] G. Notarstefano, I. Notarnicola, and A. Camisa, "Distributed Optimization for Smart Cyber-Physical Networks," *Foundations and Trends® in Systems and Control*, vol. 7, no. 3, pp. 253–383, 2019.

[61] Y. Qu et al., “Decentralized Federated Learning for UAV Networks: Architecture, Challenges, and Opportunities,” *IEEE Network*, vol. 35, no. 6, pp. 156–162, Nov. 2021.

[62] A. Clayton, *How we reach zero-carbon energy independence*. World Economic Forum Annual Meeting, May 2022.

[63] “Paris agreement under the united nations framework convention on climate change,” 2015.

[64] O. Roelofsen, K. Somers, E. Speelman, and M. Witteveen, *Plugging in: What electrification can do for industry*. McKinsey, 2020.

[65] Statista, “Share of electricity generation from wind energy sources worldwide from 2010 to 2021,” 2022.

[66] Statista, “Share of electricity generation from solar energy worldwide from 2010 to 2021,” 2022.

[67] United Nations Development Programme (UNDP), “World energy assessment.”

[68] J. Kim et al., “Decarbonizing the iron and steel industry: A systematic review of sociotechnical systems, technological innovations, and policy options,” *Energy Research & Social Scienc*e, vol 89, p. 102565, 2022.

[69] F. Chiara and M. Canova, “A review of energy consumption, management, and recovery in automotive systems, with considerations of future trends,” *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 227, no. 6, pp. 914–936, 2013.

[70] B. Sarlioglu and C. T. Morris, “More electric aircraft: Review, challenges, and opportunities for commercial transport aircraft,” *IEEE transactions on Transportation Electrification*, vol. 1, no. 1, pp. 54–64, 2015.

[71] A. G. Boulanger, A. C. Chu, S. Maxx, and D. L. Waltz, “Vehicle electrification: Status and issues,” *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1116–1138, 2011.

[72] N. K. Golla and S. K. Sudabattula, *Impact of plug-in electric vehicles on grid integration with distributed energy resources: A comprehensive review on methodology of power interaction and scheduling*. Materials Today: Proceedings, 2021.

[73] M. L. Tuballa and M. L. Abundo, “A review of the development of Smart Grid technologies,” *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 710–725, 2016.

[74] A. G. Alleyne and C. T. Aksland, “Control as an enabler for electrified mobility,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 659–688, 2022.

[75] L. Pérez-Lombard, J. Ortiz, and C. Pout, “A review on buildings energy consumption information,” *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.

[76] J. J. Muldoon et al., “Model-guided design of mammalian genetic programs,” *Science Advances*, vol. 7, no. 8, p. eabe9375, 02 2021. https://advances.sciencemag.org/ lookup/doi/10.1126/sciadv.abe9375

[77] H. Niederholtmeyer et al., “Rapid cell-free forward engineering of novel genetic ring oscillators,” *eLife*, vol. 4, p. e09771, 2015.

[78] EBRC, Research Road Maps For the Next-Generation Bioeconomy, "Engineering Biology: A Research Road Map for the Next-Generation Bioeconomy," https://roadmap.ebrc.org/2019-roadmap.

[79] J. A. Doudna and E. Charpentier, "The new frontier of genome engineering with CRISPR-Cas9," *Science*, vol. 346, no. 6213, p. 1258096, 2014.

[80] A. Burt and A. Crisanti, "Gene drive: evolved and synthetic," *ACS chemical biology*, vol. 13, no. 2, pp. 343–346, 2018.

[81] J. D. Boeke et al., "The genome project-write we need technology and an ethical framework for genome-scale engineering," *Science*, vol. 353, no. 6295, pp. 126–127, 2016.

[82] R. Mackelprang et al., "Guiding ethical principles in engineering biology research," *ACS Synthetic Biology*, vol. 10, no. 5, pp. 907–910, 2021.

[83] E. National Academies of Sciences and Medicine, *Preparing for future products of biotechnology*. National Academies Press, 2017.

[84] D. Del Vecchio, A. J. Dy, and Y. Qian, "Control theory meets synthetic biology," *Journal of the Royal Society Interface*, 2016.

[85] V. Hsiao, A. Swaminathan, and R. M. Murray, "Control theory for synthetic biology: Recent advances in system characterization, control design, and controller implementation for synthetic biology," *IEEE Control Systems Magazine*, vol. 38, no. 3, pp. 32–62, 2018.

[86] M. Khammash, M. Di Bernardo, and D. Di Bernardo, "Cybergenetics: Theory and methods for genetic control system," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 916–926.

[87] D. Mishra, P. M. Rivera, A. Lin, D. D. Vecchio, and R. Weiss, "A load driver device for engineering modularity in biological networks," *Nature Biotechnology*, vol. 32, no. 12, pp. 1268 – 1275, 11 2014. http://www.nature.com/doifinder/10.1038/nbt.3044

[88] C. Briat, A. Gupta, and M. Khammash, "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks," *Cell Systems*, vol. 2, no. 1, pp. 15–26, 2016. https://www.sciencedirect.com/science/article/pii/S2405471216000053

[89] J. E. Pinsker et al., "Evaluation of an Artificial Pancreas with Enhanced Model Predictive Control and a Glucose Prediction Trust Index with Unannounced Exercise," *Diabetes Technology & Therapeutics*, vol. 20, no. 7, pp. 455–464, 2018.

[90] F. Angaroni et al., "An Optimal Control Framework for the Automated Design of Personalized Cancer Treatments," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 523, 2020.

[91] J. S. Yu and N. Bagheri, "Modular microenvironment components reproduce vascular dynamics de novo in a multi-scale agent-based model," *Cell Systems*, vol. 12, no. 8, pp. 795–809.e9, 2021.

[92] J. J. Muldoon, Y. Chuang, N. Bagheri, and J. N. Leonard, "Macrophages employ quorum licensing to regulate collective activation," *Nature Communications*, vol. 11, no. 1, p. 878, 2020.

[93] J. Stewart-Ornstein and H. El-Samad, “Stochastic modeling of cellular networks,” *Methods in cell biology*, vol. 110, no. Nat. Rev. Microbiol.42006, pp. 111–37, 2012.

[94] M. Voliotis, P. Thomas, R. Grima, and C. G. Bowsher, “Stochastic Simulation of Biomolecular Networks in Dynamic Environments,” *arXiv*, vol. 12, no. 6, p. e1004923, 2015.

[95] A. Gupta, C. Schwab, and M. Khammash, “DeepCME: A deep learning framework for computing solution statistics of the chemical master equation,” *PLoS Computational Biology*, vol. 17, no. 12, p. e1009623, 2021.

[96] A. Ghaffarizadeh, R. Heiland, S. H. Friedman, S. M. Mumenthaler, and P. Macklin, “PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems,” *PLoS Computational Biology*, vol. 14, no. 2, p. e1005991, 2018.

[97] J. S. Yu and N. Bagheri, “Agent-Based Models Predict Emergent Behavior of Heterogeneous Cell Populations in Dynamic Microenvironments,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 249, 2020.

[98] A. J. N. Anelone, M. F. Villa-Tamayo, and P. S. Rivadeneira, “Oncolytic virus therapy benefits from control theory,” *Royal Society Open Science*, vol. 7, no. 7, p. 200473, 2020.

[99] M. F. Villa-Tamayo, A. J. N. Anelone, and P. S. Rivadeneira, “Tumor Reduction Using Oncolytic Viruses Under an Impulsive Nonlinear Estimation and Predictive Control Scheme,” *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1705–1710, 2021.

[100] K. Krawczyk *et al.*, “Electrogenetic cellular insulin release for real-time glycemic control in type 1 diabetic mice,” *Science*, vol. 368, no. 6494, pp. 993–1001, 2020.

[101] K. Deisseroth *et al.*, “Next-generation optical technologies for illuminating genetically targeted brain circuits,” *Journal of Neuroscience*, vol. 26, no. 41, pp. 10380–10386, 2006.

[102] R. Sepulchre, T. O’Leary, G. Drion, and A. Franci, “Control by neuromodulation: A tutorial,” in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 483–497.

[103] J. Bohg et al., “Interactive perception: Leveraging action in perception and perception in action,” *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1273–1291, 2017.

[104] F. Fuchs, Y. Song, E. Kaufmann, D. Scaramuzza, and P. Dürr, “Super-human performance in gran turismo sport using deep reinforcement learning,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4257–4264, 2021.

[105] C. Cadena et al., “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age,” *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.

[106] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart, “Robust visual inertial odometry using a direct ekf-based approach,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 298–304.

[107] C. Laschi, B. Mazzolai, and M. Cianchetti, “Soft robotics: Technologies and systems pushing the boundaries of robot abilities,” *Science Robotics*, vol. 1, no. 1, p. eaah3690, 2016.

[108] E. W. Hawkes, C. Majidi, and M. T. Tolley, “Hard questions for soft robotics,” *Science robotics*, vol. 6, no. 53, p. eabg6049, 2021.

[109] A. Billard and D. Kragic, “Trends and challenges in robot manipulation,” *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[110] A. Ajoudani et al., “Progress and prospects of the human–robot collaboration,” *Autonomous Robots*, vol. 42, no. 5, pp. 957–975, 2018.

[111] J. Kuffner and J. Xiao, “Motion for manipulation tasks,” in *Springer Handbook of Robotics. Springer*, 2016, pp. 897–930.

[112] K. Gryte, T. H. Bryne, and T. A. Johansen, “Unmanned aircraft flight control aided by phased-array radio navigation,” *Journal of Field Robotics*, vol. 38, pp. 532–551, 2021.

[113] A. Ollero, M. Tognon, A. Suarez, D. Lee, and A. Franchi, “Past, present, and future of aerial robotic manipulators,” *IEEE Transactions on Robotics*, vol. 38, no. 1, pp. 626–645, 2022.

[114] K. Alexis, G. Darivianakis, M. Burri, and R. Siegwart, “Aerial robotic contact-based inspection: planning and control,” *Autonomous Robots*, vol. 40, p. 631–655, 2016.

[115] J. Moyne et al., “A requirements driven digital twin framework: Specification and opportunities,” *IEEE Access*, vol. 8, pp. 107781–107801, 2020.

[116] M. Saez, F. P. Maturana, K. Barton, and D. M. Tilbury, “Real-time manufacturing machine and system performance monitoring using internet of things,” *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 4, pp. 1735–1748, 2018.

[117] M. Tranzatto et al., “CERBERUS: Autonomous legged and aerial robotic exploration in the tunnel and urban circuits of the darpa subterranean challenge,” *Field Robotics*, 2022.

[118] H. F. Grip et al., “Modeling and Identification of Hover Flight Dynamics for NASA’s Mars Helicopter,” J. Guidance, *Control and Dynamics*, vol. 43, no. 2, 2020.

[119] C. Belta, B. Yordanov, and E. A. Gol, Formal methods for discrete-time dynamical systems. *Springer*, 2017, vol. 15.

[120] P. Tabuada, Verification and control of hybrid systems: a symbolic approach. *Springer Science & Business Media*, 2009.

[121] M. Y. Vardi, “An automata-theoretic approach to linear temporal logic,” in *Logics for concurrency*. Springer, 1996, pp. 238–266.

[122] O. Maler and D. Nickovic, “Monitoring temporal properties of continuous signals,” in *Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems*. Springer, 2004, pp. 152–166.

[123] L. Lindemann and D. V. Dimarogonas, “Control barrier functions for signal temporal logic tasks,” *IEEE control systems letters*, vol. 3, no. 1, pp. 96–101, 2019.

[124] P. A. Lasota, T. Fong, J. A. Shah *et al., A survey of methods for safe human-robot interaction*. Now Publishers Delft, The Netherlands, 2017, vol. 104.

[125] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, “A survey of autonomous driving: Common practices and emerging technologies,” *IEEE access*, vol. 8, pp. 58443– 58469, 2020.

[126] A. Rudenko et al., “Human motion trajectory prediction: A survey,” *The International Journal of Robotics Research*, vol. 39, no. 8, pp. 895–935, 2020.

[127] W. Schwarting, J. Alonso-Mora, and D. Rus, “Planning and decision-making for autonomous vehicles,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 187–210, 2018.

[128] K. Muhammad, A. Ullah, J. Lloret, J. Del Ser, and V. H. C. de Albuquerque, “Deep learning for safe autonomous driving: Current challenges and future directions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4316–4336, 2020.

[129] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[130] M. M. Zavlanos and G. J. Pappas, “Distributed connectivity control of mobile networks,” *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1416–1428, 2008.

[131] T. Gustavi, D. V. Dimarogonas, M. Egerstedt, and X. Hu, “Sufficient conditions for connectivity maintenance and rendezvous in leader–follower networks,” *Automatica*, vol. 46, no. 1, pp. 133–139, 2010.

[132] D. M. Tilbury, “Cyber-physical manufacturing systems,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, pp. 427–443, 2019.

[133] L. Apilioğulları, “Digital transformation in project-based manufacturing: Developing the isa-95 model for vertical integration,” *International Journal of Production Economics*, p. 108413, 2022.

CHAPTER 4

Emerging Methodologies

Emerging Methodologies

As we proceed further into the 21st century, societal needs and goals continue to evolve. Our underlying systems become ever more complex as well, necessitating new tools for analysis and synthesis that effectively utilize advancing technologies and address these high-level objectives. This chapter showcases key methodologies that are emerging in control systems. We have focused on five central concepts: i) learning, ii) safety, iii) resilience, iv) humans, and v) architectures.

The first section focuses on the intersection between learning and data-driven control, with a specific emphasis on scalability and dynamics. The second section emphasizes the need for weaving in safety and its formal representations into control of autonomous systems. The third section concerns the resilience of CPS in the wake of cyberattacks, extreme weather, and other major disruptions. The fourth section puts humans at the center of the conversation and addresses technical challenges in the analysis and synthesis of interactions between control systems and humans in the emerging field of CPHS.

Finally, the fifth section focuses on control architectures that seek to address how various systems and their components interconnect and interact to produce desirable societal-scale control systems.

4.A Learning and Data-Driven Control

Anders Rantzer, Florian Dörfler, Na Li, Nikolai Matni, Lucy Y. Pao, Gireeja Ranade, Angela Schoellig, Claire J. Tomlin, Melanie Zeilinger

Fruitful interaction between machine learning and data-driven control plays out on a range of scales—from quantum control and molecular dynamics to swarms of autonomous robots and global energy markets. To fully exploit the synergies between these two disciplines, we need to revisit the foundations of systems and control with a focus on the intersections between dynamics, scalability, learning, and architecture.

Abstract One of the major developments in control over the past decade—and one of the most important moving forward—is the interaction of machine learning and control systems. Learning-based components are used to process sensor data, represent uncertain behavior of systems in their environments, and enable decision-making based on context and observed data. These methods have a long history in control, but technological and methodological advances in machine learning are engendering new opportunities and challenges. To take full advantage of these, we will need to revisit the theoretical foundations of learning and data-driven control. Moreover, successful applications will need to be analyzed for the efficient transfer of knowledge between the various branches of engineering and to the next generation of scholars. Real-world applications are highly diverse, ranging across scales from synthetic biology at the cellular level to the design of economic incentives in large energy networks. Therefore, this knowledge transfer will be interdisciplinary but activated by a common core of system theory. Several major research challenges should be addressed. The interplay between offline and online computation must be better understood. Theory and tools for physics-informed learning must be developed in a dynamic setting. We will need to verify learning-based solutions for safety-critical applications, build a better theoretical foundation for the exploration/exploitation tradeoff, and further examine the interaction of mechanism design and information design in economic contexts. Last but not least, we will need to develop a systematic theory for the role of system architecture in learning for control.

4.A.1 Introduction

One of the major developments is a growing intersection of machine learning and control systems. This is evidenced by the facts that learning-based components are used to process sensor data, represent uncertain behavior of systems in their environments, and enable decision-making based on context and observed data. These methods have a long history in control, but renewed interest has been sparked by a number of technological and methodological advances, including:

- Inexpensive and efficient sensors that simplify the collection of data
- Increasingly powerful processors supporting both online and offline computations
- High-bandwidth and low-latency communications that are approaching ubiquity
- New machine learning algorithms enabling the efficient processing of huge datasets
- Open-source software and data generating higher-quality and more efficient solutions

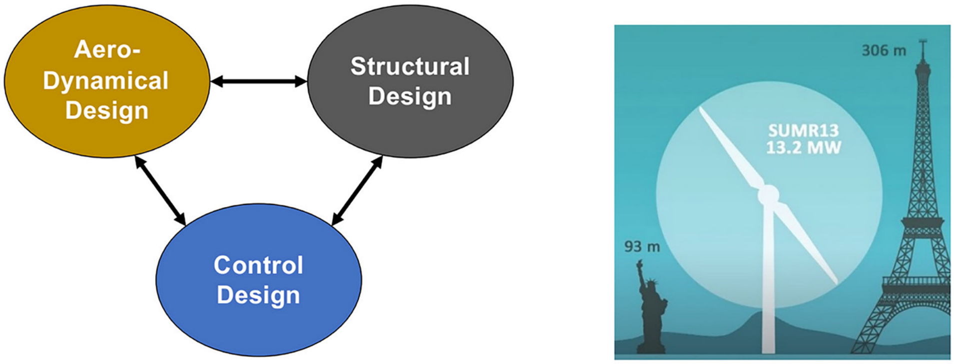


Figure 4.1: Pushing the limits of systems performance in complex systems (e.g., large wind turbines) often involves repeated controller tuning for different mechanical and aerodynamical properties in an iterative design process. This is one important application where initial automatic controller tuning techniques are already beginning to make an impact, and where future developments in automatic controller tuning will make a big difference.

These developments are creating new opportunities in classical control contexts. They are also giving rise to entirely new challenges as data-driven feedback control is being deployed in a wide range of new settings. Many aspects of modeling and learning are irrelevant in static settings but are of crucial importance in the realm of feedback control. These elements are well-known in control engineering, but the limitations of current learning theory become apparent when we try to exploit recent machine learning algorithms in a feedback setting. Below, we cover some important issues to be addressed as we proceed toward a more complete systems theory for data-driven control. We will begin with near-term challenges and gradually advance to ultimate goals.

4.A.2 Autotuning

Autotuners for PID controllers have found widespread commercial use in the process industry since the 1980s. The most common ones are based on a simple relay feedback experiment that provides enough input excitation to allow for the automatic estimation of essential system parameters. Based on these parameters, a controller can be tuned automatically. An important motivation for the use of autotuning in the process industry is the large number of control loops, which make manual tuning tedious and expensive.

It is useful to ask how these early ideas can be further developed. Certainly, there are many application areas where automatic tuning has the potential to improve efficiency and save tedious manual work. HVAC for buildings has already been identified as promising, given that it also involves many simple control loops of a similar type. More complex systems could also be of interest. Consider, for example, large mechanical structures like wind turbines, where structural and aerodynamical properties are intertwined with control systems. The design and optimization of modern wind turbines becomes an interactive process, illustrated in Figure 4.1, where the controller must be updated many times before a successful co-design can be found. Automatic tuning has the potential to speed up the design process and generate great savings. Another area is in automotive industry, where the demand for highly customized products creates a need for simple and rapid tuning of various control loops.

4.A.3 Batch Learning

Batch learning, as opposed to incremental learning, means that algorithms run offline and do not interfere with the system during data collection. This simplifying setting is exploited in classical system identification as well as iterative learning control [1]. For large data sets, such methods need to be combined with modern machine learning tools for efficient computations. However, to ensure robust optimization for complex tasks in the presence of uncertainties, joint design of experiments and algorithms remains essential. Thus, well-known difficulties associated with closed-loop identification, colored noise, time-varying system parameters, and unmodeled dynamics create new challenges in combination with learning algorithms developed under more idealized conditions.

4.A.4 Extremum Seeking and Learning

In the field of adaptive and learning systems, we can also see many opportunities with extremum seeking [2]. With the explosion of interest in learning algorithms (for all sorts of purposes, including optimization and control), it should not be overlooked that extremum seeking is also a learning-based approach (in the minimalistic, most efficient sense of learning), a model-free approach, and a data-based approach.

Extremum seeking's real-time performance, rigorous convergence guarantees, and convergence rate assignment capabilities represent what machine learning and reinforcement learning algorithms have yet to achieve. It is our hope that some of the researchers and students who are pursuing learning-based capabilities will take note of these opportunities, especially in terms of blending model-free learning and optimization with the model-based compensation of complex but physically grounded processes.

4.A.5 Off-Line vs. On-Line Computations

The trade-off between offline (design time) and online (run time) computations, which is central to reinforcement learning, has been extensively studied in the context of MPC. This is due to the transition from traditional applications in process control (with time constants of minutes or hours) to recent applications in automotive industry and power electronics that have much faster time constraints. The idea is simple and fundamental: online computations have obvious advantages, since planning and execution can be focused on the situation at hand and employ up-to-date sensor information from a variety of sources. This generally improves robustness to model errors and disturbances. However, for applications in which there is less time to gather and process information, we must increasingly rely on plans made offline. Another important reason for focusing on the distinctions between the two modes is the uncertainties that can occur in real time. These introduce a significant shift in the relative roles of the two computations, placing more emphasis on the latter rather than the former.

This raises interesting challenges for the development of new theory and methodology:

- Representations of cost functions, control policies, and dynamics that provide efficient interfaces between online and offline computations
- Tailored real-time numerical methods with inherent robustness against time-varying computational constraints
- Systematic collection of training data
- Quantification of how data properties impact the final control performance
- Analysis of the tradeoff between simulation based training and experimental data

4.A.6 Parametric vs. Non-Parametric Methods

The question of whether data-driven controllers should be based on a parameterized plant-model can be traced back to the early days of adaptive control. From an engineering perspective, this model can serve many purposes, especially if its parameters have clear interpretations. The plant-model could then play an important role in verification, stress testing, and fault diagnosis, and the acceleration of learning by incorporating a *priori* information. Furthermore, the model can enable the transfer of a learned policy from one system to another system with similar features. Most importantly, this model can then enable an analytical foundation for establishing stability and robustness guarantees. Creating models for such purposes—and, more generally, combining data-driven models with first-principles models for control—is a major challenge closely related to physics-informed machine learning.

4.A.7 Partially Observable and Controllable Systems

Most machine learning algorithms assume that essential variables can be measured. Similarly, most literature on system identification and adaptive control is focused on control variables that are accessible for manipulation. However, there are many applications where essential variables can neither be easily measured nor manipulated but where feedback control policies need to be designed anyway. Obvious examples are government decisions to fight a pandemic or to mitigate climate change, as well as the control of large-scale infrastructure systems or the deployment of cancer therapies inside a human body. Research efforts in this direction are currently being pursued in statistics and several application disciplines under the term of "causal inference." However, the development of a theory for dynamic data-driven control under such circumstances remains an overarching challenge.

4.A.8 Verification for Safety-Critical Applications

The verification of software for safety-critical applications was already a major engineering challenge and a multibillion-dollar industry before the introduction of machine learning components. The addition of such devices greatly complicates matters. Schematically speaking, the literature follows two parallel paths, one formal and one pragmatic. On the formal side, the main idea is that by severely restricting the set of allowable algorithms and simultaneously making strict assumptions about the system to be controlled, it is possible to prove safety and other properties (e.g., using so called control barrier functions). The challenge is to extend this paradigm to include components like neural networks. Most likely, this will only be possible if we are willing to restrict our attention to special classes, where verifiability is explicitly taken into account during the training process. Probabilistic, rather than deterministic, certificates for safety and performance could be the natural goal for analysis of safety-critical feedback loops.

4.A.9 Optimal Exploration-Exploitation

The role of excitation has long been recognized as an essential feature in adaptive control. The idea is that unless you explore all relevant input directions, you cannot keep your model up to date. Feedback based on incorrect models may deteriorate and cause several undesirable effects such as bursting [3], or even system instability. At the same time, excitation is costly and may itself have a negative impact on performance, so you don't want too much. This has motivated the concept of dual control. In machine learning, the corresponding tradeoff is known as exploration versus exploitation. The idea is easy to understand at a qualitative level, but a fully quantitative theory is far from being available.

To understand the role of exploration, a key consideration is how to quantify the current competence level of a controller based on the kind of training data that it has received. This includes revisiting experiment design techniques of system identification. Sample complexity may only be a coarse measure in this regard, since it is not just the number of samples but the "quality" of the samples that matters.

The real difficulties occur when exploration and exploitation are intertwined. Importantly, some states cannot be explored unless you first reach the relevant region of state-space. Hence, exploitation is very often a prerequisite for exploration. Exploitation, i.e., control, may often need to precede exploration, i.e., learning. The idea is simple to grasp qualitatively, but hard to quantify theoretically. One of the difficulties is that statistical analysis is complicated by highly nonlinear learning dynamics. Worst-case analysis inspired by robust control and adversarial learning is therefore a valuable complement.

4.A.10 Data-Driven Mechanism Design and Information Design

The control of large systems with many decision-makers has strong ties to economic theory. A good example is power networks, where price incentives are used as control variables. Other cases include recommendation algorithms and online advertising that interacts with feedback from humans. Hence, it natural to take advantage of relevant economic theory and tools for the purpose of control. Mechanism design is a branch of economic theory that explores how agents can achieve desirable social or economic outcomes given the constraints of incomplete information and individual self-interest. Similarly, information design studies how outcomes are influenced by information sharing. A concrete example is when traffic flow in a road network is influenced not only by tariffs but by designing the information shared with drivers. Theory and methodology for mechanism design and information design will be of foundational importance for future sustainable societies, with data-driven methods filling the gaps where first-principle models are missing.

4.A.11 Learning Architectures to Support High Level Objectives

Consider two future autonomous system use cases: i) a bomb-defusing rover sent into an unfamiliar GPS- and communication-denied environment (e.g., a cave or mine), which is tasked with the objective of locating and defusing an improvised explosive device, and ii) an autonomous UAV competing in a racing competition. Both systems will make decisions based on inputs from a combination of simple sensing devices (such as inertial measurement units) and more complex ones (such as cameras). However, the extreme diversity in requirements, conceptually illustrated in Figure 4.2, highlights the need for a principled approach to navigate architectural tradeoffs. Two philosophically opposing approaches to designing suitable perception/action pipelines are: i) layered, and ii) end-to-end. The more familiar layered (or modular) approach is rooted in engineering best practices, designs different layers (perception, prediction, localization, planning, control) independently, and interconnects them via prespecified (rigid) protocols. While the benefits of such approaches are well-known, this approach limits interaction between layers to often ad-hoc specifications shared between teams, thus limiting end-to-end optimization. Hence, all protocols come with a cost. In contrast, the end-to-end approach, classically favored by the deep reinforcement learning community, seeks to learn an optimized map directly from pixels (or other high-dimensional sensory inputs) to actions. Such approaches have the advantage of optimizing the full perception/action pipeline for a specific application and have led to impressive demonstrations—but the resulting systems are opaque, difficult to troubleshoot, and lack safety guarantees. Further, they often preclude prior knowledge from being incorporated into learning/design, leading to increased sample complexity and brittle algorithms.

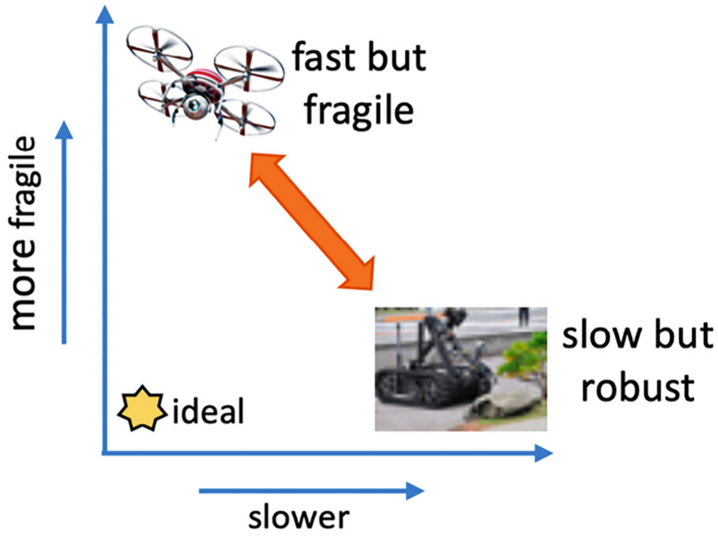


Figure 4.2: Learning architecture must reflect the fact that applications can have very different robustness and performance requirements depending on dynamics, environments, objectives, and constraints.

The two approaches above can be viewed as extreme points of a broader architectural design space, with the layered approach optimizing for interpretability, modularity, and arguably safety and robustness, and the end-to-end approach optimizing for performance and simplicity of implementation. A coherent theory of architecture should allow for a principled exploration of the entire design space. Many questions remain unanswered. For example: What algorithms can be connected? What are their input/output properties? How do errors propagate? It is a major challenge to define the architecture design space in such a way that tradeoffs of this kind can be analyzed in a structured and quantitative manner.

Recent years have seen promising first steps towards balancing the adaptability and performance of reinforcement learning with the safety and robustness of control. For example, Lyapunov and barrier constraints have found use as safety shields in reinforcement learning, while priors and regularizers rooted in model-based robust control have been used to advance model-free reinforcement learning. In this way, the machine learning and control communities are converging upon methods that explore intermediate architectures. These are important early efforts in addressing a bigger challenge: How to systematically build architectures that support safe and robust performance.

Recommendations

For young researchers: The interface between learning and data-driven control is one of the most exciting and fast-moving areas of engineering right now, with remarkable opportunities to address societal and environmental needs. We recommend mastering the basic theoretical foundations before progressing into newly emerging spaces. Theory for feedback control and system identification has been developed over several decades and its underpinnings in mathematics, statistics, and computer science share common ground with optimization theory and machine learning. We also recommend getting deeply involved in a specific application area. It is often in the interaction between disciplines that the most exciting ideas arise. For example, the natural processes of learning and control in biology will likely remain an important source of inspiration for engineers in the years to come.

For funding agencies: Recent decades have seen phenomenal growth in AI and machine learning, with a wide variety of applications in image analysis, language processing, medicine, and beyond. However, most success stories so far have related to static problems. There remain many societal and environmental challenges that will require learning on faster timescales, where dynamics and feedback become dominant. Examples include dynamic networks for energy, water, traffic, communications, healthcare, and sustainable food production. This is the home ground for systems and control theory. However, the integration of learning and control is still in its infancy, and many fundamental questions remain unresolved. With the proper resources, these practices can offer great rewards that can be harvested in a secure and reliable manner for the benefit of society

4.B Safety-Critical Systems

Aaron D. Ames, Koushil Sreenath, Claire J. Tomlin, Paulo Tabuada, Mykel J. Kochenderfer

> Ensuring the safe behavior of intelligent and autonomous systems is critical to establishing trust and enabling widespread deployment. The formal representations of safety given by control theory provide a paradigm for synthesizing, testing, and guaranteeing safe behavior that can be put into practice in real-world systems.

Abstract As intelligent systems are deployed in real-world settings, there is an increasing need to providing guarantees that these systems will not violate underlying safety constraints. For example, semiautonomous planes must fly within safety envelopes and robots must avoid collisions with environmental obstacles, other agents, and pedestrians. The failure to provide general, formal, and verifiable guarantees of safe behavior leads to ad-hoc solutions that are overly application-, domain-, and platform-specific. When edge cases are not considered, the results can be catastrophic: loss of public confidence and even the loss of human life. There is a pressing need to develop safety-critical control approaches that can undergird the design of any advanced control and autonomy feature. The real-world applications of such an approach range from autonomous cars and planes to advanced robotic and CPHS.

4.B.1 Introduction

We will focus on these three fundamental questions of safety:

- **Safety-Critical Control:** At the level of complex control systems—described by nonlinear models—what are the underlying means to realize safe behaviors?
- **Safety-Critical Autonomy:** Complex autonomous systems instantiate layered architectures that combine planning, trajectory generation, and control. How do we guarantee safety at and across these layers?
- **Safety-Critical Learning:** Given systems with unknown components, insufficient models, or perception modules, how do we combine learning with safety-critical approaches?

Addressing these questions can lead to formulations of safety in systems with black-box components (e.g., machine learning modules) and the application of safety-critical approaches to safety-critical systems with humans-in-the-loop. These concepts, taken in total, lead to open questions and problems in safety-critical systems. In this section, we present key open problems building upon a solid foundation of concepts that have been verified and realized experimentally on a wide variety of systems (e.g., multirobot systems) [4].

4.B.2 Notions of Safety and Their Relation

Our future will increasingly be filled with autonomous systems—from autonomous cars to medical devices that carry out intricate operations with minimal human intervention. To enable such systems to operate around and even inside humans, we need a strict notion of safety. However, there is no single standard definition of safety; we thus explore a few key definitions of the concept and investigate the relation between them. Notions of safety will form the basis for understanding architectures for safety and, consequently, for the synthesis of safety-critical methods in adaptation and learning.

Deterministic

The notion of safety was first introduced in the context of program correctness [5] where safety requires that "bad" things do not happen. For dynamical systems, this corresponds to ensuring forward invariance of "good" regions in space [6]. That is, early definitions of safety [7] involved defining a safety region that was given by an inequality that carved out a region in state-space, like $L = \{x \mid l(x) \geq 0\}$. Thus, safety is quantified via the zero super-level set of $l(x)$. While it is easy to define such a safe region starting from safety constraints, it is far less obvious how to ensure or verify that a system with input constraints remains within the safe region. Put another way, it is very hard to find bounded control inputs to guarantee that a system stays within a given set and is thereby safe.

One way to address this problem is to compute the set of all initial states from which the system can evolve such that it never leaves L (visualized in Figure 4.3). This is done by ensuring that there always exists at least one control input at each point in time, despite any perturbations, to keep the state trajectory in L. The set of such initial conditions is defined as the *viability kernel* [8], S, which has its roots in controlled invariance and corresponding reachability [9]. While the viability kernel can provide a strict guarantee of safety, computing the viability kernel is computationally intensive. An alternative approach to safety is through the use of *control barrier functions (CBFs)* [10]. In this case, the zero super-level set of $h(x)$ defines the safe set $C = \{x \mid h(x) \geq 0\}$ such that inside C there is an additional condition satisfied on the time-derivative, $\dot{h}$, along solutions of a control system. In particular, there exists a control input such that $\dot{h}(x,u) \geq -\alpha h(x)$, for some positive α. This ensures that the state never leaves the safe set C. While finding a control barrier function is challenging, in most cases, a heuristic $h(x)$ is chosen and an online optimization solved to find the control input u. The safe set C and its connection with h is illustrated in Figure 4.3.

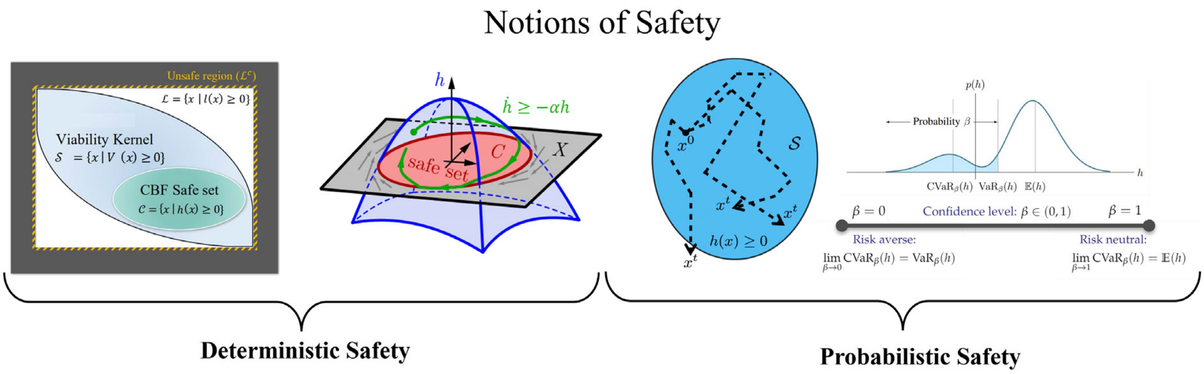


Figure 4.3: The notions of deterministic and probabilistic safety are illustrated. Deterministic: (left) illustration of the defined safety regions L,S,C (Figure adapted from [11]), and (right) the safe set C as the 0-superlevel set of a control barrier function h, and how the CBF condition ensures forward invariance of the safe set (Figure adapted from [12]). Probabilistic: (left) illustration of the evolution of a stochastic system, and (right) risk adverse notions of safety—the mean of h being positive does not capture the low probability events, while the Value at Risk (VaR) and the Conditional Value at Risk (CVaR) capture these risk-adverse scenarios (Figure adapted from [13]).

Probabilistic

In some cases, we may not require a strict guarantee of safety but rather require the safety guarantee to be satisfied with at least some probability $p < 1$. This condition may arise either because it is computationally easier or more realistic to make a probabilistic guarantee or because the inherent system dynamics or measurements are stochastic. The latter arises due to perturbation or uncertainty in the model or uncertain measurements for which a probabilistic model is available. One approach to probabilistic safety guarantees is through probabilistic safety barrier certificates [14]. In this case, given a desired confidence level p with which to ensure that the system state remains safe (i.e., stays within the safe set C), a chance constraint can be formulated as $Pr\{x \in C\} \geq p$. If the uncertainty can be modeled as a Gaussian process, then there exist explicit solutions to the chance constraint, and a feasible control input can be computed through a second-order cone optimization problem [15, 16].

Stochastic uncertainty can also be captured in the context of discrete time processes and, in this context, different notions of safety can be posed as they relate to *risk* [13]. In particular, given a safe set represented by a barrier function, h, there are multiple probabilistic interpretations of safety for a stochastic process. One could consider the mean of h and formalize ways in which to keep this mean positive—but this does not properly capture risk in a variety of settings, since there are potentially a large number of safety violations in this case. Controllers that enforce safety in this case are, therefore, *risk-neutral*. Taking inspiration from the finance and operational research communities, we can instead consider worst-case behaviors (in a distributional sense) and thereby generate *risk-averse* controllers. This can, be done, for example, via the Conditional Value at Risk (CVaR), which is a specific example of the more general notion of a coherent risk measure. As in the deterministic case, one can formulate CVaR notions of control barrier functions to ensure safety in a risk-averse fashion, i.e., ensure the positivity of the CVaR and, therefore, ensure safety in the context of low-probability events. There are still a wide range of challenges when it comes to guaranteeing probabilistic safety in real-world systems—a key example being understanding safety with respect to distributional shifts (e.g., due to changing weather conditions, time of day, and other factors that change dynamically).

Temporal and Event-Based Notions of Safety

Beyond purely deterministic or stochastic formulations of safety, there are notions of safety that are temporal in nature or that change over time due to event-based constraints. Constraints of this form appear in human-robot interaction [17]. For example, it may be desirable to ensure safety for some time interval, after which the safety constraints might change. Or, we may wish to ensure safety until a goal or event is reached, or to change controllers when an event is observed. These types of specifications can be captured in great detail through the use of temporal logic. Popular examples of specific temporal logics that are well-studied include linear temporal logic and signal temporal logic. Stochastic variants also exist, e.g., distributional temporal logic to capture complex probabilistic specifications. In all cases, safety can be encoded via the specifications, e.g., always staying in a safe region, or never reaching an unsafe region. Safety specifications can be coupled with additional constraints (e.g., that the system eventually reach a goal) and combined in a layered fashion—always eventually reaching a goal while staying in a desired safe region. It should be noted that methods deployed in the set-based definitions of safety can often be translated to this setting, e.g., through the use of control barrier functions and specifications (or fragments thereof); see [18, 19, 20, 21] for a few examples of a much broader body of work. Finally, among the different ways of verifying and enforcing temporal logic specified safety, there are methods that abstract the continuous dynamics to a discrete representation that enable the verification of temporal logic specifications [22].

Societal Interpretations of Safety

While the preceding definitions of safety rely on regions of the state-space of a system, humans instead evaluate the safety of a system by tabulating statistics of failures and using percentages to quantify safety. For instance, in 2019, traveling by car was 99.989% safe [23] while traveling by airplane was 99.99986% safe [24]. In contrast, we were only 99.8% safe from heart disease, the leading cause of death [25]. However, the range of percentages can be quite large. For instance, for drinking water to be considered safe, the percentage of lead should be under 0.0000015%. An alternative measure is *nines of safety* [26], wherein 90%, 99%, and 99.9% safe outcomes have one, two, and three nines of safety respectively. A 100% safe outcome has infinite nines of safety. The nines of safety can be captured mathematically through the formula $-log_{10}(1-p)$ with p being the probability of the safe outcome; see [26] for more details.

It is important to understand these societal interpretations and targets of safety in order to be able to reasonably deploy the associated methods and build public trust. For example, private transportation accounted for 82.5% of annual transportation-related fatalities between 2000 and 2009, resulting in 36,849 deaths [27]. If autonomous systems achieve a similar level of safety, is this sufficient for establishing public trust? Or must autonomous systems reach the level of safety for commercial transportation (accounting for 14.8% of annual transportation-related fatalities, or 6,390 deaths in the same period)? Or must they reach an even higher level of safety? This has important consequences, not just in terms of the synthesis of safety-critical methods, but in terms of the regulations that can be created around the understanding gained from these approaches.

One way to address safety at a societal level is to create standards that capture safety features, provide concrete procedures for verifying these features, and promote trust in autonomous systems. This can be done in two ways: optimizing existing standards, and creating new standards. In the former case, mathematical representations of safety (such as those described above) can be used; an example of this can be found in [28] where CBFs were used to fulfill ISO/TS 15066 regulations while maximizing performance. Alternatively, new safety standards can be created to capture new needs. An example of this is the Airborne Collision Avoidance System X (ACAS X) [29], which has become an international standard. This standard was formed through the use of notions of probabilistic safety coupled with the optimization of aircraft performance (encoded by Markov decision processes [MDPs] [30]), and was flight-tested extensively as part of the process. This is representative of the broader process of encoding safety via standards.

4.B.3 The Pillars of Safety

To quantify safety in complex systems, it is necessary to frame the ways in which safety plays a role in different contexts: at the level of control, within the broader autonomy stack, and when combined with learning. These different considerations form the foundation for the *pillars of safety* in complex systems.

Safety-Critical Control

At a fundamental level, enforcing safety in a provable fashion is best understood in the context of real-time control, since this is what will drive the actuation of a system into safe (or unsafe) regions. This includes both deterministic and probabilistic notions of safety for both continuous and discrete time systems. A variety of methods have been developed over the years for ensuring safe behavior on systems. This includes approaches from the robotics community, e.g., artificial potential fields [32]. The control community has also developed multiple methods to ensure safety, including forward and backward reachability set calculations [33], including Hamilton-Jacobi methods [34], optimization-based approaches where safety is enforced via cost functions and state constraints; reference governors that modify existing controllers to ensure safety [35]; and abstraction-based methods [36, 22]. In these settings, safety is framed as a forward set invariance problem (per the deterministic notion of safety), and controllers are synthesized via the calculated safe sets.

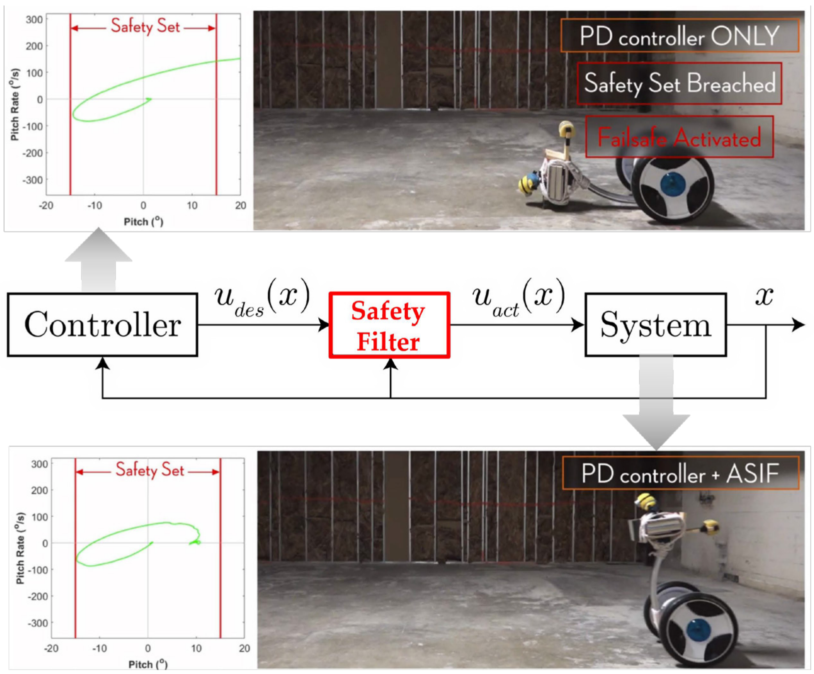


Figure 4.4: Illustration of a safety filter, wherein a nominal controller (that may result in unsafe behavior) is filtered to generate safe behavior. Figure adapted from [31].

To give a concrete example of safety-critical real-time control, consider a safe set defined as the zero super-level set of a function h with corresponding safe set $C = \{x | h(x) \geq 0\}$.

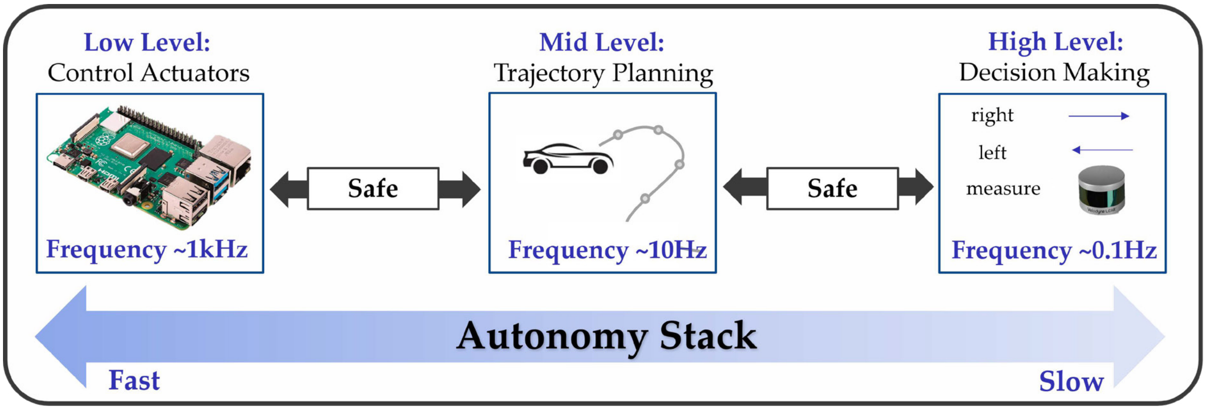


Figure 4.5: Graphic representation of the autonomy stack (adapted from [37]) where the lower level, i.e., the real-time control layer, is shown on the left, and the highest level, i.e., the decision-making layer, is shown on the right.

Per the discussion on deterministic notions of safety, the system is safe, i.e., the set C is forward invariant if there exists a control barrier function [10], i.e., for all x E C there exist control inputs u such that the CBF condition $\dot{h}(x,u) \geq -\alpha h(x)$ is satisfied. This condition can be used to synthesize a *safety filter* [31] that takes a nominal controller $u_d(x)$ and modifies it in a minimally invasive fashion such that safety is guaranteed by solving the real-time optimization problem: $\mathrm{argmin}_u \|u - u_d(x)\|^2$ subject to the constraint $\dot{h}(x,u) \geq -\alpha h(x)$. In the case of an affine control system, this is an affine constraint and thus the result is a quadratic program that can be solved in less than a millisecond. The result, as illustrated in Figure 4.4, is that even if the nominal controller is unsafe, applying the safety filter results in safe behavior on the system. Importantly, this and other safety-critical control paradigms have a variety of open problems, both in the underlying theory (e.g., synthesizing CBFs, incorporating input constraints) and the practical instantiations (e.g., uncertainty in the underlying model, sensing, and the environment). For example, recent work has utilized Hamilton-Jacobi methods to synthesize CBFs that respect input constraints [11].

Safety-Critical Autonomy

Autonomous systems leverage *autonomy* stacks which operate across multiple levels to plan, sense, and act—from real-time controllers to planning and decision-making. This requires safety to be considered across multiple time scales and multiple modes, describing the evolution of the system and each time scale. To this end, we consider safety at the following levels:

Low Level: At this level, control laws directly interact with the physical world via real-time controllers. The physical world, in this context, is described by continuous-time (nonlinear) differential equations. As a result, safety-critical control methods can be applied as previously described. Direct interaction with the physical world via digital controllers leads to the practical implementation of these controllers which can be done via sampled data or event-triggered approaches. In a safety-critical context, special care must be taken to ensure that safety is not violated in the inter-sample or interevent intervals. Finally, this real-time layer must interface with the higher layers in the autonomy stack in a safety-critical fashion. For example, it may be desirable for the safety-critical controller to track specified reference signals from the trajectory planning layer (via u_d). Certificates that these reference signals can be tracked must be provided to ensure feasibility.

Middle Level: At this level, trajectories are generated that encode desired behavior in a dynamically feasible fashion. This is often posed as an optimization problem, with a cost function encoding the control objectives (and stability constraints), and the constraints imposing dynamics and, importantly, safety and input constraints. In its most general form, the result is a constrained nonlinear optimization problem. But to realize this in a computational fashion, a linear model is often leveraged (e.g., via sequential linearization). The result is an MPC that can be used to generate desired trajectories that are passed to the real-time control level. These are safe due to the enforcement of state constraints, and the use of low-level controllers ensures safety at the continuous dynamics level. The desired behavior is then obtained from the decision-making level.

High Level: Decision-making occurs at the highest level (with the slowest update rates). This is often achieved with discrete level representations, e.g., "grid world" abstractions of the system behavior. This has a long history in planning for navigation in robotic systems, from RRTs to probabilistic road maps. From a deterministic perspective, formal methods can be employed to ensure that complex specifications are achieved, as framed via temporal logic. In a stochastic setting, MDPs and variants thereof, e.g., POMDPs, can be used to quantify uncertainty present at the decision-making level (including uncertainty in sensing and perception modules). In all cases, the decision-making level dictates the ability of an autonomous system to achieve its overarching goals, subject to safety constraints. Safety, in this context, can be encoded at the specification level (e.g., always avoid certain regions in space) or via analogues of CBFs wherein safety is guaranteed in a probabilistic sense via risk adverse approaches, e.g., CVaR. The decisions made at this level are passed down the autonomy stack to guide mid-level, and ultimately low-level, system behavior.

Safety-Critical Learning

Given a safety-critical autonomy stack, there are unique challenges and requirements for learning in this context [38]. It is essential to understand the role of learning at each layer of the autonomy stack, i.e., the decision hierarchy as it relates to learning.

At the *low level*, i.e., the real-time control level, execution speed matters. Thus, one must balance the tradeoff between time-efficiency and model accuracy. This learning must be efficient to execute—often this is accomplished by learning "residual" terms that quantify the gap between a nominal model and the observed data. Learning this residual term enables fast adaptation when there are rapid changes in the environment. Importantly, and especially at this level, this can be done in a way that guarantees physical properties of the system. For example, in the context of safety-critical control, one can learn at the level of "certificates" for safety, e.g., CBFs, wherein the residual quantifying the gap between the expected model-based value of the CBF and the observed CBF is learned [39, 40]. This is a scalar quantity, and as such it requires a minimal number of iterations to learn. The learned quantity can also be deployed on hardware at fast loop rates. At this level, the state must also be estimated—this can be done via classic methods in control, but can also be achieved with learning when perception leads to state estimation, e.g., when using vision to estimate depth in the context of collision avoidance.

At the *middle level*, learning can be used to plan multi-objective and multistep tasks. Specifically, learning can be deployed at the trajectory generation level to achieve a variety of objectives over a time window. This can be done in a model-free fashion, e.g., via reinforcement learning, or in conjunction with model-based approaches, e.g., learning the residual in a dynamics model and then employing traditional optimization-based methods. At this level, learning can also be integrated with model-based control methods that operate over time horizons, e.g., MPC [41] and adaptive control [42]. Again, this involves learning the residual dynamics and/or system parameters and encoding uncertainty in these learned quantities into the constraints [43, 44, 45, 46]. In all cases, by exploiting structure, safety-critical constraints can be added, and the result is the ability to shield learning such that the output is safe. For example, safety can be guaranteed in reinforcement learning by viewing it as a discrete-time dynamical system and avoiding "bad" regions in space when the learning is occurring (e.g., via discrete-time control barrier functions). Analogously, if learning is used in conjunction with MPC, safety can be guaranteed via state constraints in the corresponding optimization problem.

At the *high leve*l, collecting data is expensive and time-consuming since it typically comes from information-rich sensors, e.g., cameras and lidar. As a result, this level typically uses more abstract representations of sensed information. When using cameras and lidar, these are often obtained via neural networks (e.g., detecting pedestrians or classifying objects in an image). But not much is known about those components and how they contribute to a loss of safety. Therefore, the decision hierarchy can be leveraged, with the expensive (in terms of data and computation time) learning elements playing a role at the highest level, since this layer operates at the lowest frequency. Additionally, the faster elements of the autonomy stack can be utilized to keep the system stable and safe while the data-rich information sources are processed, e.g., via measurement robust CBFs [47]. Finally, it is at the highest level that probabilistic notions of safety can be deployed in conjunction with the learning process. For example, uncertainty due to learning from sensing can be encoded stochastically at the highest level via the representations described above, e.g., POMDPs wherein safety can be guaranteed in the context of covariant risk measures.

4.B.4 Guaranteeing Safety for Systems

The pillars of safety give us the means to synthesize safe behavior on systems. The corollary is also critical to understand: Given a system with unknown and black-box elements (e.g., machine learning components), how does one guarantee safety? The key is to apply safety-critical methods externally, thereby allowing a system to be robust failure, as when failures occur in the learning components.

Safe Deployment with Black-Box Elements

The ability to generate safe behaviors for systems, even in the more general context of complex autonomy stacks, naturally leads to methods for safely deploying systems with black-box elements. A specific example of this is shown in Figure 4.4 in the context of safety-critical control: a safety filter was constructed utilizing CBFs which took a nominal signal and modulated it to ensure safe behavior. Importantly, the nominal "controller" can be a black-box element, as no knowledge of this block is needed except for its output (u_d). That is, this controller may be a machine learning component without any guarantees of correct (safe) behavior. But by applying the safety filter, safety can be guaranteed. Similar ideas can be applied throughout the autonomy stack via methods developed to achieve safety-critical autonomy. For example, by combining these methods with adaptive control, these ideas may be extended to the case when parametric uncertainties occur in real time. This illustrates a key paradigm of the safety-critical framework: the ability to guarantee safety in a system even in the presence of unknown and potentially malicious components.

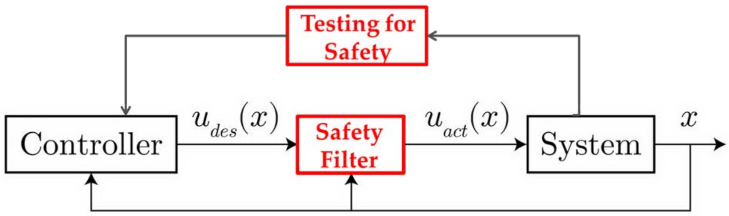


Figure 4.6: Diagram illustrating the generation of tests for safety that are filtered through safe controllers.

Testing and Evaluation of Unknown Components

The ability to generate safe behavior also gives us a way to test and evaluate (T&E) a given system for its ability to meet a set of specifications—ranging from safety (framed as forward set invariance) to general temporal logic specifications (e.g., reach-avoid), even in the presence of unknown components. There is a long history of verification, wherein tests evaluating properties can be synthesized via falsification [48] or reachability analysis [49]. Recent results focus on T&E to specifically establish safety. For example, control barrier functions can be used to encode signal temporal logic specifications [18]. As such, they can be used to monitor the satisfaction of these specifications and synthesize tests that maximally stress a system with regard to given specifications. For example, one can form a min-max game utilizing CBFs that induces a test that is minimally safe even when the controller attempts to maximize for safety [50] (as illustrated in Figure 4.6). Generalizing this methodology points to a critical application domain: leveraging safety-critical methods as *monitors* of system behavior as well as *generators* of tests that certify safety.

4.B.5 Safety for Human-in-the-Loop Systems

Safety becomes critical—a matter of life and death—when humans are in the loop. The role that humans play in CPHS is described further in Section 4.D. There are three levels at which humans interact with CPS: at the physiological level, in autonomous systems, and in population-wide networks. Safety is critical at all of these levels, but in different ways. For example, at the physiological level, safety must be considered in contexts in which humans are in direct contact with robotic systems (as with robotic assistive devices). Thus, this interaction must factor into the underlying safety-critical controllers as well as the trajectory and planning layers. Humans interacting with autonomous systems provides another example where the human might form the basis for the "desired" signal in a safety-critical controller that must be modulated to ensure safety; see [51]. Finally, at the population level, safety can be understood in the aggregate—as with large-scale power systems where safety is represented by systems operating within safe regions even under abnormal loads, or in the context of pandemics where safety-critical control can inform active intervention measures [52].

At all levels, the human's perception of safety is essential and forms the basis for *trust* in the underlying system and its behaviors. That is, if the human perceives unsafe behavior even if it is not actually unsafe (e.g., aggressive breaking on an autonomous car), it can lead to distrust. Thus, establishing trust goes beyond purely mathematical representations of safety and formal guarantees thereof, and points to a notion of *perceived safety* that must include formal representations of safety. This leads to the legal and policy aspects of safety: What levels of certification are really needed? Too many regulations can delay deployment and increase costs, while too few can erode public trust.

4.B.6 Applications and Open Challenges

The application of safety-critical methods is far-reaching. It is arguably the primary technological bottleneck to the widespread deployment of (semi-)autonomy. The public must have confidence that systems are quantifiably safe before they can trust these systems. This is certainly true of traditional applications in which safety has long been central, like aerospace, automotive, and robotic systems. But with the rise of autonomous robotic systems that interact with humans on a daily basis—from autonomous cars to robots being deployed in a wide variety of work environments—the role of safety is more critical than ever. This is especially true given the inclusion of elements that are not fully explainable, e.g., machine learning components. Novel approaches to safety can extend to a myriad of applications, including the power grid, epidemiology, transportation systems, climate change, space exploration, and economics—in short, all of the applications, societal drivers, and technological trends described throughout this road map.

With the broad scope of application for safety-critical methods comes a wealth of open problems and challenges. Here we briefly highlight open problems in each of the key topics we have touched upon:

Notions of Safety: The notions of safety presented herein largely centered on set-theoretic notions. Expanding these ideas to more general safety considerations, including trust (especially where humans are involved) poses both theoretic challenges and opportunities.

Safety-Critical Control: These methods critically depend on estimates of the state—it is through the state that safety-critical quantities are calculated, e.g., the value of a CBF and the corresponding inequality constraints. When the state is estimated from learning modules (e.g., deep neural networks processing images and video, extracting features, and inferring human intent), it is essential to determine corresponding errors and their impact on safety.

Safety-Critical Autonomy: As ML-based methods move to end-to-end solutions by capturing existing hierarchies in autonomy stacks into abstract neural network layers, it becomes increasingly hard to guarantee safety at the intermediate levels of the stack. Ensuring safety guarantees at the decision-making level, the planner level, and at the control level will become harder when they are encompassed by a single end-to-end black-box neural network.

Safety-Critical Learning: Guaranteeing safety (or constraint satisfaction) for a system with a black-box ML-based model is challenging. The technical challenges involve our inability to build mathematical frameworks for predictive models of the output of neural networks so as to provide guarantees on safety and correctness. This is especially prevalent in the “exploration (robustness) vs. exploitation (optimality)” tradeoff when safety is critical.

Safe Deployment: For systems that utilize ML-components, the domain shift between training and test datasets complicates the problem of safe deployment and makes even statistical or probabilistic estimates of failure challenging. Quantifying this, in conjunction with the use of model-based methods ensuring safe deployments, provides unique opportunities.

Test and Evaluation: Synthesizing tests that evaluate the safety of complex systems, especially those that contain black-box modules including ML components, provides challenges that are distinct from the problem of safe behavior synthesis, e.g., understanding the role of composing environment models and system behavior. It may be easier to make guarantees about systems with ML components than about the ML component itself.

Safety for Humans-in-the-Loop: This leads to an expanded notion of safety through contact where humans physically interact systems. Additionally, due to the unknown dynamics of these interactions, leveraging ML has great potential—especially when fused with model-based approaches. Finally, we can quantify higher-level interactions: characterizing intent, human preferences, and trust, and leveraging these in layered architectures.

Recommendations

For young researchers: For over a century, understanding stability has been the primary objective of control. While the stability of systems will never cease to be a vital consideration, there is now a singular opportunity to go beyond stability. As systems become more complex, a wider range of properties must be certified. Arguably, safety is the next logical generalization. There are also properties that combine and extend stability and safety that can give control theory a broader repertoire of enabling capacities. In all of these cases, hardware and the physical world can guide the next generation of theory and be used to put this theory into practice.

For funding agencies: We have an important opportunity to establish the *science of safety* by approaching safety from a formal perspective. This field can benefit from the mathematical guarantees produced by control and CPS approaches. Moreover, it has the potential to inform many other disciplines. In machine learning, for example, data-driven methods can result in unexpected (and therefore unsafe) behavior—and safety-critical controllers could ensure safe deployment of algorithms in the real world. More generally, there are expansive interdisciplinary areas that could benefit from a formal approach to safety, including autonomous systems, assistive devices, synthetic biology, and space exploration.

4.C Resilient Cyber-Physical Systems

Dan Work, Carlos Canudas de Wit, Bruno Sinopoli

> New complex CPS are being created to increase performance in mobility, energy, water, and other vital infrastructures. Designing these systems to be resilient to cyberattacks, extreme weather events, and other adverse events will enable infrastructures to quickly detect and recover from these disruptions.

Abstract Cyberattacks, climate change, and other extreme disruptions pose major threats for our complex large-scale physical infrastructures such as water, energy, and transportation systems. The rarity and variability of these events creates challenges for robust designs which operate conservatively over long periods of normality. In contrast, resilient CPS designs can allow these systems to maintain better nominal performance *and* safely operate during extreme events through real-time reconfiguration. Enabling resilience will require new methods of detecting, identifying, responding to and recovering from these disruptions. There are new opportunities for research that span system modeling and analysis, detection and identification, the design of safety-preserving methods, and reconfiguration and restoration.

4.C.1 Introduction

The term *resilience* is increasingly being discussed in the context of complex CPS within domains ranging from transportation and energy to ecology and biology. While resilience is often discussed alongside other properties of control systems such as robustness, reliability, and safety, it represents a distinct attribute. A system is robust if it can withstand perturbation without the need for adaptation. Robust systems result in conservative designs that can lead to reduced nominal performance. In contrast, a resilient system recovers in response to perturbation and restores operation quickly (with perhaps some degree of reduced functionality.) Resilience also often connotes perturbations that are severe, complex, and tantamount to a crisis. Large-scale cyberattacks are one notable example, as are concerted physical attacks and extreme weather events. A combined risk of all of these anomalies, potentially pushing systems to *extremis*, motivates the need for resilience as a key property.

As systems become increasingly large-scale, the problems of ensuring resilience are much more complex than before. Today's societal-scale infrastructures (e.g., transportation, energy, water) are spatially distributed and are composed of many intricate networked components and subsystems. To facilitate the widespread and cost-effective deployment of enabling technologies that can sense, communicate, compute, and actuate on a societal scale, it is necessary to leverage commodity equipment and shared resources. For example, smart building applications that are able to use an existing Wi-Fi network are far less expensive to deploy than systems that require a dedicated communication network. But the use of shared off-the-shelf technology poses significant risks due to the inability to introduce air gaps and its often-substandard robustness and security properties. Additionally, the extensive use of data-driven methods and technologies (e.g., machine learning) to replace analytical and model-based design tools introduces risks relating to reliability and verification of safety properties. Finally, climate change increases the probability of catastrophic events that require response and adaptation.

These trends present open challenges for the design of tools that can reason about interdependence and vulnerabilities and create systems that are resilient to large classes of disruptions. Testing the properties of robustness and resilience at scale is its own challenge. As an example, today's modeling approaches for large, open, societal-scale systems often lack the fidelity needed to serve as digital twins. At the same time, testing on real-world systems is often impossible, since they cannot be taken out of service for experimentation. These and other issues motivate the need for new approaches to unlock truly smart and resilient infrastructure systems.

4.C.2 Towards Resilience

The terms resilience and robustness have often been used interchangeably. A system is defined as robust if it can withstand a perturbation without the need for real-time adaptation. As a consequence, design for robustness needs to take potentially major disruptions into account and often results in conservative design that sacrifices performance. At the same time, implementation is easier, as it does not require the deployments of tools and algorithms to provide situational awareness and real-time reconfiguration. In contrast, a resilient system responds to perturbations that can be significant, even approaching a crisis mode, and is able to assess and identify the specific disruption in real-time. A resilient design can consequently be less conservative than a robust design, but that comes at the cost of increased complexity. The types of disruptions can be varied and severe, such as cyberattacks, natural disasters, and other unforeseen structural changes. Moreover, societal-scale infrastructure systems often operate in a complex legal and regulatory environment that can be community- and context-dependent. Because of the rarity and variability of potential disruptions, a robust approach to design can be undesirable, as the system will unnecessarily run in a conservative mode (and therefore with reduced performance) during very long periods of undisrupted operation. We argue that resilience, as defined above, is a more attractive property to pursue (notwithstanding the added design-time and run-time complexity) as it allows the system to run more efficiently during normal operation. For large-scale systems, like transportation and energy networks, this can translate to lower costs and adverse environmental impacts.

The attributes of a resilient system span design, detection, identification, response, and recovery, as shown in Figure 4.7. At each step, there are opportunities for new control approaches to enable resilience of large-scale systems. We can design systems to enable performance and security. We can increase the ability to detect attacks and disruptions by leveraging knowledge of the system. We can identify and isolate malicious agents and faulty components. Based on the identification of the disturbance, we can respond by deploying conservative measures to protect safety. At the recovery step, we can deploy responses to restore functionality.

Design	Detection	Identification	Safety-Preserving Response	Targeted Response/Recovery
1. System Design Design controller and system for performance and security	**2. Detect Attacks** Leverage system knowledge to recognize attacks	**3. Isolate Attacks** Use system knowledge to isolate malicious/faulty components	**4. Resilience I** Deploy conservative measures to preserve safety	**5. Resilience II** Deploy response to restore performance and security if appropriate

Figure 4.7: Phases of Resilience.

4.C.3 Characterization of Disruptions

Disruptions come in many forms and have different causes, both natural and manmade. A key imperative will be to classify and categorize disruptions so that that responses can be designed to address classes of disruptions, rather than single forms. Achieving scaling properties will make these problems manageable. This is particularly crucial for attaining resilience against security attacks.

Methods to Measure / Quantify Resilience

Advances in the ability to measure and quantify resilience are needed. In its simplest form, if we managed to project functionality onto a single dimensional index, we could measure loss by integrating over time, as shown in Figure 4.8. It is more likely that the index will depend upon the specific application and that several indices may be defined for the same problem.

Approaches like sensitivity analysis may be relevant for systems in which appropriate models are available, e.g., to explore the consequences of individual component failures due to external disturbances. But as the number of heterogeneous components in a system grows, exhaustive exploration of the failure combinations to determine resilience properties becomes impractical, even in even modestly sized systems. Better approaches for exploring and quantifying resilience for large-scale systems are needed. Two key elements that will need to be accounted for in resilience analysis and design are cost and constraints.

As for the first, it is obvious that it is easier to make a system more resilient with more resources, since greater redundancy can be built into it. Comparing different strategies will require some sort of normalization with respect to the use and cost of resources employed. With regard to the second, resilience will require certain properties to be guaranteed. Safety—as defined in a system theoretic manner as the ability to maintain all the relevant states within a subset of the state-space, called the safe set—is of paramount importance, in particular for human-centric systems.

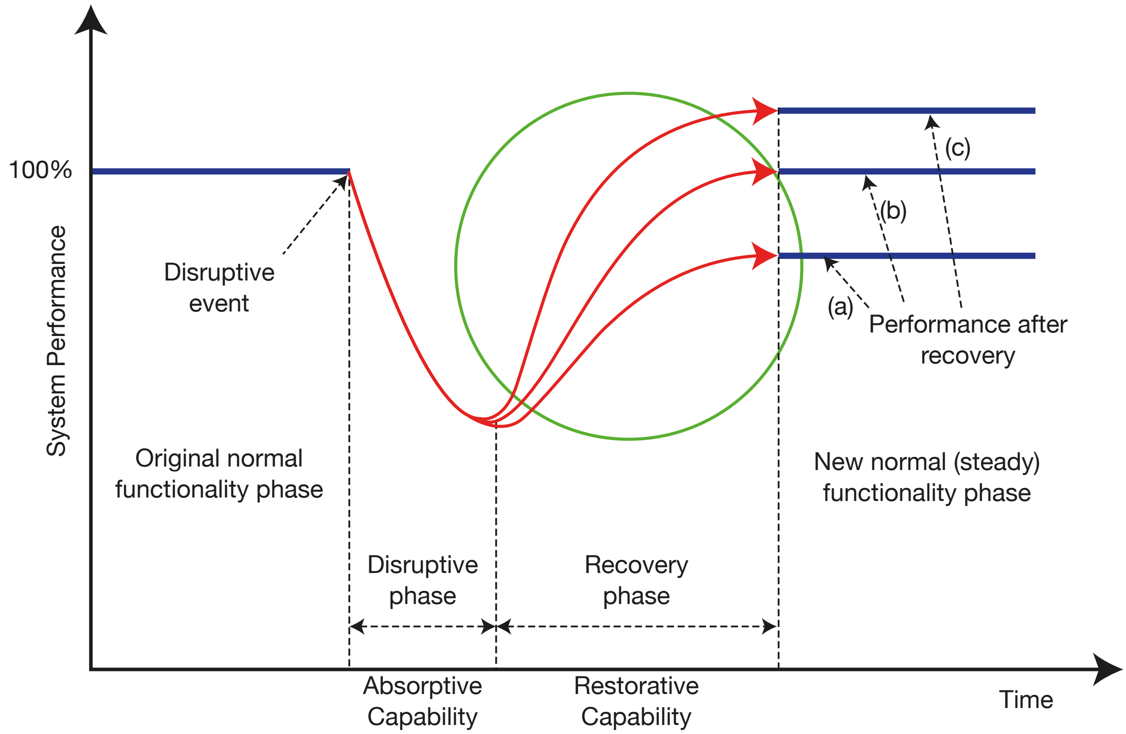


Figure 4.8: A typical resilience chart.

4.C.4 Automated Decision-Making for Resilience

The operation and management of large-scale infrastructure systems offers multiple directions for enabling resilience. Many current systems (ranging from air traffic control to water treatment) depend heavily on humans-in-the-loop to monitor day-to-day operations and handle disruptions when they occur. Given the rapid increase in system complexity on the horizon (e.g., vertical take-off and landing vehicles and UAV package delivery), human-in-the-loop management will reach a scalability bottleneck. The modeling of cyberattacks, the use of red-team/blue-team approaches, trade-offs with privacy, and overall connections with computer science-centric cybersecurity methods are all worthy of exploration. Systems that can team with human decision-makers, automate the detection and identification of threats, and reason about responses in ways that are explainable and verifiable provide directions for improving the system resilience of complex infrastructure systems.

The potential societal impact of solving these problems is immense. It is becoming apparent that critical infrastructure systems are sensitive to attacks and disruptions which could incur huge societal costs. Meanwhile, there is a risk that these issues will fall in between defined research disciplines unless a more holistic system perspective is taken. For example, there is a large scientific barrier between the control engineering and computer security research communities, and this needs to be surmounted in order to address resilience. Indeed, many cybersecurity solutions, based on encryption and authentication, are necessary to establish certain security properties.

However, many of these solutions are not currently adapted to large-scale physical systems with real-time constraints, feedback, or legacy devices with limited computational power. Furthermore, security solutions may increase overall system complexity such that safety verification is no longer feasible with today's tools. Conversely, many control and safety engineers are not familiar with the current security threat landscape and with existing solutions. It is necessary to create interdisciplinary teams to overcome these obstacles.

4.C.5 Opportunities for Control

There are multiple opportunities to advance control for resilience of CPS.

Modeling and Analysis

The complexity and openness of societal-scale infrastructure systems, and the relative lack of sensing infrastructure, creates significant challenges for modeling and assessing the impacts disruptions to these systems prior to the occurrence of such a disruption. Today's computational resources enable simulation of systems with far greater fidelity than in the past, but sources of uncertainty in the underlying systems models remain. Dynamics are often approximated, parameters are not fully characterized, and system boundaries are more challenging to identify. Digital twins, which have seen great success in well-defined physical systems, are more difficult to realize when human interaction results in fundamentally different system behaviors. When the response of individual people or large communities influences dynamics and introduces uncertainty, it is challenging to create meaningful digital twins with which generalizable experiments can be conducted. Interdependencies across other large-scale infrastructure systems reveal another promising opportunity for advancement in modeling and analysis. A systems-theoretic view and resulting control-oriented models will prove useful in the design of resilient CPS. One concrete opportunity is data-driven modeling, analysis, and control (seen as a broad generalization of system identification and adaptive control) that incorporates physical constraints and is capable of providing property guarantees.

Detection and Identification

Several opportunities may arise from leveraging methodologies from control-oriented Fault Detection and Identification (FDI) and enriching them with techniques from machine learning and security. In the case of security attacks, novel methodologies are needed to deal with intelligent attackers who may be capable of defying passive detection techniques. Over the last decade, the control community has made tremendous progress in this area, but significant work remains for when reliable system models are missing.

Safety-Preserving Methods

In CPS, it is important to preserve certain properties—often related to safety—at most or all times. While this can be achieved for normal conditions via proper design of decision-making, it is more challenging to devise safety-preserving responses to detected anomalies. Novel methods (such as those inspired by Lyapunov analysis like barrier functions) and verification techniques can be used to devise algorithms capable of preserving safety in the face of anomalous system conditions. This becomes even more challenging when such methods cannot leverage reliable models and have to resort to the use of data-driven techniques. Both stochastic and deterministic methods are relevant, depending on how important it is to preserve the property at all times.

Reconfiguration and Restoration

Concepts from optimization and system architectures can be used to develop methods of reconfiguring systems after anomalies have been detected, identified, and isolated. Economic analysis can prove useful in determining the amount of resources to commit (which will inform the length of recovery and the level of desired functionality at the end of the process).

Recommendations

For young researchers: In large, complex CPS, resiliency is a desirable property that is distinct from robustness. Resiliency is required to mitigate the consequences of rare but severe disruptions such as major cyberattacks and extreme weather. Rather than simply maintaining performance, resilient systems identify, detect, isolate, respond to, and ultimately recover from these disruptions. Adding resilience will reduce costs and improve performance, but this will require new methods of enabling real-time adaptation of complex and critical systems.

For funding agencies: Society's critical infrastructure systems are growing more complex and at the same time face more extreme disruptions. While robust systems maintain performance throughout minor contingencies, resilient systems are able to assess specific major disruptions and design specific responses in real time. Resilient design has posed many open questions in control. If these are addressed, critical systems can continue to support societal needs over the decades to come.

4.D Cyber-Physical-Human Systems

Sandra Hirche, Aaron Ames, Tariq Samad, Angela Fontan, Françoise Lamnabhi-Lagarrigue

> Cyber-physical-human systems (CPHS) is an emerging field with diverse applications in fields like in healthcare, logistics, production, and infrastructure systems, where new technical challenges toward the understanding and designing the interaction between control systems and humans are addressed. Important research questions concern the computational modeling of humans as a basis for model-based control, as well as the creation of control policies that are safe, robust, adaptive, and trustworthy.

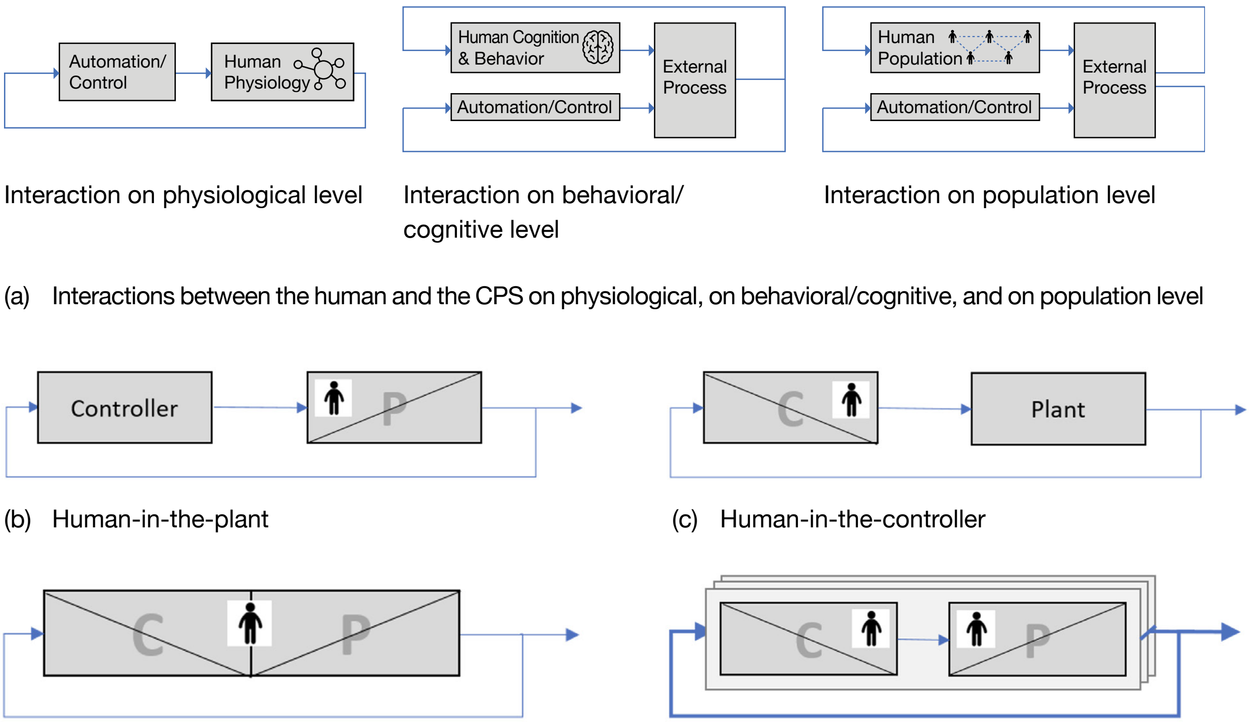
Abstract With steady advances in CPS, the interaction between humans and technical systems becomes ever more complex and traditional boundaries between humans and technology are blurred. The emerging field of CPHS grapples with a range of new technical challenges for designing interactions between control systems and humans, as well as broader questions of social domains. The applications are highly diverse; examples include artificial pancreases, prosthetics, robotic assistance devices, human augmentation technologies, smart homes, autonomous cars, and smart infrastructures. One key research challenge is characterizing models of humans and how they adapt during interaction. This will take multidisciplinary efforts to derive computational human models for prediction and control. Depending on the application and interaction level, this may require physiological modeling, cognitive-behavioral modeling, and modeling of decision-making in communities. The involvement of human beings implies significant uncertainties in the prediction of systems and also necessitates strict safety requirements. One of the main challenges in human-centered control design is developing methods that are safe, robust to uncertainty, efficient, and capable of learning and adaptation. Trustworthiness, ethics, and liability are vital dimensions as well.

4.D.1 Taxonomies of CPHS

CPHS are spread over a multitude of highly diverse application domains from biomedical engineering via autonomous robotics to infrastructure systems. Thus, the field is highly cross-disciplinary and leads to many different options for categorization. Here, we briefly introduce two ways of classifying CPHS based on i) the interaction between the human and the CPS, and ii) the role of the human within the control loop. Each taxonomy is based on some simplification and the boundaries between the different categories are fluid. An illustration is proposed in Figure 4.9.

Categorization Based on the Interaction Between Human and CPS

The interaction between CPS and humans may occur at various levels: i) at the level of physiological signal interaction, ii) on an individual cognitive and behavioral level, and iii) on an aggregate population level. The type of interaction is also strongly related to the modeling frameworks for human behavior and decision-making, from a micro (physiological, including neurons and organs, and cognitive/psychological) to a macro (interconnected individuals/community/networks) perspective.



Interaction on physiological level

Interaction on behavioral/ cognitive level

Interaction on population level

(a) Interactions between the human and the CPS on physiological, on behavioral/cognitive, and on population level

(b) Human-in-the-plant

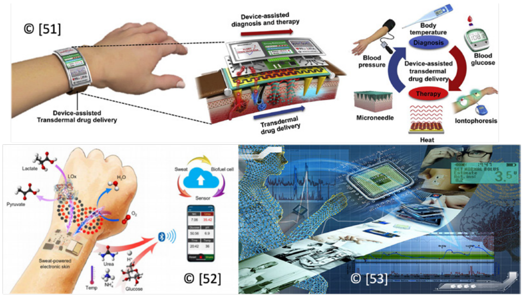
(c) Human-in-the-controller

(d) Human-machine control symbiosis

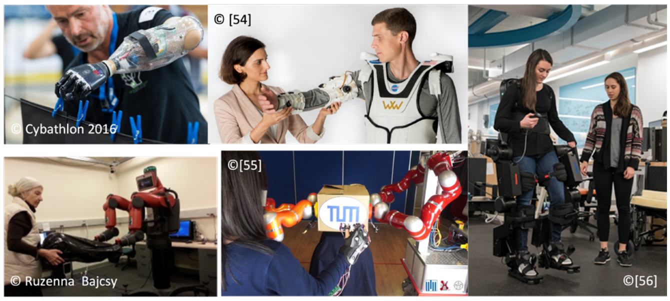
(e) Humans-in-multi-agent-loops

Figure 4.9: Taxonomies of CPHS. (a): Taxonomy based on the level of interaction between human and CPS and type of human modeling: from physiological to behavioral/cognitive and to population level. (b) to (e): Taxonomy based on the role of the human within the control loop architecture (Image source: [53]; reprinted with permission).

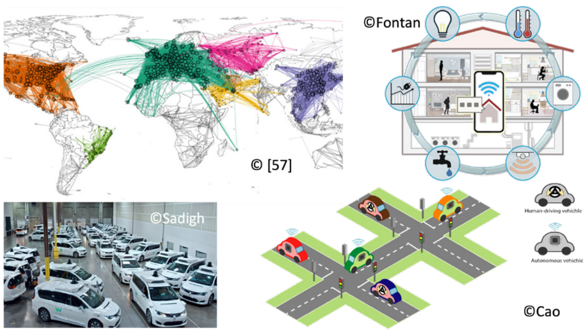
i. **Physiological Interaction:** Here, the CPS interact with the human on the level of bio-signals. That is, the controller utilizes sensed data from the human to make decisions. Bio-signals include bioelectrical activity in the brain, nerves, and muscles (e.g., EECs and EMGs), pulse, heart rate, galvanic skin response, respiration, and biochemical components such as the insulin in the blood. The CPS often augment the human body on a cell to organ level with the primary goal of achieving symbiosis between human and machine. Biomedical applications drive research in this area with the aim of providing affordable, personalized treatment of diseases (see also Section 2.B). Important applications include i) automated drug delivery, e.g., anesthesia control and artificial pancreases, ii) control of electroceuticals, e.g., deep brain stimulation, iii) wearable sensors and actuators, e.g., for spinal stimulation, and iv) smart robotic assistive devices, e.g., prostheses, brain machine interfaces, and general human augmentation. This last category overlaps with higher-level behavioral interactions in cases where interaction is achieved through the device in addition to directly sensed bio-signals from the human. See Figure 4.10a.



(a) Physiological interaction



(b) Cognitive (behavioral) interaction



(c) Population-wide interaction

Figure 4.10: Selected applications associated with physiological, cognitive (behavioral), and population-wide interaction in CPHS. Image source: [54, 55, 56, 57, 58, 59, 60].

ii. **Cognitive (Behavioral) Interaction:** In this setting, the human is considered to be a decision-maker who interacts with the CPS in a *cognitive* fashion via behaviors. The interaction signal is then communicated through the device itself or measured via external sensors, e.g., haptics, gestures, language, or even emotional expressions. The main objective for control is to design CPS behaviors that allow intuitive interactions resulting in safe and appropriately performing systems. Key applications include i) human-robot collaboration in manufacturing and construction, ii) robotic rehabilitation and assistance, iii) autonomous and semi-autonomous vehicles, iv) the operation of complex engineering systems, e.g., process plants, and v) smart homes and ambient environments. See Figure 4.10b.

iii. **Population-Wide Interaction:** Here, the focus is on large-scale interactions between networked human populations and CPS. The general objective is to facilitate the reliable, resource-efficient, and sustainable operation of infrastructure systems, either via direct control measures (policies and rules) or by incentives (see also Sections 2.C and 2.D). Important applications include i) traffic management, ii) power and general public infrastructure systems like smart cities, iii) the Sharing Economy, iv) the management of large plants, and v) the control of infectious disease. See Figure 4.10c.

Taxonomy via Role in Control Loop Architecture

An alternative taxonomy is proposed in [53], where the classical control loop is taken as an architectural template and allocates the human in one or more of its elements. In the first category, the human is "in the plant" and represents part of the process. Application domains include biomedical applications as well as temperature control in human-inhabited buildings. The second category considers the human as the controlling entity or part of it. Examples include drivers of vehicles, remote operators of UAVs and robots, and operators of building management systems and manufacturing plants. The third category is human-machine control symbiosis, where the human plays both these roles. One example application is smart prosthetics. The fourth category involves the non-trivial interaction of multiple humans with each other and with CPS in multi-agent loops. Examples include air traffic control and smart grids.

General Challenges in CPHS

While these three CPHS thrusts have significant differences in modeling, analysis, control objectives, and design tools, they share common central challenges. For example, due to complexity and often limited access to measurements, *human models are inherently highly uncertain*. Despite this uncertainty, the role of the human in the system means that *safety* plays a particularly important role. This leads to a key challenge: understanding uncertainty and safety in the context of CPHS. This raises fundamental research challenges and opportunities for the control systems community for:

- Characterizing models of humans and how they adapt during interaction
- Developing control methods that are safe, efficient, robust to uncertainty, and that can learn and adapt

Considerations of ethics, liability, and trust must be also considered at design time.

4.D.2 Computational Human Models for Prediction and Control

In acknowledging the impact of model-based control design, the modeling of human behavior becomes a vital aspect. But deriving control-oriented models of human behavior poses significant challenges. Here, we outline some of these challenges in the context of CPHS.

Generalizability vs. Specificity

All models—whether on the physiological level or on the cognitive-behavioral level—exhibit a high degree of uncertainty. There are significant limits on sensing, understanding behavior in long tails, and accounting for variability across individuals.

To address this uncertainly, learning-in-the-loop provides a promising research direction (see Section 4.A). However, in many applications, the strict safety requirements require a careful balance between robustness and flexibility in the modeling and learning process. To this end, the expression of model confidence plays an important role in actively modulating the learning process while guaranteeing safety. It is important to understand which parts of these models can be learned offline or be based on priors, and which parts need to be adapted online.

In general, the combination of classical and data-driven modeling approaches represents a promising route toward more accurate human behavior modeling in CPHS. In particular, the modeling of abstraction beyond classical ordinary differential equations becomes an interesting line of research (e.g., in terms of structures like symmetries). Such structures enable generalization—they can be used on data in the context of learning with abstract priors, while data can lead to the discovery of structures. Another challenge is quantifying these interactions via training data that is often sparse. This requires sample-efficient modeling approaches with sufficient robustness and generalization. Additionally, we need to account for behavioral changes over time—both through adaptation and via user-stated preferences (e.g., preference-based learning) which must be compared against confidence in a given model.

Understanding this interplay and utilizing the proper choice of priors is an important challenge for the control community. The solution will require collaboration with human sciences and machine learning communities.

The different types of interaction motivate the need for adequate models of human behavior at different levels, i.e., from models of physiological aspects to models of individual human behavior to models of group behavior.

Modeling for Physiological Interaction

The representation of complex physiology implies a fundamental challenge regarding the multiscale nature of such models, from individual cells to whole organs (see also Section 3.C). Being central to the field of biomedical science, control-oriented physiological modeling requires a strong transdisciplinary approach. Apart from the complexity of the models, limited access to observables and difficulties in online measurements further complicate model identification. The development of techniques for dimensionality reduction and reduced-order modeling becomes important in facilitating control design. Furthermore, physiological time delay needs to be considered. Interestingly, physiological models, despite their principal commonalities, can be individual to each human. Examples include understanding of brain function based on the analysis of dynamical processes and the role of network structure (see also Section 2.B), as well as insulin metabolism for the artificial pancreas [61].

Modeling for Individual Behavioral Interaction

The key challenge here is deriving computational models of human behavior dynamics and decision-making for system design and controller synthesis. This differs from other approaches, e.g., in psychology, where descriptive models dominate. In deriving computational models that describe the interaction between humans and CPS, there are no first-principle models to draw upon (as is the case for pure CPS). Additionally, these systems are highly nonlinear and nonstationary, e.g., they can be periodic in nature. Moreover, humans themselves are decision-making agents which adapt to autonomy—and this adaptation must be accounted for in the behavioral interactions. The applications of modeling behavioral interaction are broad: driver models for autonomous vehicle control in mixed traffic [62, 63], motor behavior of stroke patients for robotic rehabilitation (see also Section 3.D), passenger decision-making behavior in dynamic pricing for shared mobility on-demand services [64], and models for occupant behavior for comfort and energy efficiency in smart homes [65, 66, 67]. Similarly numerous are the theories derived from such fields as human movement science, cognitive psychology, and behavioral economics.

For example, optimal feedback control has been established as one of the guiding principles for human sensorimotor strategies. Research efforts have focused on the identification of optimization criteria (see [68, 69, 70]) as well as techniques to identify them based on data [71]. Other established theories are based on internal models [72, 73] and hierarchies of nested feedback loops, ranging from fast localized reflexes (proprioception) to motion generation (central pattern generators) to global long-term goal-setting. The latter can serve as a template for layered control architectures for control systems (see also Section 4.E).

High-level decision-making in humans has been captured, for example, in theories from behavioral economics in terms of Expected Utility [74, 75] and Prospect Theory [76, 77]. These models describe how humans choose among risky prospects. While the former considers the human a "rational" decision-maker, the latter allows for subjectivity and uncertainty in the decision-making process. Other relevant models include Bounded Rationality [78] and Decision Field Theory [79, 80] from cognitive psychology, and the Theory of Planned Behavior [81, 82] from social psychology.

Modeling for Population-Wide Interaction

Networks of interacting individuals represent the dominant modeling framework for modeling human behavior on a population or group level. Networked dynamical models are proposed to explain and understand the collective emergent behavior arising from the behavior of single individuals. Relevant applications include trade markets, biological networks, and infrastructure networks [83].

Another notable example is given by opinion dynamics and collective decision-making over social networks; that is, on networks of interacting individuals who exchange opinions and arrive at decisions as a community. The social ties and interactions among individuals are captured by a graph; state variables represent their opinions and a dynamical model describes the exchange and evolution of opinions over time. Several models have been proposed in the literature on opinion dynamics in the last five decades, such as for example the DeGroot model [84] and the Friedkin-Johnsen model [85] (see [86, 87, 88] for a review).

A further example concerns the modeling of epidemic spread, in which models of local disease dynamics are combined with a disease transmission network represented by a graph (see Section 2.A). More comprehensive discussions of social network dynamics in the context of CPHS are given by [89] (which draws attention to human decision-making mechanisms in social diffusion and innovation processes), and by [90] (which highlights the role of online platforms' infrastructures on the dissemination of content and the spread of information). Particularly in the context of infrastructure systems and sharing economies, the interaction of groups of individuals with algorithms and platforms plays a key role. Markets and responses to incentives (and their influence on global convergence properties) are studied within the framework of smart cities, financial networks, and electric mobility and traffic networks. Examples include the design of interventions to reduce financial contagion [91], the design of policies for incentive-based transportation demand management [92, 93, 94], and demand response in energy markets [95, 96].

4.D.3 Human-Centered Control Design

Human-centered control design for CPHS reveals novel challenges and opportunities for the systems and control community.

Safety and Trust

The safety-criticality of many CPHS application domains (many of which feature a high level of uncertainty) is one of the utmost challenges for control synthesis. Principled approaches are needed to strike the right balance between the required robustness of closed loop control systems to unmodeled uncertainties and the required adaptivity to individual dynamics. Online learning approaches within the closed loop hold promise for achieving the required level of personalization to improve user acceptance and usability. However, when it comes to learning approaches—especially for online and active learning within the loop—safety, convergence, and performance guarantees still represent a largely open research challenge (see Section 4.B). Here, certificate-based methods for safety and stability can be leveraged in a CPHS context via learning models of behavior interaction. In addition to safety (often defined as the system staying in a safe region of operation), trust also plays a vital role in user acceptance. An important interdisciplinary research question that requires input from behavioral psychology is: How can trust be enforced and modulated through control?

Architectures

A further control design challenge common to all three interaction classes is the spatial and temporal multiscale nature of the underlying models. As noted earlier in the context of modeling for behavioral interaction, this leads to notions of architectures that are layered and hierarchical. At each layer, different control methodologies are required. For example, at the highest layer—capturing human cognition and planning in the CPS—we can leverage discrete representation. We can thereby leverage formal methods, MDPs, POMDPs, and formal methods for achieving a given specification. At the middle layer—corresponding to trajectory representations of behaviors—we can utilize tools from optimization, including optimization-based model predictive controllers (MPC). Finally, at the lowest layer—encoding real-time feedback controllers—classical control methods that leverage real-time force interactions (e.g., Lyapunov-based controller synthesis) can be employed both for linear and nonlinear systems. In all cases and at all levels, models of the human at different timescales and representing different classes of behaviors can be used. This includes the use of data-driven approaches.

Validation

Another important question is: How can we evaluate human-centered control designs? Closing the gap between theoretic instantiations and real-world implementations to validate CPHS is a tremendous challenge. We must go beyond traditional validation and verification (V&V) methods due to the human-in-the-loop and the resulting safety-critical consequences. Because of high interpersonal variability, it is rarely sufficient to perform experiments with only one human subject. Multiple subjects need to perform the testing in order to achieve empirical results of statistical significance. Particular care needs to be exercised in the choice and sequence of experiments and in the choice of subjects. This points to the importance of experimental design in the V&V of CPHS. Specifically, it reveals important opportunities in the domain of T&E. It also implies collaborative opportunities between the control community and the social sciences with the common goal of validating human-centered control designs.

Control for Interaction

A unique aspect of CPHS is the interaction between humans and cyber-physical components. A key challenge is comprehending the role of closed-loop feedback in this interaction, especially when these feedback loops take different forms and occur at different timescales. This requires leveraging an understanding of architectures, as described above, and the corresponding impact on associated feedback loops. For any interaction, at any layer of the architecture, the role of sensing and signals in the feedback loop must be quantified: How certain is the sensing, can the essential attributes of the system be sensed, and how are these signals utilized in the controller? Answers to these questions can vary depending on the timescale at hand. Crucially, one must account for a variety of ways in which the signals and resulting controllers can degrade via the interaction with humans, including time delays and rate-limited control. Signals quantifying the interaction can also serve as a basis for learning and data-driven methods, for which the above questions take on further significance. Learning from inaccurate signals will result in inaccurate learned quantities. Finally, at the control synthesis level utilizing interactive signals, quantifying the objectives and constraints must be paired with the type of interaction and the layer of the architecture for which it occurs. The challenges in this space sit at the core of CPHS: good controller synthesis requires good information about these interactions.

Ethics, Liability, Trustworthiness, and User Acceptance

Going beyond traditional applications associated with CPHS, we can consider the formulation of control systems that can inform policy-level decision-making. For example, we can consider the ethical implications of public policy—representing different policies as control inputs to a population-wide CPHS—and employ data-driven models to simulate the consequences of these policies (see Chapter 7 for a deeper examination of these questions). These considerations raise issues related to privacy; we must consider the privacy of human data and how we can learn under these privacy constraints (e.g., incorporating aspects of federated learning.) In addition, when considering the impact of CPHS on humans, it is essential to ensure fairness and avoid bias—thus expanding what is being guaranteed beyond traditional metrics. Other important issues include user acceptance, explainability, transparency, and trustworthiness. In the end, to enable the adaptation of new technologies that tightly couple with humans, all aspects of interaction must be considered in a way that ensures confidence in all components.

4.D.4 Opportunities for Systems and Control

As illustrated throughout this section, there are numerous opportunities for the development of new control methods in the context of CPHS. Here we briefly summarize the key observations and the resulting open problems.

Modeling: At the core of controller development are the models used to synthesize these controllers. The uncertainty present in the human component of the system requires new paradigms in understanding the models of these systems. These can be understood from a hierarchical perspective via a layered architecture that operates at different timescales and utilizes different modeling techniques. At all levels, data-driven methods must be utilized in order to properly represent the uncertainty inherent in humans. These data-driven approaches can leverage novel forms of sensing, system identification, and machine learning. Special care must be given to the inherently nonlinear and nonstationary behaviors of humans, the difficulty in collecting data, and privacy and ethical considerations. In all cases, first-principle models can be combined with data-driven approaches to create a complete picture of the human system and set the stage for the synthesis of controllers that utilize these models.

Controller Synthesis: CPHS models provide a basis for the synthesis of controllers. As with the models themselves, these controllers must be attuned to the layered nature of these systems. The specific control methodology will depend on the layer of the hierarchy and the nature of the human interactions occurring at that level. Additionally, due to the uncertainty present in the human component of CPHS, data-driven models must be incorporated into the controller synthesis—from offline model identification and learning to the use of real-time sensing in the loop. Safety is critical in the controller synthesis problem, potentially requiring a shift from stability being the primary objective. Additional considerations that are unique to CPHS include establishing trust, protecting privacy, and resolving ethical conflicts. Validation of synthesized controllers requires new paradigms in V&V and T&E that account for the human-centered nature of CPHS, including limited data, adaptation, the use of sufficient numbers of subjects, and the corresponding experimental design. In summary, CPHS share many of the overarching challenges common to next-generation control systems, while presenting unique risks and opportunities due to the central role of the human. CPHS have the potential to have a clear and demonstrable impact on quality of life at the personal and societal level.

Recommendations

For young researchers and funding agencies: Important research questions concern the development of adaptive, data-driven control approaches that are probably safe, robust to uncertainties, and attuned to trustworthiness, intuitiveness, and ethics. Important challenges in the biomedical and cognitive sciences include the derivation of computational human models that are suitable for online prediction and control.

The safe and seamless interaction of human and smart technologies represents a significant opportunity—and a formidable challenge. The concept of CPHS recognizes the fact that ultimate societal outcomes of future CPS technologies will depend on a deeper understanding of the interactions between CPS and humans.

4.E Control Architecture

Nikolai Matni, Na Li, John C. Doyle, Anders Rantzer, Marie Csete, Steven Low, Aaron D. Ames

Societal-scale control systems are composed of a wide variety of elements, such as sensors, actuators, computers, communication devices, algorithms, software, and human-machine interfaces. Control architecture describes how these components are connected and how they interact. The goal of this section is to promote the study of control as a control architecture and the development of an integrated, unifying framework of architecture.

Abstract There is a growing awareness that architectural decisions play an important role in deploying control technologies in real-world societal systems. Virtually all contemporary engineering systems of moderate complexity (e.g., aviation, processor design, and autonomous vehicles) take a top-down, hierarchical, and layered approach to architectural design. However, today's societal-scale systems (e.g., power grids, social networks, and policy) are the result of an evolutionary, bottom-up, and ad-hoc design—and this can result in unforeseen, mysterious, and at times catastrophic failures. The goal of this section is to promote the study of control as a control architecture and the development of an integrated, unifying framework of architecture. In this section, we point to methods that show particular promise in the study of control as a control architecture, including layered multi-rate controller design via timescale separation, intralayer actuation/sensing/communication architecture and controller co-design, and the application of these concepts to redesign the power grid to integrate distributed renewable energy resources.

4.E.1 Introduction

In its broadest sense, the term *architecture* simply describes the *structure* of something. Our goal in this section is to identify design principles for successful architectures in the context of control systems, as well as to quantify the consequences of this control architecture along a variety of axes. To make the concept of control architecture accessible, we start with an example of a layered control architecture motivated by multi-rate control. We then shift our focus from cross-layer to intralayer architecture design. Finally, we ground our discussion in the context of a crucial societal-scale system architecture design problem—upgrading the power grid. We argue that quantitative study of the role and design of architecture in complex systems is essential to creating sustainable infrastructures that support societal needs. Systems-level thinking in terms of layering, feedback, and abstraction is crucial to understanding control architecture at a societal scale. New tools will be required to deal with currently intractable problems, and control theory will be central in the development of this theory of architecture. While an integrated, unifying framework of architecture and of control as a control architecture remains a long-term objective, we point to some promising methods and opportunities for advancement.

What Is Control Architecture, and Why Should We Care?

All of the societal systems discussed thus far are composed of interconnections of interacting elements, such as sensors, actuators, computers, communication devices, human-machine interfaces, humans, algorithms, and software. Control architecture describes how these system components are connected and how they interact. There is a growing awareness that *architectural decisions* play an important role in putting control technologies into real-world societal systems. While virtually all contemporary engineering systems of moderate complexity (e.g., aviation, processor design, autonomous vehicles, power grids) take a top-down hierarchical and layered approach to architecture design, many of today's societal-scale systems (e.g., social networks and public policy) are the result of an evolutionary, bottom-up and ad-hoc design, which can result in unforeseen, mysterious, and at times catastrophic failures. Furthermore, due to socioeconomic constraints, changes to these systems are subject to constraints inherited from initial designs, as in the discussion below about upgrading the power grid. This example serves as a stark reminder of the importance of design principles for control architectures.

4.E.2 Layered Architectures

Layered architectures are the most familiar and ubiquitous design pattern in complex engineered systems (e.g., planning and reflex, guidance, navigation, control, and software and hardware) and are crucial in decomposing seemingly intractable problems into well-understood subproblems. Despite their prevalence, a quantitative theory of layered architecture for control systems remains elusive. Instead, we often rely on the intuition and experience of system designers to identify the abstractions and layers defining the control architecture for a given system. This subsection will introduce some of the key technical challenges that need to be overcome to develop such a theory. We will also outline plausible next steps. We begin with a motivating example to ground our discussion.

Model Predictive Control as a Multi-Rate Layered Architecture

Consider the following finite-horizon optimal control problem (OCP) with $x(t)$ the system state, ξ a known initial condition, $u(t)$ the control input, $w(t)$ a process disturbance, $C(x,u,t)$ an instantaneous cost encoding the (negative) utility of state-input pair (x,u) at time t, and $C_T(x)$ the terminal cost.

$$\begin{aligned} &\text{minimize}_{x(t),u(t)} && \int_0^T C(x(t), u(t), t)dt + C_T(x(T)) \\ &\text{subject to} && \dot{x}(t) = f(x(t),u(t)) + w(t),\ x(0) = \xi, \end{aligned} \tag{4.1}$$

A standard approach to solving the OCP (4.1) is via nonlinear model predictive control (NMPC). A typical NMPC loop consists of repeatedly solving OCP (4.1) every τ seconds assuming the nominal dynamics $\dot{x}(t) = f(x(t),u(t))$, with initial condition ξ set to the current system state, applying the computed control signal $u(0:\tau)$, and then repeating. The implicit feedback introduced by the NMPC loop, via updating the initial condition, introduces robustness to the unmodeled disturbances $w(t)$. While the OCP (4.1) is typically intractable to solve exactly, even under the assumption of nominal dynamics, good heuristics exist for finding an approximate solution. However, due to their computational complexity, these heuristics typically lead to the update time τ being unacceptably large for applications with fast dynamics such as robotics, power-system control, and aeronautics.

One approach to resolving this difficulty is to recognize that the OCP (4.1) is in fact accomplishing two tasks at once. First, observe that the state trajectory $x(t)$ is chosen to minimize a cost function that encodes some notion of practical utility (e.g., optimal path to a goal, throughput in a network or utility maximizing power flows), i.e., the solution to the OCP is *planning an optimal state trajectory*. Second, a control trajectory $u(t)$ (or feedback policy u(x)) is computed such that the desired trajectory $x(t)$ is realized subject to the perturbed system dynamics $\dot{x}(t) = f(x(t),u(t)) + w(t)$, i.e., the solution to the OCP is also *designing a feedback controller to track the optimal state trajectory*. This observation suggests a natural decomposition of the NMPC problem into planning (trajectory generation) and tracking (feedback control) layers. Indeed, this is exactly the architecture proposed in the autonomy stack shown in Figure 4.5 of Section 4.B.[1]

A typical approach to implementing this layered architecture is to *abstract* the nonlinear dynamics of the true system, and replace them with *simplified dynamics* in the planning layer. For example, suppose that the dynamics $\dot{x}(t) = f(x(t),u(t))$ arise from a mechanical system described by double integrator dynamics $\ddot{q}(t) = F(q(t),\dot{q}(t),u(t))$, where q and $\dot{q}$ are generalized coordinates and velocities, respectively. Then, a reasonable approach would be to approximate the system dynamics as a single integrator at the planning layer, resulting in the planning problem (4.2) which we observe is *linear*, and hence tractable to solve quickly, thus reducing the update time τ. Furthermore, given a reference trajectory $(r(t),\dot{r}(t) = v(t))$, standard tools from nonlinear control can now be used to design a feedback controller that ensures that $x(t) \rightarrow r(t)$. For example, a proportional derivative (PD) controller of the form $u(x(t)) = P(x(t) - r(t)) + D(\dot{x}(t) - \dot{r}(t))$ could be designed using Lyapunov techniques to ensure the desired convergence of the system state $x(t)$ to the computed reference trajectory $r(t)$. Crucially, this feedback layer controller is simple enough to be operated in real time. The resulting layered architecture is illustrated in Figure 4.11.

$$\begin{array}{ll} \text{minimize}_{r(t),v(t)} & \int_0^T C(r(t), v(t), t)dt + C_T(r(T)) \\ \text{subject to} & \dot{r}(t) = v(t),\ r(0) = \xi, \end{array} \tag{4.2}$$

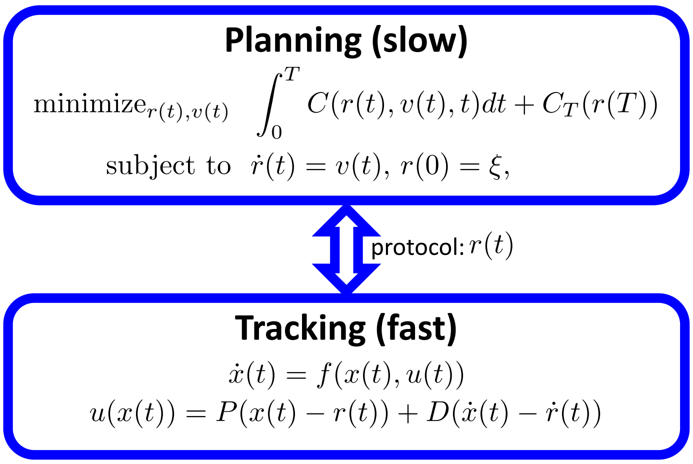


Figure 4.11: A layered architecture exploiting timescale separation to minimize approximately solve the OCP (4.1).

[1] We note here the use of the term *level* in place of the term *layer* in Section 4.B. This inconsistency highlights the need for a unified theory and language for discussing control architectures.

Challenges and Opportunities

This simple and hopefully familiar example illustrates many of the fundamental trade-offs inherent to layered architecture design. We will next attempt to generalize and formalize these concepts and highlight opportunities for their quantitative study. We begin by highlighting the advantages of the layered approach.

- **Modular:** By introducing a reference trajectory as a *protocol* between the tracking and planning layers, we modularize the controller into layers with well-defined and interpretable roles. This, in turn, allows for adjustments to be made to the planning layer without changing the tracking layer, and vice versa. This notion of modularity, achieved through an *hourglass design pattern* defined by layering and protocols (see Figure 4.12), is ubiquitous across complex engineered systems and extends well beyond planning and control to the full decision/control/software/hardware stack; see [97, Ch. 15] for a more extensive discussion of hourglasses and bowties (defined later in the section).
- **Tractable:** While the original OCP (4.1) is difficult to solve, the planning layer problem (4.2) admits an analytic solution, and the feedback layer PD controller can be easily computed in real time. Similarly, decomposing complex engineering design tasks into modular and interpretable layers allows for the global problem to be tackled via nearly independent *tractable subproblems*.
- **Multi-Rate Control:** The original OCP (4.1) naturally decomposes across two control rates: a slower rate, at which the reference trajectory $r(t : t+\tau)$ is updated, and a faster rate, at which the actual control action $u(t)$ is updated. While this timescale separation is not exploited in the NMPC approach, it is heavily exploited in the layered approach. More generally, timescale separation is widely exploited across system architecture design.

Of course, the reader should be suspicious of what is lost in imposing the layered architecture proposed above on the controller structure, as we have made no rigorous justification for our choices. We next list some possible disadvantages to the layered approach.

- **Abstraction and Layering Is Currently an Art:** In the above example, we made two seemingly arbitrary design decisions: i) we decided that there should be two layers, one for planning (trajectory generation) and one for tracking (feedback control), and ii) we used single-integrator linear dynamics in the planning problem (4.2). While experience suggests that i) this is an appropriate layered architecture for the above example, and ii) that single-integrator dynamics are useful for generating trajectories for systems with dynamics arising from mechanical systems, no rigorous theory justifying our choice exists. The role of experience, intuition, and artful choices in system architecture design becomes even more prominent as the complexity of the system grows.
- **Infeasible Reference Trajectories:** By introducing simplified dynamics in the trajectory generation problem (4.2), we can no longer guarantee that the resulting reference trajectory $r(t)$ can be tracked by the true system dynamics $\dot{x}(t) = f(x(t),u(t))$. This could have detrimental effects on safety and performance, in addition to running the risk of saturating actuators. More generally, the abstractions and simplifications used to define layers and protocols can lead to cascading errors up and down a layered architecture. This, in turn, can have catastrophic consequences on performance, robustness, and safety.
- **Suboptimal Performance:** By decoupling the design of the trajectory generation and feedback control components of the system, we restrict the feasible space of control actions our system can take. This invariably leads to a loss in performance.

However, despite these potentially significant drawbacks, layered architectures are ubiquitous across complex engineered systems. This suggests that the trade-off between the advantages and disadvantages favors the use of layering in an overwhelming majority of applications. Our goal as a research community should be to understand why and, in doing so, develop a systematic method for designing such layered architectures. We outline below some research opportunities towards meeting this goal.

- **Layering as Optimization Decomposition for Dynamical Systems:** A principled approach to deriving layered architectures based on optimization decomposition is introduced in [98]. The authors are motivated by steady-state resource allocation across the protocol stack of networked communication networks and argue that *Lagrange multipliers* associated with resource constraints are natural protocols for coordination across the resulting layered architectures. While many of the insights and approaches from [98] can be applied to general optimization problems (including optimal control problems), a comprehensive theory of layering as optimization decomposition for dynamical systems is lacking. This presents an exciting opportunity for future research. Possible paths forward include explicitly leveraging timescale separation (e.g., formalized via perturbation theory) to define layers and identifying the dynamical system analog of the resource constraint/Lagrange multiplier definition of layering suggested in [98].
- **Quantifying Performance Limits of Layered Architectures:** There is a tension between system performance and modularity. How to quantify this performance gap—and more importantly, how to identify circumstances under which this gap is small—is a crucial challenge that must be overcome in developing a quantitative theory of layered architectures. Furthermore, it is not clear *how many layers* are needed to achieve a desired goal. Too few layers may lead to intractable subproblems, whereas too many layers may lead to overly conservative and inefficient control architectures. While it is clear that there is a principle of diminishing returns at play, it is unclear how to effectively navigate this design space. One possible approach is to leverage recent tools from discrete optimization that exploit this diminishing-returns property (e.g., submodular optimization [99]) to directly tackle the problem.
- **Richer Protocols for Cross-Layer Communication and Co-Design:** Using rigid and fixed protocols between layers limits both the kinds of guarantees that can be made for the full system, as well as the co-design of layers. For example, suppose that a state constraint $x(t) \in \mathrm{X}$, for X some set encoding safety, is added to the OCP (4.1). It is not immediately obvious how to ensure that this constraint is satisfied in the layered architecture, as the planning problem (4.2) may result in a dynamically infeasible trajectory and the proposed PD controller offers no tracking guarantees. This simple example highlights the need for co-design of the planning and feedback control layers. For example, if the PD controller can be certified to remain within a bounded distance of all possible trajectories r(t) generated by the planner, e.g., via a control Lyapunov function or contraction metric, then it suffices to incorporate the constraint $r(t) \in \overline{\mathrm{X}}$ in the trajectory generation problem (4.2), where $\overline{\mathrm{X}} \in \mathrm{X}$ is a suitably tightened constraint set. While intuitive, there is no theoretical or algorithmic framework that would output this approach naturally. Yet, bidirectional flow of information between layers, as well as layer co-design, is crucial to allow for system-level guarantees and optimization.

4.E.3 Intralayer Control Architecture

Our discussion has so far focused on the challenges inherent to defining a quantitative theory of layered architectures. We now turn our attention to the architecture design challenges that exist within a given layer, with a particular focus on the feedback control layer. We first discuss the control architecture co-design problem, and then highlight existing progress in systematically defining coordination protocols in distributed systems.

Control Architecture Co-Design

In the feedback control layer, control architecture is naturally defined as the actuation, sensing, and communication network used to implement a given control policy. Once again, a natural trade-off arises: controllers with more actuators, sensors, and communication links will typically outperform controllers with sparser architectures. However, these hardware components may be costly to deploy, maintain, and power. Therefore, a balance between architectural density and closed-loop performance must be struck, suggesting that this problem can be naturally posed as a co-design problem where one seeks to minimize both control and architectural cost simultaneously.

We begin by providing a brief survey of existing methods for the *co-design* of control policies and the architectures needed to deploy them. We then outline challenges and opportunities for extending these concepts to all layers of a control system.

Fundamental Limits and Control Architecture: While $\mathcal{H}_\infty$ optimal control may be the most celebrated result from robust control [100], one of the most impactful outcomes of this line of work is the identification of the fundamental limits of achievable control for a given system (see for example [101, Ch. 6]). In particular, these results relate the unstable poles and zeros of the open-loop plant to lower-bounds on the norm of the (complementary) sensitivity function of the closed-loop system under *any possible controller*. Most detrimental to closed-loop performance are near unstable pole-zero cancellations, which correspond to near loss of controllability/observability of an unstable mode. When such a condition is observed, a natural remedy is to alter the actuation and sensing of the system (e.g., by adding more or moving existing components) to move the unstable zero. Although not explicitly characterized as such, this corresponds to a controller/control architecture co-design cycle.

Large-Scale Controller/Architecture Co-Design: While the above approach of hand-designing architectures is tractable for single systems with a small number of actuators and sensors, this task becomes unwieldy for societal-scale distributed control systems such as the power grid. The general controller/control architecture co-design problem can be informally posed as the following multicriteria optimization problem

$$\begin{array}{ll} \text{minimize} & \text{(control cost)} + \lambda\text{(architecture cost)} \\ \text{subject to} & \text{dynamics and communication constraints.} \end{array} \tag{4.3}$$

The challenge in solving problem (4.3) is that minimizing architectural cost is inherently a discrete optimization problem, as only integer numbers of actuators, sensors, and communication links can be deployed. In contrast, control policy design is inherently a continuous optimization problem. In order to overcome this challenge, two complementary approaches have been developed:

- **Sparsity-Inducing Optimization:** By recognizing that sparse control architectures can be associated with specific sparsity patterns in maps defining feedback controllers, the discrete architecture cost can be replaced by suitable sparsity-promoting convex penalties [102, 103]. For example, consider a static linear state-feedback policy $u = Kx$: if the first row K_1 of the matrix K is identically zero, then this means that the first actuator, associated with u_1, can be eliminated. Therefore, a natural penalty to add to the co-design problem (4.3) would be the group-norm penalty that promotes row-wise sparsity. This idea can be generalized to actuator, sensor, and communication network co-design through the use of atomic norm [104]-based penalties.
- **Discrete Optimization:** Another line of work recognized that certain natural control objectives, e.g., functions of the controllability Gramian, are *submodular*, and hence actuator placement problems over such objectives can be efficiently solved up to a provable approximation ratio. See, for example, [105] and references therein.

Horizontal Decomposition Within Layers

Subsection 4.E.2 proposed the use of *vertical decomposition of optimization problems* as a path toward a principled theory of layered architectures. Here we highlight that a complementary *horizontal decomposition* can be applied to design distributed architectures and protocols within a layer. A canonical example from [98] of such a horizontal decomposition is that of transmission control protocol (TCP) congestion control, which can be modeled as a distributed solution to the network utility maximization problem

$$\begin{array}{ll} \text{maximize} & \sum_s U_s(x_s) \\ \text{subject to} & Rx \leq c. \end{array} \tag{4.4}$$

Problem (4.4) assigns to each source s a utility function $U_s(x_s)$ that encodes the utility of its rate x_s, and combined network utility is then maximized subject to resource constraints encoded via the routing matrix R and the link capacity vector c. It was shown that the various flavors of existing TCP algorithms (e.g., TCP Reno and Vegas) could be reverse engineered using this framework and interpreted as practical approximations to a distributed primal dual algorithm solving problem (4.4). These insights were then used to forward engineer new TCP variants (e.g., TCP FAST) that significantly outperformed existing algorithms in certain settings. The theoretical underpinning of these approaches is the use of a dual decomposition of the resource constraints $Rx \leq c$, where once again *Lagrange multipliers* serve as an *intralayer coordination protocol* across agents in the resulting distributed optimization algorithm.

Challenges and Opportunities

While certain aspects of intralayer architecture design have effective solutions (e.g., actuator and sensor placement at the feedback layer), other challenges remain unresolved. Here we highlight some possible challenges and opportunities for progress.

- **Horizontal Decompositions for Distributed Control Problems:** The network utility maximization problem (4.4) is a static optimization problem aimed at achieving an optimal static and steady-state rate allocation. In contrast, a dynamic version of this problem would penalize transients and control effort. Existing results in distributed control do not typically take this approach, and coordination protocols are often imposed a priori. Is there an appropriate optimization decomposition technique in such a setting, and what coordination protocol does it naturally give rise to? One path forward could be to apply an operator theoretic perspective to the resulting optimization problem, which would lead to the Lagrange multipliers parameterized as dynamical systems, see for example [106].

- **Control Across Spatiotemporal Scales:** Many of the problems in this section are motivated by large-scale distributed systems. In such settings, we often have not only timescale separation, but also spatial-scale separation (e.g., in communication networks, these scales could be roughly identified as a household network, a local area network, a regional network, and a national network). How can *locality* in spatiotemporal timescales be exploited both within and across layers in a systematic way? One possible approach is to extend localized distributed control, enabled by the system-level synthesis framework [107], to more general settings.
- **Fundamental Limits of Learning-Enabled Systems:** As described in Section 4.A, learning-enabled and data-driven components will be ubiquitous across all layers. Can a theory of fundamental limits for learning-enabled control, akin to that described above from a robust control perspective, be developed to inform the co-design of learning-enabled control architectures? Promising initial results of this flavor exist wherein information theoretic lower-bounds are combined with concepts from robust control suggest a possible path forward, see for example [108, 109, 110] and references therein.

4.E.4 A Societal-Scale Example: The Future Power Grid

This subsection aims to ground the previous concepts, challenges, and opportunities in the context of upgrading the power grid.

Timescale Separation

A major source of complexity of operating a large-scale network, such as a power grid, is ubiquitous stochastic and time-varying uncertainties. Dealing with these uncertainties often requires control mechanisms that span multiple timescales. For example, frequency control maintains the frequency of a power system tightly around its nominal value when demand or supply fluctuates. Frequency control is traditionally implemented on the generation side and consists of three mechanisms. Primary frequency control operates at a timescale up to tens of seconds and uses a governor to adjust, around a setpoint, the mechanical power input to a generator based on the local frequency deviation. This implementation is called droop control and is completely decentralized. The primary control can rebalance power and stabilize the frequency but does not in itself restore the nominal frequency. Secondary frequency control operates at a timescale up to a minute or so and adjusts the setpoints of governors in a control area in a centralized fashion to drive the frequency back to its nominal value and the inter-area power flows to their scheduled values. Economic dispatch operates at a timescale of several minutes or more, and schedules the output levels of generators that are online and the inter-area power flows to minimize the overall dispatch costs. See [111] for a more detailed introduction to these three mechanisms in the context of wide-area control of power systems.

These three control mechanisms can be interpreted as three components of a layered architecture, where division across layers is naturally defined by timescale separation. Further, since traditional loads change slowly and are much more predictable, this timescale separation has been quite efficient where the mechanism at a faster timescale tracks the setpoint scheduled by a mechanism at a slower timescale, in a manner analogous to the planning/tracking architecture we proposed in subsection 4.E.2 and the autonomy stack proposed in Figure 4.5.

However, as uncertainty increases, this architecture becomes less and less economically efficient, highlighting the inherent trade-off between modularity, robustness, and efficiency. Recently, there have been efforts to modify these mechanisms to allow a lower layer control mechanism, operating at a faster timescale, to share some objectives from a slower timescale, e.g., for secondary frequency control to attain an optimal power dispatch. While these *richer protocols* between layers can help improve system efficiency, they also raise questions of potential fragilities due to breaking the clean modularity between layers. This highlights the need for a characterization of performance/robustness trade-offs in layered architectures, as well as the need for a principled approach to designing richer protocols between layers.

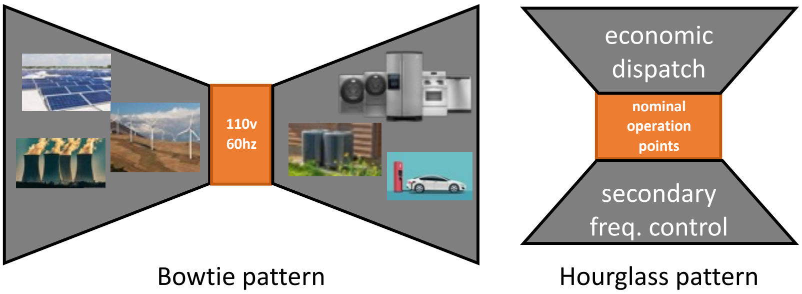


Figure 4.12: Examples of bowtie and hourglass design patterns in the power grid. Bowties allow producers and consumers within a layer to interact in a modular way via an interface: here the interface of standardized voltage/frequency at an outlet allows for a diversity of power sources and sinks to be interconnected. Hourglasses allow layers to interact with each other in a modular way via a protocol: in this case, the protocol is a nominal voltage setpoint computed at the economic dispatch (planning) layer, which is then tracked by the secondary frequency control (tracking) layer. See [112] and [97, Ch.15] for more details.

Spatial-Scale Architecture

The future grid may be similar to the traditional grid in spatial coverage, but its decision and control architecture will encompass many orders of magnitude more sensors and actuators. Traditionally, more than 90% of power is generated by less than 20,000 bulk generators in the U.S. and transported to hundreds of millions of users by high-voltage transmission systems and middle- and low-voltage distribution systems. The small number of bulk generators are connected through transmission networks, where most of the control is applied—from capacity planning to generation commitment and dispatch to frequency regulation—all factors that determine the dynamics and efficiency of the entire network. A distribution feeder, which may itself be a network that feeds a small city, is treated simply as a single network node in such a control architecture. The control paradigm is to adapt the small number of large generation nodes to meet demand over the network.

This pattern will be turned upside down in the future: Bulk generators such as coal and gas plants will be replaced by renewable sources, such as wind and solar farms. There will also be an increasing diversity of other distributed energy resources (DERs), such as smart buildings, electric vehicles, HVACs, batteries, and inverters, especially in distribution feeders. Instead of a small number of large endpoints, the new control architecture needs to coordinate dramatically more active endpoints, exacerbating scalability challenges. Furthermore, these DERs will introduce large, frequent, and random fluctuations in supply, demand, voltage, and frequency, necessitating real-time closed-loop control. Distributed generation will also upset the conventional protection architecture that assumes one-way power flow from generators to end-users. To address these challenges, on the one hand, distributed algorithms are developed in which each active endpoint could make its own decisions based on local measurements and local communications with neighbors. On the other hand, concepts like virtual power plants have been proposed where the flexibility of DERs in one region, e.g., on a distribution feeder, is aggregated in the planning and scheduling phase and then a target collective behavior can be disaggregated to individual DERs during operation. This incoming architectural shift highlights the need for principled approaches to control architecture co-design, (i.e., where actuators, sensors, and communication networks should be installed), as well as to coordination protocol design. Concretely, we believe that the future power grid will need a layered architecture wherein each layer collects and communicates aggregate information that summarizes its internal state to layers above and below. Within a layer, distributed solutions where each entity coordinates with neighbors through local information exchange takes local control decisions. We will need distributed algorithms that induce stable, robust, and efficient global behavior implemented via a control architecture (both within and across layers) that facilitates the evolvability of these algorithms.

Interaction Among Engineering, Economics, and Policy: Another complexity of societal-scale systems is the coexistence of physics and humans. Components in these systems belong to heterogeneous authorities, who may act strategically to maximize profits rather than blindly follow algorithmic protocols. Another related concern is that coordination policies may require eliciting private information from owners, e.g., as with networked appliances. These concerns necessitate the design of markets, policies, and protocols to incentivize individual entities to take actions aligned with the goals of the global system, as well as to protect users' private information. To ensure the successful coordination of the overarching systems, the engineering architecture and the economic/policy structure should be integrated and jointly designed, subject to economic and social constraints imposed by the legacy system.

This complex multicriteria design problem will necessarily require collaboration amongst experts in different areas (using different languages around architecture) including communication, computer networking, power engineering, architecture, economics, social sciences, industry, and urban planners, among many others. This convergence of disciplines within a single system once again highlights the need for a unified, general, and rigorous theory of architecture that can be used as a common language across domains. Our hope is that a well-defined theory of control architecture will make this daunting task tractable by defining suitable interfaces (within layers) and protocols (across layers) that allow for respective domains of expertise to be integrated in a modular yet synergistic way.

4.E.5 Road Map for a Theory of Control Architecture

We end this section with a summary of some of the commonalities and themes that emerge in the above discussion and offer a plausible road map for ambitious control researchers interested in demystifying the art of control architecture design.

Architecture Design Is a Multicriteria Optimization Problem: A common theme across this section is the need to balance various criteria when (co)designing a control architecture. For example, the cost function of the architecture co-design problem (4.3) is a scalarization of the vector-valued cost (control cost, architecture cost). In this context, as the two design axes are clear, the main challenge is in overcoming the mixed continuous/discrete nature of the problem. While similar vector-valued costs can be defined for designing layered architectures, we are faced with the additional challenge that some of the metrics are inherently qualitative. For example, it is unclear how a numeric cost can be assigned to the tractability or modularity of an architecture. Nevertheless, being able to quantify and navigate the Pareto surface of this design space lies at the core of a theory of layered control architecture design.

From Static to Dynamic Control Architectures: Viewing architecture design as a multicriteria optimization problem allows us to associate a point on the Pareto surface with a particular architecture and vice-versa. Traditional architecture design has looked to find so-called "sweet spots" on this Pareto surface that are best suited for the task at hand, and to then build the corresponding architecture. This focus on a single architecture was a necessity, as system architecture design was an offline task, typically realized in hardware. However, with contemporary advances in hardware/software abstraction and virtualization, many architectural components can either be implemented solely in software (e.g., a neural network used as a perception map) or hardware components can be spun up/spun down on demand (e.g., compute resources in a data-center). This opens up the exciting possibility of developing a theory of dynamic control architectures, wherein we move optimally along the Pareto surface as a function of environmental and control task properties.

Reverse and Forward Engineering Control Architectures: Questions of architecture design permeate this document and are prominently featured in sections 4.A, 4.B, and 4.D. Any progress made toward a general theory of (forward engineering) control architecture will have an immediate impact in these application domains. Perhaps less obvious is the opportunity for theoretical developments in reverse engineering existing layered architectures, such as in (synthetic) biology (see Section 3.C) and neuroscience (see Section 2.B). Reverse engineering of architectures in the internet enabled significant algorithmic and technological breakthroughs [98], and the potential for similar insights in other domains is exciting motivation to further pursue the systematic study of control architecture.

Software Tools for Control Architecture Design: Existing modeling languages for system architecture design, for example SysML [113], offer a promising blueprint to operationalizing the concepts introduced in this section. Can a control architecture version of such software packages, tailored to dynamic systems evolving in dynamic environments, be developed?

Architecture, More Broadly: While this section has focused on aspects of control architecture design that are most familiar to the control theory community, it is important to recognize that architecture is an overloaded and broadly used term across various engineering disciplines and scientific domains; see for example [114, 115, 116, 117, 118]. Nevertheless, common themes emerge across all of these, including layering and protocols. Other concepts that we did not touch upon but are crucially important include physical substrate levels (not to be confused with the autonomy levels of Figure 4.5), interfaces for the transmission of quantities within and across layers (resulting in bowtie design patterns; see Figure 4.12), and the distinction between passive and active system components. We refer the interested reader to [112] and [97, Ch.15] for an accessible discussion of many of these concepts.

Recommendations

For young researchers: There is a growing awareness that architectural decisions play an vital role in implementing control technologies within real-world societal systems. Virtually all contemporary engineering systems of moderate complexity (e.g., aviation, processor design, and autonomous vehicles) take a top-down, hierarchical, and layered approach to architectural design. However, today's societal-scale systems (e.g., power grids, social networks, and policy) are the result of an evolutionary, bottom-up, and ad-hoc design—and this can result in unforeseen, mysterious, and at times catastrophic, failures. In this section, we point to methods that show particular promise in the study of control as a control architecture, including layered multi-rate controller design via time-scale separation, intralayer actuation/sensing/communication architecture and controller co-design, and the application of these concepts to redesign the power grid to integrate distributed renewable energy resources.

For funding agencies: The quantitative study of the role and design of architecture in complex systems is essential to creating sustainable infrastructures that support societal needs. To understand control architecture at a societal scale, we need systems-level thinking in terms of layering, feedback, and abstraction. Control theory will be central to the development of this theory of architecture, and new tools will be needed to solve currently intractable problems. While an integrated, unifying framework of architecture remains a long-term vision, recent developments in optimization, controls, robotics, and systems engineering show early promise as intermediate objectives in this challenging area of research.

BIBLIOGRAPHY

[1] H.-S. Ahn, Y. Chen, and K. L. Moore, “Iterative learning control: Brief survey and categorization,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1099–1121, 2007.

[2] M. Krstic and H.-H. Wang, “Stability of extremum seeking feedback for general nonlinear dynamic systems,” *Automatica-Kidlington*, vol. 36, no. 4, pp. 595–602, 2000.

[3] A. M. Annaswamy and A. L. Fradkov, “A historical perspective of adaptive control and learning,” *Annual Reviews in Control*, vol. 52, pp. 18–41, 2021.

[4] L. Wang, A. D. Ames, and M. Egerstedt, “Safety barrier certificates for collisions-free multirobot systems,” *IEEE Transactions on Robotics*, vol. 33, no. 3, pp. 661–674, 2017.

[5] L. Lamport, “Proving the correctness of multiprocess programs,” *IEEE transactions on software engineering*, no. 2, pp. 125–143, 1977.

[6] M. Nagumo, “Über die lage der integralkurven gewöhnlicher differentialgleichungen,” *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, vol. 24, pp. 551– 559, 1942.

[7] F. Blanchini, “Set invariance in control,” *Automatica*, vol. 35, no. 11, pp. 1747–1767, 1999.

[8] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability theory: new directions*. Springer Science & Business Media, 2011.

[9] D. P. Bertsekas and I. B. Rhodes, “On the minimax reachability of target sets and target tubes,” *Automatica*, vol. 7, no. 2, pp. 233–247, 1971.

[10] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, “Control barrier function based quadratic programs for safety critical systems,” *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2016.

[11] J. J. Choi, D. Lee, K. Sreenath, C. J. Tomlin, and S. L. Herbert, “Robust control barrier-value functions for safety-critical control,” in *IEEE Conference on Decision and Control*, Dec. 2021.

[12] T. G. Molnar, R. K. Cosner, A. W. Singletary, W. Ubellacker, and A. D. Ames, “Modelfree safety-critical control for robotic systems,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 944–951, 2021.

[13] M. Ahmadi, X. Xiong, and A. D. Ames, “Risk-averse control via cvar barrier functions: Application to bipedal robot locomotion,” *IEEE Control Systems Letters*, vol. 6, pp. 878–883, 2021.

[14] W. Luo, W. Sun, and A. Kapoor, “Multi-robot collision avoidance under uncertainty with probabilistic safety barrier certificates,” in *Conference on Neural Information Processing Systems*, Dec. 2020.

[15] M. J. Khojasteh, V. Dhiman, M. Franceschetti, and N. Atanasov, “Probabilistic safety constraints for learned high relative degree system dynamics,” in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, 10–11 Jun 2020, pp. 781–792.

[16] F. Castañeda, J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, “Gaussian processbased min-norm stabilizing controller for control-affine systems with uncertain input effects and dynamics,” in *American Control Conference*, New Orleans, LA, May 2021.

[17] P. A. Lasota, T. Fong, J. A. Shah et al., “A survey of methods for safe human-robot interaction,” *Foundations and Trends® in Robotics*, vol. 5, no. 4, pp. 261–349, 2017.

[18] L. Lindemann and D. V. Dimarogonas, “Control barrier functions for signal temporal logic tasks,” *IEEE control systems letters*, vol. 3, no. 1, pp. 96–101, 2018.

[19] G. Yang, C. Belta, and R. Tron, “Continuous-time signal temporal logic planning with control barrier functions,” in *2020 American Control Conference (ACC). IEEE*, 2020, pp. 4612–4618.

[20] M. Srinivasan and S. Coogan, “Control of mobile robots using barrier functions under temporal logic specifications,” *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 363– 374, 2020.

[21] M. Ahmadi, A. Singletary, J. W. Burdick, and A. D. Ames, “Safe policy synthesis in multi-agent pomdps via discrete-time barrier functions,” in *2019 IEEE 58th Conference on Decision and Control (CDC). IEEE*, 2019, pp. 4797–4803.

[22] P. Tabuada, *Verification and control of hybrid systems: a symbolic approach*. Springer Science & Business Media, 2009.

[23] “Fatality facts 2019,” IIHS, https://www.iihs.org/topics/fatality-statistics/detail/state-by-state, 2019.

[24] “Airplane crashes,” NSC, https://injuryfacts.nsc.org/home-and-community/safety-topics/airplane-crashes/, 2019.

[25] “Deaths and mortality,” CDC, https://www.cdc.gov/nchs/fastats/deaths.htm, 2019.

[26] T. Tao, “Nines of safety: a proposed unit of measurement of risk,” What’s new, https://terrytao.wordpress.com/2021/10/03/nines-of-safety-a-proposed-unit-of-measurement-of-risk/, Oct. 2021.

[27] I. Savage, “Comparing the fatality risks in united states transportation across modes and over time,” *Research in transportation economics*, vol. 43, no. 1, pp. 9–22, 2013.

[28] F. Ferraguti et al., “A control barrier function approach for maximizing performance while fulfilling to iso/ts 15066 regulations,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5921–5928, 2020.

[29] M. J. Kochenderfer, J. E. Holland, and J. P. Chryssanthacopoulos, “Next-generation airborne collision avoidance system,” Massachusetts Institute of Technology-Lincoln Laboratory Lexington United States, Tech. Rep., 2012.

[30] M. J. Kochenderfer, *Decision making under uncertainty: theory and application*. MIT press, 2015.

[31] T. Gurriet et al., “Towards a framework for realizable safety critical control through active set invariance,” in 2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS). IEEE, 2018, pp. 98–106.

[32] C. W. Warren, “Global path planning using artificial potential fields,” in 1989 *IEEE International Conference on Robotics and Automation*. IEEE Computer Society, 1989, pp. 316–317.

[33] M. Althoff, G. Frehse, and A. Girard, “Set propagation techniques for reachability analysis,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, 2021.

[34] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, “A time-dependent hamilton-jacobi formulation of reachable sets for continuous dynamic games,” *IEEE Transactions on automatic control*, vol. 50, no. 7, pp. 947–957, 2005.

[35] E. Garone, S. Di Cairano, and I. Kolmanovsky, “Reference and command governors for systems with constraints: A survey on theory and applications,” *Automatica*, vol. 75, pp. 306–328, 2017.

[36] C. Belta, B. Yordanov, and E. A. Gol, *Formal methods for discrete-time dynamical systems*. Springer, 2017, vol. 15.

[37] U. Rosolia, A. Singletary, and A. D. Ames, “Unified multi-rate control: from low level actuation to high level planning,” *arXiv preprint arXiv:2012.06558*, 2020.

[38] L. Brunke et al., “Safe learning in robotics: From learning-based control to safe reinforcement learning,” *arXiv preprint arXiv:2108.06266*, 2021.

[39] A. Taylor, A. Singletary, Y. Yue, and A. Ames, “Learning for safety-critical control with control barrier functions,” in *Learning for Dynamics and Control*. PMLR, 2020, pp. 708–717.

[40] J. Choi, F. Castañeda, C. J. Tomlin, and K. Sreenath, “Reinforcement learning for safety-critical control under model uncertainty, using control lyapunov functions and control barrier functions,” *arXiv preprint arXiv:2004.07584*, 2020.

[41] F. Allgöwer and A. Zheng, *Nonlinear model predictive control*. Birkhäuser, 2012, vol. 26.

[42] A. M. Annaswamy and A. L. Fradkov, “A historical perspective of adaptive control and learning,” Annual Reviews in Control, vol. 52, pp. 18–41, 2021.

[43] R. Cheng, G. Orosz, R. M. Murray, and J. W. Burdick, “End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 3387–3395.

[44] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, “Learning-based model predictive control: Toward safe learning in control,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 269–296, 2020.

[45] A. J. Taylor and A. D. Ames, “Adaptive safety with control barrier functions,” in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1399–1405.

[46] B. T. Lopez, J.-J. E. Slotine, and J. P. How, “Robust adaptive control barrier functions: An adaptive and data-driven approach to safety,” IEEE Control Systems Letters, vol. 5, no. 3, pp. 1031–1036, 2020.

[47] S. Dean, A. Taylor, R. Cosner, B. Recht, and A. Ames, “Guaranteeing safety of learned perception modules via measurement-robust control barrier functions,” in Conference on Robot Learning. PMLR, 2021, pp. 654–670.

[48] Y. Annpureddy, C. Liu, G. Fainekos, and S. Sankaranarayanan, “S-taliro: A tool for temporal logic falsification for hybrid systems,” in *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 2011, pp. 254–257.

[49] G. Frehse et al., “SpaceEx: Scalable verification of hybrid systems,” in *International Conference on Computer Aided Verification*. Springer, 2011, pp. 379–395.

[50] P. Akella, M. Ahmadi, R. M. Murray, and A. D. Ames, “Formal test synthesis for safety-critical autonomous systems based on control barrier functions,” in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 790–795.

[51] B. Xu and K. Sreenath, “Safe teleoperation of dynamic uavs through control barrier functions,” in *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, May 2018, pp. 7848–7855.

[52] A. D. Ames, T. G. Molnár, A. W. Singletary, and G. Orosz, “Safety-critical control of active interventions for covid-19 mitigation,” *IEEE Access*, vol. 8, pp. 188454–188474, 2020.

[53] Human-in-the-Loop Control: Applications and Categorization,” IFAC-PapersOnLine, v. 53 (5), 2020, pp. 311-317.

[54] H. Lee et al., “Device-assisted transdermal drug delivery,” *Advanced Drug Delivery Reviews*, vol. 127, pp. 35–45, 2018.

[55] Y. Yu et al., “Biofuel-powered soft electronic skin with multiplexed and wireless sensing for human-machine interfaces,” *Science Robotics*, vol. 5, no. 41, 2020.

[56] F. J. Doyle III, L. M. Huyett, J. B. Lee, H. C. Zisser, and E. Dassau, “Closed-loop artificial pancreas systems: Engineering the algorithms,” *Diabetes Care*, vol. 37, pp. 1191–1197, 2014.

[57] Z. Kadivar, C. E. Beck, R. N. Rovekamp, and M. K. O’Malley, “Single limb cable driven wearable robotic device for upper extremity movement support after traumatic brain injury,” J*ournal of Rehabilitation and Assistive Technologies Engineering*, vol. 8, 2021.

[58] S. Music et al., “Human-robot team interaction through wearable haptics for cooperative manipulation,” *IEEE Transactions on Haptics*, 2019.

[59] M. Tucker et al., “Preference-based learning for exoskeleton gait optimization,” 2020, pp. 2351–2357.

[60] A. A. Matamalas, Joan T. and S. Gómez, “Effective approach to epidemic containment using link equations in complex networks,” *Science Advances*, vol. 4, 2018.

[61] J. E. Pinsker et al., “Evaluation of an Artificial Pancreas with Enhanced Model Predictive Control and a Glucose Prediction Trust Index with Unannounced Exercise,” *Diabetes Technology & Therapeutics*, vol. 20, no. 7, pp. 455–464, 2018.

[62] D. Sadigh, S. Sastry, S. A. Seshia, and A. D. Dragan, “Planning for autonomous cars that leverage effects on human actions,” in *Proceedings of Robotics: Science and Systems*, AnnArbor, Michigan, June 2016.

[63] E. Stefansson, F. J. Jiang, E. Nekouei, H. Nilsson, and K. H. Johansson, “Modeling the decision-making in human driver overtaking,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 15338–15345, 2020, 21st IFAC World Congress. https://www.sciencedirect.com/science/article/pii/S2405896320330159

[64] Y. Guan, A. M. Annaswamy, and H. Eric Tseng, "Cumulative Prospect Theory Based Dynamic Pricing for Shared Mobility on Demand Services," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. Nice, France: IEEE, 12 2019, pp. 2239–2244. https://ieeexplore.ieee.org/document/9029508/

[65] C. Wilson and H. Dowlatabadi, "Models of decision making and residential energy use," *Annual Review of Environment and Resources*, vol. 32, no. 1, pp. 169–203, 2007. https://doi.org/10.1146/annurev.energy.32.053006.141137

[66] H. Allcott and T. Rogers, "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation," *American Economic Review*, vol. 104, no. 10, pp. 3003–37, October 2014. https://www.aeaweb.org/articles?id=10.1257/aer.104.10.3003

[67] T. Harputlugil and P. de Wilde, "The interaction between humans and buildings for energy efficiency: A critical review," *Energy Research & Social Science*, vol. 71, p. 101828, 2021. https://www.sciencedirect.com/science/article/pii/S2214629620304035

[68] T. Flash and N. Hogan, "The coordination of arm movements: An experimentally confirmed mathematical model," *Journal of Neuroscience*, vol. 5, no. 7, p. 1688–1703, 1985.

[69] C. Harris and D. M. Wolpert, "Signal-dependent noise determines motor planning," *Nature*, no. 394, p. 780–784, 1998.

[70] E. Todorov and M. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature Neuroscience*, no. 5, p. 1226–1235, 2002.

[71] K. Mombaur and D. Clever, "Inverse optimal control as a tool to understand human movement," in *Geometric and Numerical Foundations of Movements*, J. Laumond, N. Mansard, and J. Lasserre, Eds, 2017.

[72] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, no. 12, p. 457–465, 1976.

[73] M. Kawato, "Internal models for motor control and trajectory planning," *Current Opinion in Neurobiology*, vol. 9, no. 6, p. 718–727, 1999.

[74] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1953.

[75] P. J. H. Schoemaker, "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations," J*ournal of Economic Literature*, vol. 20, no. 2, pp. 529–563, 1982.

[76] D. Kahneman and A. Tversky, "Prospect theory: An Analysis of Decision Under Risk," *Econometrica*, vol. 47, no. 2, pp. 263–291, 1979.

[77] A. Tversky and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, vol. 5, no. 4, pp. 297–323, 10 1992. http://link.springer.com/10.1007/BF00122574

[78] H. A. Simon, Models of Bounded Rationality, 1982.

[79] J. R. Busemeyer and J. T. Townsend, "Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment," *Psychological Review*, vol. 100, no. 3, pp. 432–459, 1993.

[80] J. R. Busemeyer and A. Diederich, “Survey of decision field theory,” *Mathematical Social Sciences*, vol. 43, no. 3, pp. 345–370, 7 2002.

[81] I. Ajzen, “The theory of planned behavior,” *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 12 1991. https://linkinghub.elsevier.com/retrieve/pii/074959789190020T

[82] I. Ajzen, “From Intentions to Actions: A Theory of Planned Behavior," In: Kuhl, J., Beckmann, J. (eds) Action Control. Springer Series in Social Psychology. Springer, Berlin, 1985. https://doi.org/10.1007/978-3-642-69746-3_2.

[83] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets - Reasoning about a Highly Connected World*, 1st ed. Cambridge: Cambridge University Press, 2010.

[84] M. H. DeGroot, “Reaching a consensus,” *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[85] N. E. Friedkin and E. Johnsen, “Social Influence Networks and Opinion Change,” *Advances in Group Processes*, vol. 16, pp. 1–29, 1999.

[86] A. V. Proskurnikov and R. Tempo, “A tutorial on modeling and analysis of dynamic social networks. Part I,” *Annual Reviews in Control*, vol. 43, pp. 65–79, 2017.

[87] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo, “Opinion dynamics and the evolution of social power in influence networks,” *SIAM Review*, vol. 57, no. 3, pp. 367–397, 2015. https://doi.org/10.1137/130913250

[88] G. Shi, C. Altafini, and J. S. Baras, “Dynamics over signed networks,” *SIAM Review*, vol. 61, no. 2, pp. 229–257, 2019. https://doi.org/10.1137/17M1134172

[89] L. Zino and M. Cao, “Social diffusion dynamics in cyber-physical-human systems,” in *Cyber-Physical-Human Systems: Fundamentals and Applications,* A. M. Annaswamy, P. Khargonekar, F. Lamnabhi-Lagarrigue, and S. K. Spurgeon, Eds. UK: Wiley, 2022.

[90] M. Castaldo, P. Frasca, and T. Venturini, “On online attention dynamics,” in *Cyber-Physical-Human Systems: Fundamentals and Applications*, A. M. Annaswamy, P. Khargonekar, F. Lamnabhi-Lagarrigue, and S. K. Spurgeon, Eds. UK: Wiley, 2022.

[91] G. C. Calafiore, G. Fracastoro, and A. V. Proskurnikov, “Control of dynamic financial networks,” *IEEE Control Systems Letters*, vol. 6, pp. 3206–3211, 2022.

[92] C. Cenedese, M. Cucuzzella, J. M. A. Scherpen, S. Grammatico, and M. Cao, “Highway traffic control via smart e-mobility – part i: Theory,” 2021. https://arxiv.org/abs/2102.09354

[93] C. Cenedese, P. Stokkink, N. Geroliminis, and J. Lygeros, “Incentive-based electric vehicle charging for managing bottleneck congestion,” *European Journal of Control*, p. 100697, 2022. https://www.sciencedirect.com/science/article/pii/S0947358022000905

[94] Z.-C. Li, H.-J. Huang, and H. Yang, “Fifty years of the bottleneck model: A bibliometric review and future research directions,” *Transportation Research Part B: Methodological*, vol. 139, pp. 311–342, 2020.

[95] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 2015.

[96] M. Fochesato, C. Cenedese, and J. Lygeros, "A stackelberg game for incentive-based demand response in energy markets," 2022. https://arxiv.org/abs/2204.08730

[97] K. J. Åström and R. M. Murray, *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.

[98] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, 2007.

[99] S. Fujishige, *Submodular functions and optimization*. Elsevier, 2005.

[100] J. C. Doyle and K. Zhou, "Essentials of robust control," *Prentice Hall, Englewood Cliffs NY*, 1998.

[101] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, Feedback control theory. Courier Corporation, 2013.

[102] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2426–2431, 2013.

[103] N. Matni and V. Chandrasekaran, "Regularization for design," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3991–4006, 2016.

[104] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[105] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 67–78, 2015.

[106] A. Rantzer, "Dynamic dual decomposition for distributed control," in *2009 American control conference*. IEEE, 2009, pp. 884–888.

[107] Y.-S. Wang, N. Matni, and J. C. Doyle, "A system-level approach to controller synthesis," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, p. 4079–4093, Oct 2019.

[108] A. Tsiamis, I. M. Ziemann, M. Morari, N. Matni, and G. J. Pappas, "Learning to control linear systems can be hard," in *Conference on Learning Theory*. PMLR, 2022, pp. 3820–3857.

[109] A. Tsiamis and G. J. Pappas, "Linear systems can be hard to learn," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 2903–2910.

[110] Y. Jedra and A. Proutiere, "Sample complexity lower bounds for linear system identification," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 2676–2681.

[111] A. Chakrabortty and P. P. Khargonekar, "Introduction to wide-area control of power systems," in *2013 American control conference*. IEEE, 2013, pp. 6758–6770.

[112] M. E. Csete and J. C. Doyle, “Reverse engineering of biological complexity,” *science*, vol. 295, no. 5560, pp. 1664–1669, 2002.

[113] S. Friedenthal, A. Moore, and R. Steiner, *A practical guide to SysML: the systems modeling languag*e. Morgan Kaufmann, 2014.

[114] T. Samad, P. McLaughlin, and J. Lu, “System architecture for process automation: Review and trends,” *Journal of Process Control*, vol. 17, no. 3, pp. 191–201, 2007.

[115] T. Samad and A. M. Annaswamy, “Controls for smart grids: Architectures and applications,” *Proceedings of the IEEE*, vol. 105, no. 11, pp. 2244–2261, 2017.

[116] J. L. Hennessy and D. A. Patterson, *Computer architecture: a quantitative approach*. Elsevier, 2011.

[117] F. Kendoul, “Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems,” *Journal of Field Robotics*, vol. 29, no. 2, pp. 315–378, 2012.

[118] J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems,” *Manufacturing letters*, vol. 3, pp. 18–23, 2015

CHAPTER 5

Technology Validation & Translation

Technology Validation & Translation

Stefano Di Cairano, Johan Eker, Sonia Martinez, Thomas Parisini, Tariq Samad

The industrial ecosystem needs control systems. In this section, we propose measures and recommendations for next-generation research validation that can bridge the gap between academia and industry and enable the development of control technologies and products with high societal impact.

Abstract Over the next decade, the control systems community (including academics, young professionals, practitioners, and entrepreneurs) has the opportunity to transform scientific results into beneficial technologies and products for society at large. In this section, we discuss the path from academic research toward industrial uptake and product development. We provide data about the importance of systems and control to industry, identify key ways that the academic community has historically failed to engage with the industrial ecosystem, and discuss potential solutions. Particular attention is given to the validation of research outcomes at a range of levels, from lab-controlled to real-world environments. We believe that the lack of validation is a major obstacle for technology transfer, and that the academic community can play a crucial role in closing this gap. We further suggest a pathway for better engagement with collaborative innovation ecosystems. This can reveal new avenues of research in academia, and ultimately enable a paradigmatic shift in the societal impact of control systems.

5.A Introduction

Systems and control is a horizontal scientific discipline that does not directly translate into a specific technology. Thus, it has suffered from the theory/applications dichotomy for decades. To quote, Steven Low's column in *Control Systems Magazine* [1], "The gap between theory and practice is a lot bigger in practice than in theory." What is the way forward for systems and control to shift into an overall go-to-market strategy? Will entrepreneurial communities play a game-changing role?

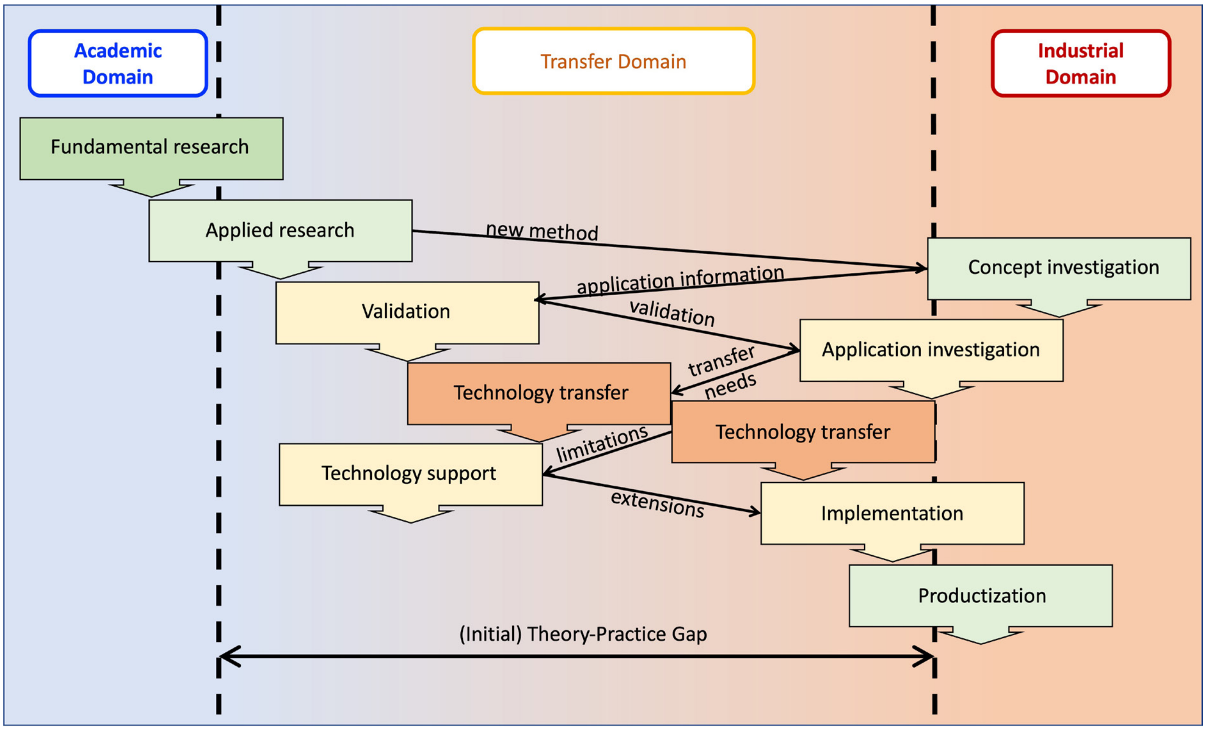


Figure 5.1: The road from academic research to industrial uptake is long and winding. However, by clearly identifying the steps involved and the current barriers, we believe the uptake of research results may be significantly improved.

Over the next decade, the control systems community (including academics, young professionals, practitioners, and entrepreneurs) has the opportunity to transform scientific results into beneficial technologies and products for society at large. Figure 5.1 shows a schematic view of the path from academic research toward industrial uptake and product development. We provide data about the importance of systems and control to industry, identify key ways that the academic community has historically failed to engage with the industrial ecosystem, and discuss potential solutions. Particular attention is given to the validation of research outcomes at a range of levels, from lab-controlled to real-world environments. We believe that the lack of validation is a major obstacle for technology transfer, and that the academic community can play a crucial role in closing this gap. We further suggest a pathway for better engagement with collaborative innovation ecosystems. This can reveal new avenues of research in academia, and ultimately enable a paradigmatic shift in the societal impact of systems and control.

5.B Engagement in Industrial Ecosystems

Control systems play a central role across a remarkable breadth of application domains and industry sectors: aerospace, automotive, biomedical, building, and many others. The vast and ever-expanding array of control-enabled products, services, systems, and solutions is a testament to the societal importance of the discipline.

Despite this fact—or perhaps because of it, since meeting such diverse needs requires extensive abstraction of fundamental theories and methodologies—industry's level of engagement with the control research community is currently poor. This is a disadvantage for both sides: it takes longer for industry to benefit from new research results, and research efforts are poorly informed by real-world requirements. A paradigmatic shift in the industry's engagement is necessary and timely. We believe that enhancing the control research community's understanding of the industry and practical applications is essential for bridging the gap. In [2], 10 messages for the control research community are enumerated to this end (see Table 5.1). Here, we further abstract these messages, identifying five crucial "dimensions" of understanding.

Table 5.1: Ten messages for the control research community (reprinted from [2], copyright 2020, with permission from Elsevier).

1. Advanced control technologies vary significantly in their impact and perceptions thereof
2. The control research community is broadly unaware of the impact of advanced control
3. Real-world success requires domain understanding
4. Control technology implementation infrastructures and architectures are industry-specific
5. Advanced control is more than feedback control...It is a systems-oriented, rigorous mindset
6. Control science has broad-based relevance for new and emerging technologies
7. Corporate r&d can (sometimes) serve as a bridge for technology transfer of academic research
8. Cost reduction is a high priority for industrial innovation in control
9. Economic expectations influence industry investment in research
10. The industry-academy disconnect extends to education

In elaborating on the dimensions for understanding industry and applications, it is worth highlighting different industry sectors, types of products and services, companies, and even different groups or businesses within companies. In each case, specific and often idiosyncratic contexts, requirements, and considerations can apply. We can define five dimensions of knowledge and understanding, all of which should be traversed by researchers seeking industry impact in their work:

- **Scientific**: Here we refer to physics, chemistry, biology, or other scientific fields that underpin the problem to be addressed by automation and control. In developing a control algorithm for a nonlinear petrochemical process, for example, an understanding of the reaction chemistry is important. Similarly, for flight control, aerodynamics must be understood. The field involved does not have to be a natural science. Computer and communications science lies behind software and computational applications and are relevant to today's CPS.

The dynamics of the phenomena involved are presented in a language familiar to control systems researchers—which is, of course, mathematics. Thus, this aspect of industry application is one that is well within the comfort zone of these researchers. Indeed, the control research community may mistakenly believe that, once they have grasped the scientific basis of the application domain, their understanding is largely complete.

- **Engineering**: In this context, we mean the objectives, requirements, and constraints associated with a given application. Examples include safety and reliability, performance as measured by speed or precision of control, efficiency or control effort, cybersecurity, and cost reduction. The last is worth emphasizing: a recent survey of industry sectors, reported in [2], found that cost reduction was one of the top three requirements for all but one of the sectors (the odd one out is medical technology). Yet the focus of most control research projects is performance and robustness, and cost reduction is often not considered.
- **Technological**: The engineering of an automation system involves technological platforms and components that are usually specific to an industry sector or application domain. An partial list of these components and platforms would include:
  - Sensors and actuators, whose limitations and capabilities must be understood
  - Computational engines, which can be embedded microprocessors, centralized computational facilities, and, most recently, cloud-hosted applications
  - Communication networks, whose bandwidth and latency will affect control designs
  - Real-time operating systems
  - User interfaces

  The technological elements of industry-specific automation are often standardized. The standards are often, but not always, open and industrywide. Suppliers of control systems and components, or large companies themselves, may require adherence to their prescribed communication protocols or interfaces.

  Familiarity with technological standards and practices will facilitate the translation of research results into functional products.

- **Economic**: The overarching objective of any public corporation is to make money. Decisions at high levels of companies are made with this objective foremost in mind. Resources, whether finances for investment or people for development, are always limited, and executives must consider risk/reward trade-offs in their decision-making. Important considerations include: How much will a new development cost to commercialize, and what return on investment can be expected? How do those numbers compare with other opportunities, whether in the same general area or not? Is the targeted application sector growing, and if so, what is the compound annual growth rate (CAGR)? What is the competitive landscape and does it support the investment required? See Table 5.2 for an itemization of recent data related to market size and growth rates for several industry sectors and cross-industry technology areas.
- **Sociological**: We use this term somewhat loosely to refer to the human element of an industry sector. For example, what type of operators or engineers will be available to install, commission, and maintain an advanced control solution? The education level of operational staff and the availability of Ph.D.-level resources in the organization can make the difference between a research development being

incorporated in commercial practice and being relegated to the shelf. Whether a workforce is unionized can affect training needs and the acceptance of solutions that could result in jobs being lost in one area and gained in another.

In summary, bridging the gap between research results and their usability in industry is a key step in achieving structural industry engagement. In the rest of this chapter, we will focus on the technological dimension and the research–validation–transition workflow as a future key enabling factor.

Table 5.2: Sales volumes and compound annual growth rates (CAGR) for selected control-related industry sectors and cross-industry technologies. Sources listed are market reports and press releases thereof and were accessed Dec. 27, 2019. (Reprinted from [2], copyright 2020, with permission from Elsevier).

Control industry sectors in $B (year)	**Sales volume**	**Growth (CAGR)**	**Period for CAGR**	**Source**
Industrial Control	117 (2017)	5.3%	2018-2025	https://tinyurl.com/yxd2gya7
Automotive Control	63.6 (2017)	4.4%	2019-2025	https://tinyurl.com/uy78wfp
Aircraft Flight Control	11.1 (2017)	3.52%	2018-2023	https://tinyurl.com/ydf7uqo8
Industrial Robotics	16.5 (2017)	12.0%	2020-2022	https://tinyurl.com/wkd3d23
Smart Home Automation	75 (2018)	11.8%	2019-2025	https://tinyurl.com/vz2ozk5
Cross-Industry Technology Areas				
Cybersecurity	119 (2018)	14.5%	2019-2024	https://tinyurl.com/ve82q32
Digital Transformation	262 (2018)	18.2%	2019-2026	https://tinyurl.com/yynu6plb
Internet of Things	190 (2018)	24.7%	2019-2026	https://tinyurl.com/y2czseqh

5.C Validation

The path that takes fundamental research to real-world problem-solving is long and complex. New control methods may be analytically demonstrated to provide significant benefits with respect to standard methods, but in order to gain the trust of final implementer (whether a company or a government regulator), extensive validation work must be completed. Experimental validation in the final preproduction setup is too time-consuming and costly. Therefore, we usually take a sequence of intermediate validation steps of increasing complexity, from simulation models and hardware-in-the-loop digital twins based on first principles and/or data to full-fledged experimental setups.

While there is a clear economic benefit to validating theoretical results in environments that are closer to a final target application, the validation process can raise issues that are not yet ideally addressed by the original control design, thus spurring new research. In addition to the generalization of the theory, the context provided by these applications is invaluable. It can help provide further understanding of why things work the way they do and enable extraction of greater value from the theory. Closing the gap between theory and its impact on the real world is also crucial to attracting new researchers and students to the field and to engaging the general public. This is done very effectively in other communities. One example is the ImageNet dataset [3] which has been instrumental in advancing research in Computer Vision and Machine Learning by providing a common benchmark for comparing algorithms and methodologies. It has also been used to benchmark computation aspects like hardware acceleration. OpenAI Gym [4] provides a similar environment for evaluating reinforcement learning algorithms. Kaggle [5] is a crowdsourced platform for Machine Learning research. It started out hosting competitions (which continue this day), but it has also become a public data platform offering a wide collection of datasets and tutorials. Anybody can set up a Kaggle competition and invite the community to participate. This kind of gamification of research is largely lacking in the control community, although some competitions are starting to appear at conferences, such as at the 2020 IFAC World Congress. Control publications are also behind in engaging with the *reproducible research* initiatives by IEEE [6].

In the next three sections, we discuss the main components, challenges, and potential benefits of validation infrastructure in enabling new control developments. We aim to lay out a road map to validation infrastructure for deployable large-scale, real-world control systems. Key elements include suggesting a sequence of validation steps, advocating for open-source validation resources, and calling for the creation of standardized tests and benchmarks that can freely be used by the community to increase the impact of control system technologies. By validation infrastructure, we refer to the entire set of tools, including experimental hardware, simulation software, datasets, and benchmarks which can help demonstrate the benefits of a control method. A validation process encompasses the application of the control methodology on benchmark problems, simulations of different complexities, laboratory experimental testbeds, and scaled (dimension and extension) experimental prototypes. We start by presenting a brief, account of various efforts that focus on one or more of these stages of validation, followed by a more detailed description of these steps.

5.D Current Status of Benchmark and Testbeds

There are multiple levels of validation of a control method or technology, going from detailed simulations to hardware-in-the-loop experiments and digital twins, each with different levels of precision.

Control research has often leveraged standard abstract problems such as the inverted pendulum, (possibly) interconnected mass-spring-damper systems, and (multiple) water tank(s). All of these paradigmatic examples can easily be expressed as control systems, and thus are readily available for control method evaluation.

In assessing a control method in a specific target application, the first step before moving to actual experiments is usually the verification of the control method on detailed simulation models. In many cases, these models and tools have been built and shared with the control and research communities at large. Some examples include benchmarks for fault-tolerant control of wind turbines and wind farms [7, 8], pulp mills for process control [9], the ADMIRE aircraft benchmark [10] and the A320 model in the SMAC toolbox [11], and the Toyota Prius model included in PSAT [12] (now Autonomie). There are numerous simulation tools for power grid simulation and control. Of special note are those produced by energy national laboratories and the EPRI (Electric Power Research Institute), which include OpenDSS [13] and GridLAB-D [14] for simulation of power distribution networks. In this case, simulation models are usually not in a form that can be used directly for control purposes, so the extraction of a control-oriented model that is suitable for the applied method requires time and effort. However, these benchmarks provide valuable proof that the control methods have the potential to work in real-world applications.

Datasets are also effective for such types of validation. In many branches of signal processing, validation datasets are quite common; one of the oldest open datasets is the UCI Machine Learning Repository [15]. For control, the use of prerecorded datasets is more challenging, because if plant behavior is changed by the control method, the data needs to change as well. Still, datasets can be used as real disturbances, for validating estimation and fault detection methods, and to determine parameters for simulation models. Some key examples of relevant datasets (primarily from power domain) are NREL's repository [16] on buildings, grid, solar, wind and transportation data; DataHUB at PNNL [17]; the Open Data Initiative at LLNL and CityBES [18, 19, 20]; and EPRI's EPRI10 [21]. Additional examples are the PeMS Performance Measurement System (PeMS) data source [22] with data from California's metropolitan area freeways; the Building Performance Database (BPD) [23] from DoE; and NuScenes for perception and autonomous driving [24]. In recent years, datasets have become increasingly available through public, private, and government efforts. At the time of writing, the aforementioned Kaggle [5] returned 2,880 datasets on COVID-19, six on robotics navigation, and 22 on fault detection. GitHub [25] and Google's dataset search engine [26, 27] are other good sources of datasets. Commercial publishers are also offering repositories for datasets, such as IEEE DataPort, the Open Data tool from Elsevier, reference [28] which provides links to datasets, and software through ArXiv.

Moving into the physical world, the first validation steps involve implementation on laboratory testbeds, which on a small scale can replicate real application systems. Examples include [29, 30, 31, 32, 33, 34, 35, 36, 37, 38], which span automotive, aerospace, robotics, and air conditioning applications. In power systems and energy research, there are several tools that include hardware-in-the-loop as part of their testing. Some examples are the commercial tools RTDS [39], Typhoon-HIL [40], and the Grid Simulation and Power Hardware-in-the-Loop from NREL [41]. Proprietary digital twins in the process and metal industry are also becoming common, even though they use commercial platforms and are not publicly available (see, for example, the gPROMS platform [42]).

In recent years, there have been successful efforts in providing shared and remotely-accessible testbeds that can serve as standard platforms for experimental evaluation. One example is the SPHERES [43] testbed from MIT, which allows researchers to remotely deploy control algorithms in the International Space Station (ISS) through a guest scientist program (GSP). Besides a flight testbed that is operated by astronauts, the SPHERES/GSP testbed consists of a simulator for researchers to run locally. A more recent example is the Robotarium [44] at the Georgia Institute of Technology, which enables deployment of algorithms for multirobot control from anywhere in the world. Users can test and develop algorithms locally using the Robotarium simulation API and then deploy them on the actual plant by uploading the code. The code is stored indefinitely, allowing for experiments to be repeated and verified later with identical configuration. A slightly different approach is proposed by Duckietown [45], an open-source platform for experiments with self-driving vehicles. This is not a remote testbed but rather a blueprint for developing standardized local deployment. DERConnect is an NSF-sponsored national infrastructure for experimenting with real and simulated distributed energy resources and distributed control at a large scale [46]. This national resource, which is currently under construction, will be remotely available for researchers and practitioners in academia and industry alike.

Before moving to actual "preproduction-like" testing, the final and most challenging experimental validations leverage real-size testbeds in environments that are open yet limited and possibly controlled. In these, the control method is deployed on a prototype that closely resembles a potential final product and needs to contend with similar conditions to those encountered in real life. But its operation is more limited, because some of the externally imposed environs are shaped to excite specific behaviors but are also likely less random and rich than those in real life. Some examples (drawn specifically from automated driving, smart cities, and robotics) are Toyota's Woven City [47] for testing tomorrow's smart city technologies, Mcity [48] for automated driving, competitions such as the DARPA Grand Challenge [49], the Urban Challenge [50] for automated driving, the Robotics Challenges [51], and the Subterranean Challenge [52].

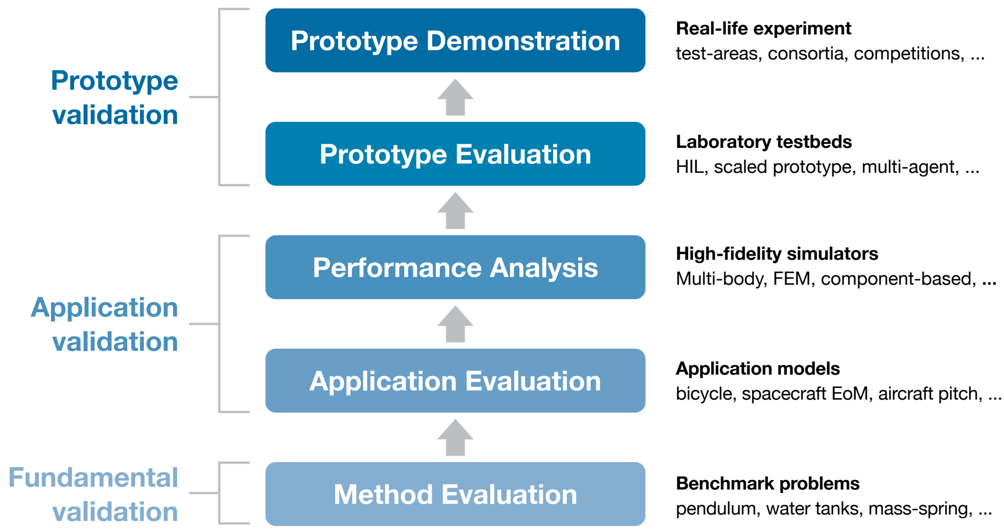


Figure 5.2: Validation steps and tools for each step.

5.E Validation Steps and Corresponding Tools

The validation of a new control method and its implementation on an actual plant proceeds through multiple steps. These steps, from fundamental research to its real-world application, have different definitions according to the process being applied. One particularly famous example is the NASA Technology Readiness Levels (TRLs) [53]. TRLs cover the evolution of new technology from basic principles investigation (TRL-1) to deployment in the final system (TRL-9) for NASA spacecraft. The higher TRLs are naturally adapted to a specific development organization (usually a company or a government entity) while most of the open-access research in academia, government, and industry (and collaboration among those) tends to cover TRL-1 through TRL-5 or TRL-6. Each TRL requires a specific level of validation, in relation to the maturity of the prototype, to the type of tests executed and to the environment in which those tests are developed.

Inspired by TRLs, but with a slightly modified structure due to our focus on the validation of new technologies in the open-access research domain, we chart below some of the major levels of validation for control. They proceed from fundamental research to final-deployment organization, i.e., when they have provided enough evidence to indicate they can work in practice. See Figure 5.2 for a schematic representation of these levels.

1. **Method Evaluation**: This level typically consists of simulation studies on relatively simple and standardized benchmarks. These studies are aimed at confirming theoretical findings in a “clean,” unperturbed environment, assessing the limits of the method when departing from theoretical assumptions, and providing data for comparing the new technology with existing ones. Here, it is key to have a proper set of widely accepted benchmarks or to identify challenging benchmarks that can highlight the relevant and novel properties of the technologies in question. Naturally, as industry standards progress, so would this collection of benchmarks. Such a collection should be thoroughly explained in a handbook and maintained in available open-source code for some (or most) of the standard development platform (e.g., Python, Julia, MATLAB). Examples: standard control benchmarks, such as the inverted pendulum, water tank systems, mass-spring-damper systems, or a DC-motor with flexible shaft.

2. **Application Evaluation**: The previous validation step should provide a range of hypotheses on how the new technology will handle specific issues of a given application. It should also address particular limitations of the current state-of-the-art methods used for that application (relative to performance, cost, robustness, etc.). This second validation step usually consists of simulation studies using simple models of the plant. These models are similar but not equal to those used for control design; they may have low/medium fidelity for quick assessment of the technology. Here, the key is to use simulation models that are simple enough to provide clear data and easy and fast to run, while still capturing the most challenging characteristics of the final application. Examples: classical models for application such as vehicle dynamics/kinematics [54, 55, 56], spacecraft equations of motion [57, 58], robot operation [59, 60], circuits/grids using IEEE standards, biological systems [61].
3. **Performance Analysis**: This refers to a realistic assessment of the potential impact of the new technology, before its actual implementation, on a more realistic model of the actual plan. This third step usually consists of simulations employing a high-fidelity model of the plant, which are far from the control-oriented model and which can include finite-elements methods, multi-body dynamics, acausal modeling languages, and program-oriented simulation models. These simulations are typically much slower to execute. They can also include limited hardware-in-the-loop. It is vital to have access to realistic simulation models. They are often provided by third parties to avoid biasing the results. They should also be simple to connect and use, so as to enable fast development and analysis of the results. Reliability of the models can be further assessed by "tuning," using real or synthetic data from the target application (or a similar one). Examples: high-fidelity models developed in simulation-oriented software packages, either for general physics, such as Modelica, Simscape, OpenFOAM, COMSOL [62, 63, 64, 65], or for specific applications, such as CarSim, CarMaker for automotive [66, 67] and Ansys STK for spacecraft [68], as well as real-life datasets, when available for the application and appropriate for the developed function, possibly also used to calibrate the simulation models.
4. **Prototype Evaluation**: This consists of an evaluation of the technology on a scaled system in a controlled laboratory environment, subject to disturbances and imperfections that are typical of the actual plant. This validation aims to verify the operation of the new technology with a physical plant. In a laboratory setting, its robustness is verified to typical disturbances and imperfection and its execution is verified in real time and with real communication constraints and computational and communication delays. Although it may be scaled down, a testbed must be affected by disturbances and modeling errors that capture the principal corresponding ones in the actual application. It should also be simple to manage, modify, and extend. Examples: Hardware-in-the-loop setups [39, 40, 41, 46], laboratory testbeds [29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 45], and remotely accessible testbeds [43, 44, 46].
5. **Prototype Demonstration**: This refers to the final evaluation of the full plan in a limited, monitored open environment. This validation assesses the operation of the technology in real conditions with the full-size plant or a minimally scaled plant. The key is to have access to an open environment where the technology can be safely tested while being constantly monitored by a set of resources that are not available in the final applications and are not used by the technology directly. Examples: Prototypes of energy efficient/smart buildings [69] and smart cities [47], realistic test areas for autonomous vehicles [48, 70], and competitions such as [49, 50, 51, 52].

5.F Desired Features of Validation Infrastructure

A good overview of traditional control system testbeds is given in [71], which outlines a set of important criteria for a successful industrial control testbed, based on earlier work from the Emulab testbed [72]:

- **Fidelity** – The testbed should represent the physical system as accurately as possible.
- **Repeatability** – Repeated experiments should yield the same or statistically consistent results.
- **Measurement Accuracy** – Monitoring should be accurate and not interfere with the experimental results.
- **Safe Execution** – Research activities should not be able to cause any devastating effect on the physical system and personal safety.

Control systems research is constantly uncovering new target applications—from automated vehicles to smart cities, from synthetic biology to social dynamics, and from autonomous spacecraft to swarms of UAVs. Thus, the validation infrastructure for proving the impact of new control methods must evolve as well. Based on the known and predicted target applications of control systems in the near to mid-term future, we outline some additional features, beyond those in [72], that validation infrastructure should have in order to provide effective and convincing results:

1. **Multi-Domain**: As systems become more complex—involving elements from multiple electromechanical, biological, information, and human/social domains—their validation infrastructure should consistently support all of these domains, or at least several of them. Thus, it is important to investigate these domain interactions as well as the performance of the control method on each domain. Examples with multi-domain characteristics include bio-inspired robots, hybrid/electric vehicles, HVAC/building control, crowd motion control, opinion dynamics, deep-space spacecrafts, and epidemic dynamics.
2. **Complexity-Scalable**: Another dimension of system complexity is the "parts" the systems are composed of and their coupling (in, e.g., multi-agent and possibly heterogeneous systems, multicomponent systems, layered control architectures, and interconnected systems.) The validation infrastructure should be able to scale and validate the interactions between these various components. Examples include: disaster relief robot teams, connected automated vehicles, autonomous driving software architecture, power grids, smart cities and infrastructures, and, more generally, large CPS.
3. **Enabling Integration of First Principles and Data**: All applications are based on fundamental physical principles, but no physical principle alone perfectly models observed reality. Validation infrastructure must therefore enable first-principle modeling, data-based modeling, and, especially, the integration of physics-based modeling with data-based modeling (i.e., so-called data-augmented model-based design). Examples: powertrains (conventional and hybrid), vapor compression cycles, robot and automated vehicle motion, crowd modeling, and social information dynamics.
4. **Digitization and Networked Architecture**s: At the core of control application expansion is the prevalence of computational resources interconnected by communication networks: microprocessors, embedded systems, GPUs, edge-computers, and cloud resources. While widespread, these resources are still limited, so the validation infrastructure needs to be able to assess the impact of software architecture, connectivity requirements, computational load, and communication bandwidth on control methods—and vice versa. Examples include: automatic manufacturing, connected automated vehicles, autonomous logistics, multi-spacecraft/UAV systems.

5. **Movable Boundary**: Each system is part of more complex system and/or environment, and different validation phases may draw boundary lines for modelled interactions in different places. The validation infrastructure must enable moving such boundary lines to model less (in the initial phases) or more (in later phases) interaction of the system components and of the system with the environment. Examples include: electric vehicles, power grids, autonomous vehicles, vapor-compression systems, human social interaction systems.
6. **Rapidly Prototypeable**: Unlike the final application, validation infrastructure needs to be easily (re) programmable to allow rapid testing of and rapid changes to the method, without excessive time spent implementing the controller. The architecture should be open as much as possible and enable the use of open-source software, possibly coded using different platforms. Examples include: automated code generation, cross-compiling, virtualization, ROS integration, modular architecture, and digital twins.

5.G Translation of Research Outcomes to Innovation and Products

The previous sections highlighted the key elements enabling the translation of R&D outcomes into innovation and, eventually, products. Significantly increasing the active presence of control technology startups and spinoff companies in innovation districts[1] is a potentially disruptive opportunity. There are already significant use cases that support this outlook. Around the globe, innovation districts are bringing together researchers, early stage entrepreneurs, and corporate leaders to work and live in open ecosystems.

Innovation districts have a strategy and business model that have changed dramatically over the last few years. Companies have moved away from closed innovation models to more open approaches in which organizations work in collaborative networks and form novel partnerships between previously unrelated industries.

Innovation districts differentiate themselves from traditional science or business parks. Unlike incubators, they include a wide range of businesses across the ecosystem, from startups to large corporations. They lend themselves well to complex and multidisciplinary activities in areas of convergence between different sectors. As a key enabling methodologies and technologies, systems and control are ideally placed to grow in Innovation district-like ecosystems with effective pipeline mechanisms sourcing talent and research outputs from one or more leading universities. Major institutions can generate significant market value, particularly those that are “research-intensive.”

Innovation districts bring together large and small companies, universities, and research centers in order to foster leading-edge advancements in key industrial sectors. These ecosystems follow the Open Innovation paradigm, which holds that synergies between different fields and firms can accelerate development and bring powerful new solutions to market. All players share and collaborate throughout the entire journey, from initial brainstorming to product launch. This sharing includes intellectual property rights (IPR), and only at higher TRLs do the different companies customize their activities to their product. This is an ideal ecosystem for horizontal technologies like control engineering.

[1] https://en.wikipedia.org/wiki/Innovation_district

5.H Concluding Remarks

In this chapter, we focused on how to narrow the gap between academic research and industrial needs. We identified obstacles in various different dimensions, from technical to sociological, which have historically prevented the translation of research results into impactful products. We believe validation will be a key component in technology transfer and we have proposed a structured method to introduce validation at different levels. This paradigmatic shift will require new common testbeds and use cases, as well as a new spirit of free collaboration between scholars and practitioners.

Recommendations

In this section, we list a set of recommendations for advancing the area of validation and facilitating translation to innovation and technology transfer.

Short Term

- Stress the important contribution of validation in the application of fundamental research. Leverage journals, e.g., IEEE Transaction on Control Systems Technology and conferences e.g., IEEE Conference on Control Technology and Applications should highlight the validation of control research (especially steps 3–5). Promote special sessions and special issues focusing on validation and make validation visible in calls for papers.
- Refocus CSS publications, including Control Systems Magazine—CSM, to regularly feature content of interest to industry and articles from which researchers could learn about industry perspectives. Alternatively, a new industry- and application-oriented publication could be founded.
- Support initiatives for open-source code distributions associated with novel control research. Code distribution will promote reproducibility of results, help to disseminate knowledge, speed up progress, and strengthen the trust of final implementers of control research.
- Offer industry- and innovation-relevant training to interested control students, faculty, and young professionals. This could be accomplished through workshops, short courses, webinars, and more. The offerings could cover business foundation and innovation management topics, as well as overviews of industry applications.
- Launch an “Industry Experts” program at research-oriented control conferences (e.g., IEEE CSS and IFAC) to promote industry participation. This program could give select industry leaders a professional distinction and support travel to our conferences for presentations and curated discussions.

Medium Term

- Extend the set of standard benchmarks and make digital twins of challenging industrial applications openly available to control researchers. Beyond the inverted pendulum, double integrator, water-tank control, continuously stirred tank reactor, simplistic quadrotors, etc., new benchmarks should enable the comparison and analysis of new control methods. These benchmarks should be real-world motivated, simple enough to allow clear early assessment of the methods, and general enough to address multiple aspects of future control research.
- Promote and support initiatives that enable the acquisition, refinement, validation, and open distribution of data for configuring and calibrating realistic simulation environments. Incentivizing making this information available to the entire community will empower further comparative analysis of new control methods on a standard test base.
- Support and encourage the development of shared, realistic, small-scale prototypes to evaluate control in safe and controlled environments like laboratories and small testing facility. Enabling remote accessibility will allow the entire community to benefit from a standardized method for testing the outputs of control research.

Long Term

- Aggregate, develop, and promote high-fidelity validation platforms—especially simulators—focusing on realistic open-source platforms that provide a high degree of property preservation when translating methods to real-world applications.
- Support and incentivize partnership with companies and governmental entities to establish larger, more realistic testbeds at full- and quasi-full scale. Such controlled and monitored environments will provide the strongest level of validation before transfer of the technologies to final implementers.
- Foster the growth of a research culture that facilitates translation to high-impact products and bridges—once and for all—the theory/practice gap.

[1] S. Low, “The theory–practice–entrepreneurship gap,” *IEEE CONTROL SYSTEMS Magazine*, 2021.

[2] T. Samad et al., “Industry engagement with control research: Perspective and messages,” *Annual Reviews in Control*, vol. 49, pp. 1–14, 2020.

[3] J. Deng *et al*., “ImageNet: A large-scale hierarchical image database.”

[4] G. Brockman et al., \OpenAI gym," arXiv preprint arXiv:1606.01540, 2016.

[5] Kaggle, https://www.kaggle.com.

[6] IEEE, “Research Reproducibility,” https://journals.ieeeauthorcenter.ieee.org/create-your-ieee-journal-article/research-reproducibility/.

[7] P. F. Odgaard, J. Stoustrup, and M. Kinnaert, “Fault-tolerant control of wind turbines: A benchmark model,” *IEEE Transactions on control systems Technology*, vol. 21, no. 4, pp. 1168–1182, 2013.

[8] P. F. Odgaard and J. Stoustrup, “A benchmark evaluation of fault tolerant wind turbine control concepts,” *IEEE Transactions on Control Systems Technology*, vol. 23, no. 3, pp. 1221–1228, 2014.

[9] J. Castro and F. Doyle III, “A pulp mill benchmark problem for control: problem description,” *Journal of Process Control*, vol. 14, no. 1, pp. 17–29, 2004.

[10] M. Hagström, “The ADMIRE benchmark aircraft model,” in *Nonlinear analysis and synthesis techniques for aircraft control*. Springer, 2007, pp. 35–54.

[11] J. Biannic and J. Boada-Bauxell, “A civilian aircraft landing challenge,” *On-line available from the aerospace benchmark section of the SMAC Toolbox*, 2016. http://w3.onera.fr/smac

[12] A. Rousseau, J. Kwon, P. Sharer, S. Pagerit, and M. Duoba, “Integrating data, performing quality assurance, and validating the vehicle model for the 2004 Prius using PSAT,” in *SAE 2006 World Congress & Exhibition*, no. 2006-01-0667, 2005.

[13] “OpenDSS,” https://www.epri.com/pages/sa/opendss.

[14] “GridLAB-D,” https://www.gridlabd.org/.

[15] “The UCI Machine Learning Repository,” https://archive.ics.uci.edu/ml/index.php.

[16] “NREL Data and Tools website,” https://nrel.gov/research/data-tools.html.

[17] “PNNL Data Hub,” https://data.pnnl.gov.

[18] “LLNL Open Data initiative,” https://data-science.llnl.gov/open-data-initiative.

[19] “Urban science group at lawrence berkeley national lab,” https://buildings.lbl.gov/urban-science/tools.

[20] Y. Chen, T. Hong, X. Luo, and B. Hooper, “Development of city buildings dataset for urban building energy modeling,” *Energy and Buildings*, vol. 183, 2019.

[21] “The EPRI10 datasets,” https://www.epri.com/thought-leadership/artificial-intelligence/data-sets.

[22] “PeMS Data Source. Caltrans traffic operations,” https://dot.ca.gov/programs/traffic-operations/mpr/pems-source.

[23] “Buildings performance database,” Buildings Technology Office. US Department of Energy, 2014, https://www.energy.gov/eere/buildings/building-performance-database-bpd.

[24] H. Caesar et al., “Nuscenes: A multimodal dataset for autonomous driving,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11621–11631.

[25] “Github,” https://github.com/datasets.

[26] “Google dataset engine,” https://datasetsearch.research.google.com/.

[27] “Discovering millions of datasets on the web,” https://www.blog.google/products/search/discovering-millions-datasets-web/.

[28] “Papers with Code. Medium publication.” https://paperswithcode.com/datasets.

[29] A. Bicchi, A. Balluchi, D. Prattichizzo, and A. Gorelli, “Introducing the SPHERICLE: an experimental testbed for research and teaching in nonholonomy,” in *Proc. Int. Conf. on Robotics and Automation*, 1997.

[30] S. Brennan and A. Alleyne, “Using a scale testbed: Controller design and evaluation,” *IEEE Control Systems Magazine*, vol. 21, no. 3, pp. 15–26, 2001.

[31] L. Cremean et al., “The Caltech multi-vehicle wireless testbed,” in *Proc. 41st IEEE Conf. Decision and Control*, 2002.

[32] S. Cho, J. Shen, and N. H. McClamroch, “Mathematical models for the triaxial attitude control testbed,” *Mathematical and Computer Modelling of Dynamical Systems*, vol. 9, no. 2, pp. 165–192, 2003.

[33] C. Chevallereau et al., “Rabbit: A testbed for advanced control theory,” *IEEE Control Systems Magazine*, vol. 23, no. 5, pp. 57–79, 2003.

[34] M. W. Regehr et al., “The formation control testbed,” in *Proc. IEEE Aerospace Conf.*, vol. 1, 2004, pp. 557–564.

[35] G. Hoffmann et al., “The Stanford testbed of autonomous rotorcraft for multi-agent control (STARMAC),” in *Proc. 23rd IEEE Digital Avionics Systems Conference*, 2004.

[36] J. P. How, B. Behihke, A. Frank, D. Dale, and J. Vian, “Real-time indoor autonomous vehicle test environment,” *IEEE Control Systems Magazine*, vol. 28, no. 2, pp. 51–64, 2008.

[37] H. Qiao, C. R. Laughman, D. Burns, and S. Bortoff, “Dynamic characteristics of an R410a multi-split variable refrigerant flow air conditioning system,” in *12th IEA Heat Pump Conference*, 2017.

[38] K. Berntorp, T. Hoang, R. Quirynen, and S. Di Cairano, “Control architecture design for autonomous vehicles,” in *IEEE Conf. Control Technology and Applications*, 2018, pp. 404–411.

[39] "RTDS Simulator Technologies," https://www.rtds.com/.

[40] "Typhoon HIL," https://www.typhoon-hil.com/.

[41] "Grid Simulation and Power Hardware-in-the-Loop," https://www.nrel.gov/grid/simulation-phil.html.

[42] Siemens PSE, "The gPROMS platform," https://www.psenterprise.com/products/gproms/platform.

[43] J. Enright, M. Hilstad, A. Saenz-Otero, and D. Miller, "The SPHERES Guest Scientist Program: collaborative science on the ISS," in *2004 IEEE Aerospace Conference Proceedings (IEEE Cat. No.04TH8720)*, vol. 1, 2004, p. 46.

[44] S. Wilson *et al*., "The robotarium: Globally impactful opportunities, challenges, and lessons learned in remote-access, distributed control of multirobot systems," *IEEE Control Systems Magazine*, vol. 40, no. 1, pp. 26–44, 2020.

[45] L. Paull *et al*., "Duckietown: An open, inexpensive and flexible platform for autonomy education and research," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 1497–1504.

[46] "UCSD DERCconnect," https://sites.google.com/ucsd.edu/derconnect/home.

[47] J. Savoie, "Toyota's Woven City," *The E&S Magazine*, vol. 65, no. 1, p. 13, 2022.

[48] U. Briefs, "Mcity grand opening," *Research Review*, vol. 46, no. 3, 2015.

[49] M. Buehler, K. Iagnemma, and S. Singh, *The 2005 DARPA grand challenge: the great robot race.* Springer, 2007, vol. 36.

[50] M. Buehler, K. Iagnemma, and S. Singh, The DARPA urban challenge: autonomous vehicles in city traffic. Springer, 2009, vol. 56.

[51] E. Krotkov *et al.,* "The DARPA robotics challenge finals: Results and perspectives," in *The DARPA Robotics Challenge Finals: Humanoid Robots To The Rescue*. Springer, 2018, pp. 1–26.

[52] M. Tranzatto *et al*., "CERBERUS in the DARPA subterranean challenge," *Science Robotics*, vol. 7, no. 66, 2022.

[53] J. C. Mankins *et al*., "Technology readiness levels," *White Paper, NASA, Tech. Rep*., 1995.

[54] R. Rajamani, *Vehicle dynamics and control.* Springer Science & Business Media, 2011.

[55] R. N. Jazar, *Vehicle dynamics.* Springer, 2008, vol. 1.

[56] J. B. Heywood, *Internal combustion engine fundamentals*. McGraw-Hill Education, 2018.

[57] M. J. Sidi, *Spacecraft dynamics and control: a practical engineering approach.* Cambridge university press, 1997, vol. 7.

[58] W. Fehse, *Automated rendezvous and docking of spacecraft*. Cambridge university press, 2003, vol. 16.

[59] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[60] S. M. LaValle, *Planning algorithms*. Cambridge university press, 2006.

[61] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[62] M. Tiller, *Introduction to physical modeling with Modelica*. Springer Science & Business Media, 2001.

[63] "Simscape - MathWorks," https://www.mathworks.com/products/simscape.html (accessed July 12, 2022).

[64] H. Jasak, A. Jemcov, Z. Tukovic *et al*., "OpenFOAM: A C++ library for complex physics simulations," in *International workshop on coupled methods in numerical dynamics*, vol. 1000. IUC Dubrovnik Croatia, 2007, pp. 1–20.

[65] "COMSOL," https://www.comsol.com (accessed July 12, 2022).

[66] "CarSim - Mechanical Simulation," https://www.carsim.com/products/carsim/index.php (accessed July 12, 2022).

[67] "CarMaker - IPG Automotive," https://ipg-automotive.com/en/products-solutions/software/carmaker (accessed July 12, 2022).

[68] "STK - AGI," https://www.ansys.com/products/missions/ansys-stk (accessed July 12, 2022).

[69] Y. Ma et al., "Model predictive control for the operation of building cooling systems," *IEEE Trans. control systems technology*, vol. 20, no. 3, pp. 796–803, 2011.

[70] J. Jacobson, P. Janevik, and P. Wallin, "Challenges in creating AstaZero, the active safety test area," in *5th Conf. Transport Research Arena: Transport Solutions from Research to Deployment*, 2014.

[71] Y. Geng et al., "A survey of industrial control system testbeds," *IOP Conference Series: Materials Science and Engineering*, vol. 569, no. 4, p. 042030, jul 2019. https://doi.org/10.1088/1757-899x/569/4/042030

[72] C. Siaterlis, A. P. Garcia, and B. Genge, "On the use of emulab testbeds for scientifically rigorous experiments," *IEEE Communications Surveys Tutorials*, vol. 15, no. 2, pp. 929– 942, 2013.

CHAPTER 6

Education

Education

João P. Hespanha, Christos G. Cassandras, Panos J. Antsaklis, Gireeja Ranade, J. Anthony Rossiter, Florian Dörfler, Steven L. Brunton

Education is the prime catalyst for growth and prosperity in Control Systems. This chapter discusses university curriculum changes in this field of study. It proposes specific suggestions for concepts and methodologies to emphasize and ways to adapt to a new generation of students.

Abstract In the face of rapidly developing societal challenges, the continued growth and success of the Control Systems field will rely on continued learning and education of rapidly developing societal scale challenges. While we recognize the need for greater outreach to secondary and high-school populations, we believe there are much broader needs across all areas of science and engineering. Thus, our emphasis in this report is on curricular changes at the university level. We assembled a panel of experts in the field whose deliberations have led to three key recommendations:

1. Introductory courses in Control Systems should emphasize the broader applicability of control.
2. Control Systems courses should be made available earlier in university curricula with the aim of teaching fundamental concepts to a broader group of students.
3. Given the scope of the Control Systems field, modular organization of teaching material may be a more effective way to impart knowledge.

These recommendations are based on the recognition that new mathematical models and methodologies are required for the control design of systems that dominate 21st century technology. This is especially true for nonlinear systems with increased levels of uncertainty, those with humans-in-the-loop, and those with the emergent properties of safety and resilience. Also worthy of special attention are systems with hybrid dynamic characteristics, i.e., a combination of time-driven dynamics (for physical processes) with event-driven dynamics (for computer-based and networking components). This report provides specific suggestions for how to structure and name new courses, adapt to the new generation of students, decide on different concepts and methodologies to emphasize, and identify appropriate modules from which courses can be constructed. We also include an example of a new introductory course currently in use.

6.A Present State and Future Outlook

In the field of Control Systems, our overall goal is to make a system behave in a desirable way. It stands to reason that the field is driven by the systems that society is interested in and relies on. In studying and developing control mechanisms, we seek to understand the characteristics and dynamic behavior of these systems. As the world's needs and aspirations evolve, these vital systems become increasingly more complex. Not only are the problems we need to address more challenging, but the performance expectations are more demanding and specifications are much tighter.

Among the key societal drivers that define the applications where control plays an increasingly important role are:

- Feeding the world: digital farming
- Climate change
- Space: the final frontier
- Underwater exploration
- Biomedical engineering and disease control
- The Sharing Economy
- Cybersecurity
- Geostrategy

These examples—spanning numerous disciplines like aerospace, biology, and economics—illustrate the breadth of scope for Control Systems. Education will be key to addressing these rapidly evolving challenges.

Education includes the K–12 levels as well as undergraduate and graduate-level curricula. It also goes beyond formalized curricula and into professional training that fills in the gaps between courses and addresses specific industry needs. Here, we restrict our inquiry to university-level curricula in control theory and engineering. We do, however, recognize the need for outreach to secondary and high-school populations, as well as the cultivation of entrepreneurship in postsecondary learning environments.

The topics of training and industry engagement are also of critical importance when it comes to education, and we refer the reader to Chapter 5.

6.B Curriculum Changes

Three significant recommendations were identified by our expert panelists:

1. Introductory courses in Control Systems should emphasize the broader applicability of control.
2. Control Systems courses should be made available earlier in university curricula with the aim of teaching fundamental concepts to a broader group of students.
3. Given the scope of the Control Systems field, modular organization of teaching material may be a more effective way to impart knowledge.

6.B.1 Updating the First Course In Control for Broader Applicability

Why is updating introductory control courses an important and urgent need? Unlike the technological innovations of the mid- and late 20th century, the world's emerging systems are nonlinear, feature increased uncertainties (e.g., with humans-in-the-loop), and feature hybrid dynamic characteristics, i.e., a combination of time-driven dynamics (for physical processes) with event-driven dynamics (for computer-based and networking components). Such systems are now often referred to as cyber-physical. Novel mathematical models and methodologies are needed for the control design of these systems.

Frequency vs. Time Domain Methodologies: Frequency domain methods—so prevalent in electrical engineering, with their strong connection to linear constant coefficient Ordinary Differential Equations (ODEs)—are often the only control design methods that beginning students are exposed to. But they may not be the best way to convey the generality and importance of control. In introductory courses, it is no longer sufficient to exclusively focus on the design of LTI controllers using traditional methodologies like Bode plots, Nyquist plots, and the Root Locus method. It is important to introduce more general systems than the ones described by linear ODEs with constant coefficients, including the incorporation of nonlinear models, automata, and hybrid automata. A modern understanding of Control Systems must address systems in a student's daily life, from digital control components embedded in smartphones and portable healthcare devices to large-scale networked systems in mobility-on-demand and autonomous UAV teams.

This raises the question: Should an introductory course emphasize modeling, optimization-based control, data-driven control, and time domain methods? Although it may be exciting to discuss ingenious approaches to design that use approximate relations and shortcuts in the frequency domain that were introduced before the digital era, that does not address current students' interests.

Embracing Computational Methods: Conventional control courses primarily focus on analytic methods, but in their professional lives, students will rely on computational techniques. Rather than resisting this paradigm shift, control courses should locate and highlight synergies between analysis and computation, e.g., between classical lead/lag compensation and numerical loop shaping techniques, or between dynamic programming and reinforcement learning. One of the main challenges in this new approach is to identify a set of essential computational tools that address the types of problems that capture the imagination of 21st-century students. Examples could include LMIs, Sum of Squares (SOS) programs, and possibly the rudiments of MPC. Computational techniques such as least-squares (for system identification or other data-driven control) can be introduced in first-year curricula through Jupyter Notebooks or other such tools.

Incorporating Machine Learning: The rise of machine learning and artificial intelligence provides another opportunity to innovate control theory education. Machine learning is undoubtedly changing our modern engineered world, and students are expressing a strong demand for coursework that prepares them to be a part of this exciting future. Viewed as the process of building models from data using optimization, it is clear that machine learning is closely related to the field of system identification, which has been a cornerstone of control theory since the 1950s. There are several major opportunities for research at the intersection of control theory and machine learning, each of which motivates a new and compelling curriculum. "Learning for control" is a direct descendant of system identification, enabling the data-driven characterization of complex systems with challenging and often unstructured dynamics that cannot be easily expressed in a simple analytic form. This data-driven approach creates new challenges related to certifiability and guarantees for autonomous systems and other safety-critical applications. Finally, the field of reinforcement learning, which is at the intersection of control theory and machine learning, is one of the most rapidly growing areas of both fields and provides promising avenue for general artificial intelligence. Students are drawn to these new technologies, both for their ability to shape the future and the added value in an education that has core strength in both classical and modern approaches.

Adapting to Our Students: Raised in the era of computer games and smartphones, today's students yearn to see control theory's immediate connection to truly exciting applications. They want to see the big picture first before delving into the technical details. The traditional advice—"study mathematics and physics for a few years and then everything will fall into place"—simply does not work anymore. It requires attention spans that have been drastically shortened. Instead of bemoaning this new reality, we should adjust to it and adopt a new approach—perhaps completely reversing the instructional order.

When we fail to connect control principles to applications that excite students, we sell Control Systems short. Rather than showing how powerful these principles are and how important the field is to every aspect of our daily lives, we revert to the fallback position that control is "hidden technology"—a catchy motto that does not help the field to grow.

To adapt to the new generation of students, we have to remember that a substantial proportion of college students now come from nontraditional (and in many cases underprivileged) communities. Even the word "engineering" may be unfamiliar, and social and educational priorities may be very different. We need to be sensitive to this fact in designing courses that introduce control principles and their use in addressing societal needs. While it is not clear if this trend is connected to specific geographical regions or societal structures, it is certainly well-documented and it calls for our increased attention.

As an example, a course called Designing Information Devices and Systems (EECS16AB) at the University of California, Berkeley[1], takes a modular approach, as suggested later in Section 6.B.3. It also introduces linear algebra and control concepts through practical applications that motivate students to enter the field (e.g., medical imaging and tomography, GPS design, and brain-machine interfaces and robotics.) These applications allow students to see the real power of the mathematics that underlies control theory. Students actually build simplified demonstrations of these applications in hands-on hardware labs. Figure 6.1 illustrates the first lab that students are assigned in the course sequence—the medical imaging lab. Students project a series of masks on an object they wish to image. Through this, they can set up a system of linear equations (the unknown variables being the pixels that make up the image) as in a tomography setup, and solving this system reconstructs the image. This simplified version of tomography demonstrates how linear algebra can be used to model a system, while giving students a sense of what happens in real-world engineering contexts. Matrix inversion and the Gaussian elimination algorithm—which might otherwise be seen as cumbersome—can be also motivated through such practical applications.

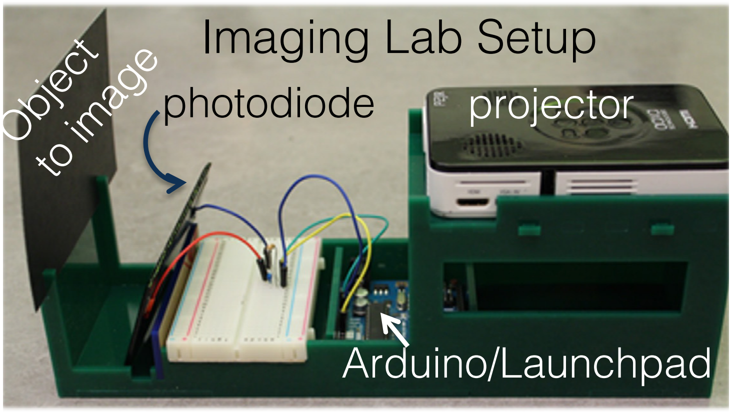


Figure 6.1: Illustration of the Imaging Lab in EECS16A. The projector illuminates the object to image, and the reflected light is collected by the photodiode, which is connected to the Arduino. By using different illumination masks, students collect different linear combinations of pixels and can use matrix inversion to generate the image.

New and Meaningful Course Titles: An elective course called Control Systems may not be attractive. It may be preferable to use a name like Building Autonomous Systems. Autonomous systems are, in fact, more general and more ambitious control systems. They can have many objectives achievable under significant uncertainties. Depending on the specific institution's programs, priorities, and community role, new names could be adopted for introductory courses in Control Systems.

[1] For more details, see https://eecs16a.org.

6.B.2 Introducing a Control Systems Course Earlier, for Broader Audiences

Control is an important concept in many different fields ranging from biology and political science to economics and engineering. Thus, we need to significantly broaden the intended audience for introductory control classes, while maintaining the field's integrity as a distinct area of learning.

One of the realizations that emerged from our international panel is that there are significant differences across educational systems around the world—in terms of class sizes, industry influence, required versus elective courses, etc. In some parts of the world, every engineering student has to take a course on control theory basics. In typical U.S. electrical engineering curricula, the first control course is often a junior or senior elective, while in mechanical engineering, control concepts are embedded in several undergraduate courses.

The broad applicability of control concepts is both a strength and a weakness. A course for a wide student audience could start with simple problems relying on computational models and introduce mathematical descriptions later. An introductory course need not delve into theoretical aspects—it can instead focus on practical tools and foundational concepts. Such curricular changes can attract new talent to the field, but they need to be implemented now, as they will take several years to have full impact.

Key Concepts to Emphasize: Perhaps the foremost concept in control is that of modeling. Models encapsulate what we know about any process of interest. They can be linear or nonlinear ODEs, PDEs, DES models such as automata and Petri nets, digital simulations, word descriptions, sets of data with rules of engagement, etc. Linear ODEs with constant coefficients represent the simplest mathematical dynamic models we have and can be used as key examples because we know a lot about them. It should be made clear that models are always approximations of real processes and that there will always be uncertainties in processes and environments which typically cannot be directly measured.

To effectively control the system's behavior, we first need stability, so that when a system is perturbed (e.g., airplane in a gust of wind) it is able to return to its previous state. Along the same lines, adaptation is needed for a system to be responsive to large changes. The current definition of "adaptive control" is restricted to rather small changes. Adaptation needs to be taught in a more general sense, and it should be linked to failure diagnosis, control reconfiguration, machine learning, etc. Elements of learning theory are also an essential ingredient of modern Control Systems.

The most powerful concept involved in controlling a system with uncertainties is feedback. It is important to recognize that feedback transcends models. Feedback mechanisms in engineered systems (including vehicles, robotics, manufacturing, and chemical process control), biology, economics, politics, and the environment are ubiquitous and can be introduced with simple explanations and examples.

Key Methodologies: A Control Systems course for a broad audience should start with time domain approaches using state variable descriptions (e.g., state feedback approaches) that can also be used for time-varying systems, discrete event systems, etc. Methodologies for systems described by ODEs with constant coefficients should continue to be taught, as should frequency domain approaches using transfer functions. It is important to emphasize the role that optimization plays in control theory, and it should be connected to powerful and popular methodologies like linear programming, dynamic programming, and linear quadratic regulation.

New methodologies for highly autonomous systems include CS/AI approaches. They may be conveniently described in terms of a hierarchical functional architecture, where to achieve higher autonomy we need to add higher layers of methodologies drawn from areas like planning and learning. However, it is important to clearly differentiate the role of control from that of learning.

A course for broad audiences should maximize the use of simulation—possibly including competitions that can stimulate student participation and creativity (e.g., see Figure 6.2). Physical experiments are desirable because they make a lasting impression and represent real-world uncertainties that simulations may not be able to capture. However, they are more expensive and need more maintenance. On the other hand, simulations are much more flexible and are able to represent a variety of situations (including large-scale systems that may not even exist yet) with increasing realism. It can be beneficial to have a publicly available collection of recommended simulations for particular cases (e.g., MATLAB and Simulink examples).

What would an updated introductory course look like? We envision a course that consists of a number of distinct educational modules as described next.

Figure 6.2: A proud first-year student showing off her final project for EECS16B at UC Berkeley: a voice-controlled robot car. This project uses principal-components analysis to learn voice commands and pole placement for control of the car.

6.B.3 Modularizing the Teaching Experience (Short Modules, Instead of Courses)

The motivation for a modular approach to the design of an introductory course in control is based on the following facts:

- It will require significant effort from instructors to change the thrust of control courses. The free availability of modules on several basic topics—along with simulated examples and experiments—would greatly facilitate the process of new course creation.
- Instructors who want to update introductory control courses cannot afford to wait for a new textbook. The process would be greatly facilitated by access to the CSS or IFAC websites, from which one could download well-designed modules of lecture material (with, at minimum, detailed summaries and selected definitions and methods), and of simulation (e.g., MATLAB-based) modules.
- Such control modules can be easily added onto existing courses on, e.g., robotics, communication, autonomous vehicles, or power systems.

The quantity and depth of these modules will depend on the time allocated and the students' background. Some thought needs to be put into the design of reasonably sized modules on certain topics. We believe that the CSS together with the IFAC should establish an Undergraduate Curriculum Committee (the IFAC may already have something along these lines). This Committee could review module submissions, certify a selected subset, and organize and recommend those that are certified.

In a prototype broad-appeal course, students would be introduced to general concepts such as fundamentals of systems modeling, sensors and actuators, and decision-making. The broader role of Systems Science, along with Control Systems, should be one of the ingredients of such a course. An award-winning module on the "Power of Feedback" could be used to spread the word to many other disciplines.

Possible modules could include:

- Feedback
- Models and uncertainty
- Mathematical models for linear, nonlinear, and hybrid (time-driven and event-driven) systems and simplifications (e.g., reduced models, abstractions, MATLAB-based computational models, and data-based models)
- Stability
- Adaptation and learning
- Data-driven control
- Applications (e.g., robotics, autonomous driving, smart grids)
- Important emerging topics such as systems of systems and security

The notion of modularizing basic Control Systems knowledge is also based on our increasing capacity for leveraging technology. For example, a student's learning experience can be greatly enhanced by bringing experimentation to the student's home through:

- The availability of small, inexpensive hardware
- The ability to share resources around the world in the form of virtual labs

Education has experienced significant changes in the past two decades with the advent of massively open online courses (MOOCs), hybrid courses, and online educational material more generally. Since the COVID pandemic, these alternative educational formats have only strengthened their positions. There are several benefits to open and online course material. Online content can be modular, flexible, and tailored to specific sub-audiences, without requiring several duplicate courses for every combination of programming language, mathematical preparation, or application domain. They also provide hierarchical access to concepts, so that the student is able to obtain the most important ideas with the given time available, whether that is 10 minutes, one hour, one week, or an entire university-level course. Professors are, therefore, increasingly leveraging online content for their own classes. Online content also means that high-quality educational material is not just available to students at rich and premier institutions, thus democratizing education. Prime examples are YouTube channels youtube.com/@eigensteve and youtube.com/@ControlLectures (see Figure 6.3), which have international viewership, including students in rural areas and others that would not otherwise have access to first-rate lectures in our field. In addition to broadened access, online material is also useful for industry professionals who seek to improve their skills on their own time. Although the verdict is still not out on how MOOCs can productively co-exist with in-person education, it is critical that the control community continues to embrace and explore these alternative educational formats, both to provide the highest quality education possible and also to broaden access and participation beyond elite institutions.

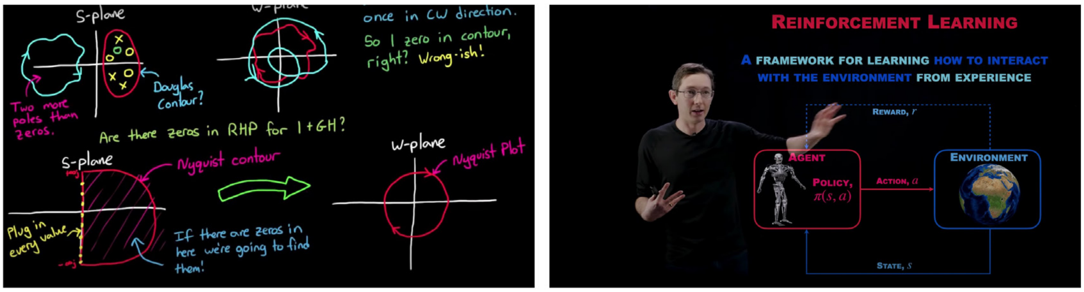


Figure 6.3: Screenshot from Brian Douglas's Control Lectures at engineeringmedia.com (left). Screenshot from Steven Brunton's reinforcement learning lecture at youtube.com/@eigensteve (right).

6.C Creating Success Stories in Curriculum Changes

The three main recommendations presented above delineate the basic structure and content of a revised curriculum in Control Systems. Given the scope of this chapter, we have not attempted to address the question of how to teach the proposed content nor to provide examples of specific case studies that promote these recommendations. However, we refer the reader to at least one prior effort along these lines as described in "Report on the NSF/CSS Workshop on New Directions in Control Engineering Education," *IEEE Control Systems*, pp. 53-58, October 1999. The reader is also pointed to a comprehensive paper in the 2023 Annual Review in Control [1], which includes three concrete examples of how the Control Systems curriculum can be brought into the 21st century in terms of i) The topics we teach and the emphasis different aspects are given, ii) How we manage delivery and student engagement, and iii) How we use laboratory exercises to support student development. The first example focuses on the holistic design of a large first course. The second looks at the expanding area of take-home laboratories and how these give students the opportunity to develop independent learning skills and a deeper understanding of core principles. The final case study addresses the role of virtual and remote laboratories in supporting inclusive learning.

Finally, we point out that the IFAC TC9.4 on Control Education has worked on the definition of the content of a first (i.e., unique) control course (see [2]) and is currently working on the collection on existing web resources for those topics that have been considered to be more relevant.

Recommendations

For young researchers: Introductory courses in Control Systems should emphasize the broader applicability of control. Such courses should be made available earlier in university curricula with the aim of teaching fundamental control concepts to a broader group of students. Given the scope of the field, modular organization of teaching material may be a more effective way to impart knowledge.

For funding agencies: Funding agencies can solicit proposals to develop modules for different control theory topics that are suitable for introductory courses.

BIBLIOGRAPHY

[1] J. A. Rossiter, C. G. Cassandras, J. Hespanha, S. Dormido, L. de la Torre, G. Ranade, A. Visioli, J. Hedengren, R. M. Murray, and P. Antsaklis, “Control Education for Societal Scale Challenges: a Community Road Map," to appear in *Annual Review in Control*, 2023

[2] J. A. Rossiter, A. Serbezov, A. Visioli, K. Žáková, and M. Huba, “A survey of international views on a first course in systems and control for engineering undergraduates,” *IFAC Journal of Systems and Control*, vol. 13, p. 100092, 2020.

CHAPTER 7

Ethics, Fairness, and Regulatory Issues

Ethics, Fairness, and Regulatory Issues

Anuradha M. Annaswamy, Munther Dahleh, Matin Jafarian, Françoise Lamnabhi-Lagarrigue, Mariana Netto

The Control Systems community will play a vital role in designing technologies that respect human values, ensure ethics and fairness, and meet regulatory guidelines, even while safeguarding our environment and our natural resources.

Abstract The true scope of control systems is significantly broader than what current and past research implies. Research objectives related to ethics, equity, fairness, and social responsibility have been conspicuously absent. This chapter outlines some key topics related to the ethical implications of autonomous systems, fairness and discrimination, and intersections with regulatory issues.

As the magnitude of the problems that our community seeks to address increases, the underlying domains and boundaries of the systems that we tackle necessarily expand as well. As we move from single systems to networks of systems where humans play an increasingly integral role—and from domain-specific performance metrics to much larger societal goals—we shift from a pure engineering perspective to a larger purview of science and technology. A whole range of emerging technologies (e.g., artificial intelligence, machine learning, pervasive sensing, internet of things, and nano-, neuro-, genetic-, and biotechnologies) are increasingly being incorporated into complex engineered systems (e.g., smart cities, autonomous vehicles, advanced weapon systems, and assistive devices). To ensure responsible innovation in these areas, researchers need to focus on ethical considerations. This advancement will be imperative for realizing a more sustainable and equitable future.

In this chapter, we outline some of the relevant areas where our community has begun to apply systems and control tools. Questions that have arisen in this context include:

- How do we develop principles and guidelines that assign specific duties and responsibilities to the various stages of research, design, development, and adoption of technologies?
- How do we ensure diversity, equity, and inclusion when designing and deploying these new technologies?
- How should we manage systems with unforeseen reactions and less than fully reliable outcomes?
- How do we assess and control the short-, medium-, and long-term impacts (physical, mental, biological, social, and ecological) on humans and the environment that arise from the application of emerging technologies, especially automation?
- How do we foster education and public awareness at all levels (including engineers, other scientists, and policymakers) towards the development of ethically aware and responsible products and services?

7.A Ethics of Autonomous Systems

Recently, dedicated efforts have been undertaken within various research communities to define ethics guidelines. For instance, in the area of autonomous and intelligent systems, the report *Ethically Aligned Design*, published by the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (A/IS) [1] is based on three pillars that capture the anthropological, political, and technical aspects of ethics and design:

1. Universal human values
2. Political self-determination and data agency
3. Technical dependability

These pillars form the basis of eight general principles that are considered to be imperatives for the ethical design [2]:

i. Human Rights–A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights.

ii. Well-Being–A/IS creators shall adopt increased human well-being as a primary success criterion for development.

iii. Data Agency–A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity.

iv. Effectiveness–A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS.

v. Transparency–The basis of a particular A/IS decision should always be discoverable.

vi. Accountability–A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.

vii. Awareness of Misuse–A/IS creators shall guard against all potential misuses and risks of A/IS in operation.

viii. Competence—A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.

This overall ethical framework also appears suited to a Control Systems approach to autonomous systems. We can aim to ensure that these principles are fulfilled at each stage of problem formulation and technology development.

7.B Ethics and Fairness

As technology progresses, its benefits do not equally reach all sectors of society. It has been argued that the development of technology has often been agnostic to or even detrimental to fairness and justice. As resources become scarce—especially in zero-sum situations—any computational framework that is developed should account for notions of bias, fairness, and justice. It is imperative that metrics that address these notions be interwoven into overall problem statements and computational solutions.

One salient line of investigation is the impact of automation on the global community [3, 4]. For instance, in [4], it is emphasized that "although advances in automation have been a key contributor to much of modern life and quality of life, they have also raised concerns in various respects. If the technology-society nexus was complicated before, with the emerging technologies of today a compounding of potential interplays is imminent. With self-driving cars, facial recognition systems, gene editing technology, quantum encryption devices, robotics systems, exoskeletons, and other new and emerging tech, the concerns go well beyond loss of employment and increasing inequality. The evolution of automation is rapidly resulting in, or exacerbating, major concerns about societal impacts, beyond the benefits." In robotics circles, this issue is of obvious importance [5]. On a broader level, NSF initiatives in the U.S. have focused on the future of work at the human technology frontier [6]. All of these discussions become urgent given the explosive growth in AI and ML and their applications to physical and safety-critical infrastructures. A relevant framework in this context can be found in [7]. The above discussions raise a broader question of ethics, i.e., all aspects of the service that new technologies provide pertaining to fairness, equity, impartiality, and honesty [8]. The problems addressed by our community must therefore include an overall framework for ethics.

7.C Regulatory Issues

Any product that is designed to be used by society needs to be approved by a regulatory body. Whether in healthcare, robotics, energy, or transportation, every sector in every country around the world is subject to oversight from appropriate entities to ensure the safety of operators and the general public. As technologies advance and new products emerge, technology transitions must accommodate regulatory considerations. Early examples can be found in the area of medical devices, with the Food and Drug Administration (FDA) in the United States as the regulatory body that is responsible for assuring safety, efficacy, and security of smart medical devices (see Section 2.B of this document). At times, regulations evolve in parallel with technology. For example, the Vienna convention related to road traffic was amended after new Advanced Driving Assistance Systems (ADAS) were designed and deployed. This suggests that there could be a coevolution of technology and regulation leading to more efficient societal functioning. As our community's research loci extend deeper into domains that involve societal needs, such intersections with regulatory issues are bound to increase. When technological advances are not merely incremental but represent major breakthroughs, these intersections become even more significant, requiring the research to inform and direct changes in policy. One example is the current paradigm shift in the power grid. Rapid decarbonization has necessitated significant conversations in the policy sector regarding electricity market structures, as well as ownerships, rights, and privileges for various emerging stakeholders [9].

Across the board, safety has taken center stage as a pervasive system property. Whether in robotics, self-driving vehicles, or any other application that is witnessing increased automation and societal integration, issues of safety are paramount. As such, safety standards have to be formulated, reformulated, or revisited to address this changing landscape. Safety standards are developed and maintained by the International Organization for Standardization (ISO). ISO 10218-1:2011, for instance, which prespecifies a safe distance between robots and their environment, has led to a more stringent standard, ISO/TS 15066:2016, so to support a collaborative environment between humans and robots. A similar effort is afoot in EU to introduce a uniform regulation across all EU nations for the operation of UAVs in civilian airspace. Regulatory concerns need to be carefully addressed in areas related to privacy and cybersecurity. Defining data property rights and addressing damage from breaches are just two of numerous issues that regulators need to focus on. It will be critical to establish guidelines for incentive design in order to proactively secure infrastructure across robotics, AI, and Industry 4.0.

In the context of ethics and fairness, particular regulations could be designed to ensure that no misuse of useful technology occurs. Along with the coevolution of technology and regulation, parallel developments in legal and technological pursuits are essential as well.

7.D Intersection With Control

Here are some important control-theoretic issues to consider:

1. **Ethical Issues of Emerging Technologies:** In addition to autonomy, another new technology is the reprogramming of cells (e.g., CRISPR), which poses serious ethical dilemmas. Reprogramming is an emerging area of intervention/control that is intertwined with learning on networks (as discussed in the bio section). While this area is fascinating and may enable us to eliminate serious genetic diseases, it brings out ethical issues in terms of what society could do with such technology. Much of that impact is not very well understood or appreciated. Many other domains raise ethical issues as well [4]: human enhancement technologies, emotional robotics (Figure 7.1), social media, cybersecurity and smart cities, neurotechnologies, synthetic biology, quantum technologies, nanotechnologies, autonomous weapons (Figure 7.2), And virtual and augmented reality. As pointed out in [1], an important general principle is that “the behavior of autonomous functions should be predictable to their operators, that those creating these technologies understand the implications of their work, that professional ethical codes are developed to appropriately address the development of autonomous systems and autonomous systems intended to cause harm, that designers not only take stands to ensure meaningful human control, but be proactive about providing quality situational awareness to operators and commanders using those systems.” Control systems have a major role to play in the understanding of areas like machine learning, control theory, and optimization. Ensuring robustness, precision, and predictability is key to addressing ethical issues.
2. **Platforms and Intrinsic Bias:** Platforms are typically predicated on designing a system that allows for interaction between multiple agents—and those are often designed following market principles. The online ad market has become one of the largest markets of the 21st century. Ad markets are basically auction mechanisms in which firms (companies) compete to get access to a particular user on the platform. People have observed that such systems can result in nontrivial bias. One case is the advertisement of STEM course. Data collected from such markets has demonstrated that women typically see many fewer STEM ads than men. Curiously, this is not due to companies favoring men,

but rather to the inability of those companies to win at auction given the fierce competition among retail firms targeting women. Similar bias has been demonstrated in other platforms including Zillow and other real estate apps.

3. **Design of Public Signals:** The field of control theory has addressed the question of designing private and public signals to nudge people to perform in a particular way. For example, it has been shown that the publication of COVID-19 infection data has impacted infection rates. As people see that numbers are rising, they exercise more restraint in interacting with others, resulting in negative feedback that mitigates spread. These public signals can sometimes be misrepresented and even result in exactly the opposite of the intended effect. The design of such signals, where they are published, and how they are presented to different groups, can all be critical in sending a consistent message to the public.

4. **Unintended Consequences of Control/Decisions:** Typically, control decisions are made to optimize certain performance objectives. We are quite familiar with trade-offs in performance objectives. However, it is possible to miss these trade-offs as humans interact with such systems. An example of such a problem is toll pricing. The implementation of transportation demand management through dynamic toll pricing (which is, in principle, an adaptive control problem) has long been recognized to potentially result in discrimination against low-income people. Such consequences may not be directly captured by a system's performance objective.

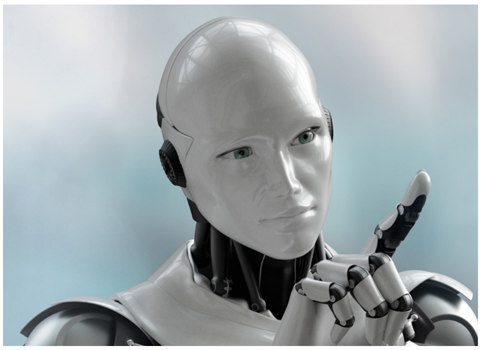


Figure 7.1: New technologies in emotional robotics require collaboration with emotion researchers in psychology, neuroscience, and computer scientists [10] (reprinted with permission).

Figure 7.2: Control systems could bring complementary expertise in order to ensure meaningful human control of autonomous weapons systems [11] (courtesy STM Photo).

Recommendations

For young researchers: Societal considerations such as ethics, fairness, justice, and social responsibility are becoming more important as the digital way of life is becoming more common across the globe. Control theory has much to offer in quantifying these issues, even as it focuses on the foundations of enabling technologies. New research is needed to explore how quantitative thinking, modeling, and analysis can be carried out to address these societal-level challenges.

For funding agencies: Funding agencies should support joint initiatives that address ethics, fairness, and social responsibility in the context of ubiquitous digital technology. To address societal-scale challenges over the coming decades, we will need a better understanding of the fundamental principles of socially responsible automation. Interdisciplinary and cross-disciplinary perspectives will be essential to the analysis and synthesis of tools. Among these, control systems-based perspectives will be of special importance and will require the attention of funding agencies.

BIBLIOGRAPHY

[1] IEEE, *Ethically Aligned Design*. IEEE Press. https://standards.ieee.org/news/ead_v2/

[2] *From Principles to Practice Ethically Aligned Design Conceptual Framework.* IEEE. https://ethicsinaction.ieee.org/wp-content/uploads/ead1e_principles_to_practice.pdf

[3] D. Rotman, *The Relentless Pace of Automation.* MIT Technology Review, 2017.

[4] F. Lamnabhi-Lagarrigue and T. Samad, *"Social, Organizational, and Individual Impacts of Automation" in Handbook of Automation*, S. Y. Nof, Ed. Springer (in press), 2022.

[5] A. Edlich and V. Sohoni, *Burned by the bots: Why robotic automation is stumbling.* McKinsey & Company: Digital McKinsey, 2017.

[6] NSF, "Future of work at the human-technology frontier." https://www.nsf.gov/news/special_reports/big_ideas/human_tech.jsp

[7] M. A. Sampath and P. P. Khargonekar, "Socially responsible automation: a framework for shaping the future," *Winter Bridge on Frontiers of Engineering,* vol. 48, no. 4, pp. 45–52, 2018.

[8] P. P. Khargonekar and M. A. Sampath, "Framework for ethics in cyber-physical-human systems," *IFAC-PapersOnLine*, vol. 1, no. 53, p. 2, 2020.

[9] National Academy of Sciences Engineering and Medicine, *The Future of Electric Power in the United States*. Washington, DC: The National Academies Press, 2021

[10] "New 'emotional' robots aim to read human feelings," https://www.news24.com/fin24/tech/new-emotional-robots-aim-to-read-human-feelings-20180111 (accessed March 14, 2023).

[11] "Companies Developing Lethal Autonomous Weapons, As Groups Seek Ban, Report Says," https://www.aviationtoday.com/2019/12/02/companies-developing-lethal-autonomous-weapons-as-groups-seek-ban-report-says/ (accessed March 14, 2023).

CHAPTER 8

Recommendations

Chapter 2 Societal Drivers

2.A Climate Change Mitigation and Adaptation

Climate change is the gravest challenge facing humanity. The urgency and scale of the problem present tremendous opportunities for collaborative research on various topics from infrastructures and communities, energy, transportation, industry and manufacturing, and food and agriculture.

2.B Healthcare and Ensuring Quality of Life

The control systems community can positively impact society by tackling healthcare and quality-of-life challenges. From mitigating the risks of pandemics to leveraging the power of the medical internet of things, neuroengineering, and the closed-loop control of medical devices, tools derived from a control systems approach can enable breakthroughs on a large scale.

2.C Smart Infrastructure Systems

The scale and scope of overcoming the challenges in building smart infrastructure systems will require collaboration across multiple domains beyond control systems, ranging from other engineering disciplines to public policy and administration. Realistic testbeds are needed to produce critical datasets and validate new methodologies.

2.D The Sharing Economy

The economy of sharing should be on the radar of the control systems community. Engineers are often at the forefront of the battle to make more efficient devices while ignoring the stimulus these provide to economic consumption. Future business models will need to decouple economic growth from increased consumption. Control systems has much to offer in this context through the design of systems where access to resources is governed efficiently and justly.

2.E Resilience of Societal-Scale systems

To build resilience in next-generation societal infrastructures, comprehensive and contextual design approaches are needed, especially in the face of emerging technological, environmental, and societal risks. Foundational tools such as network system science, control and incentive mechanisms, and optimal resource allocation designs are important resources needed to achieve this goal.

Chapter 3 Technological Trends

3.A AI and Big Data

As humans and machines become mingled in sociotechnical systems, it is essential to understand the interplay between control systems and machine learning, data science, reinforcement learning, and AI. Complexity measures, decision-making across multiple timescales, mitigating the sim-to-real gap, and developing analytically rigorous tools for robust performance are only a few examples of the myriad opportunities in this environment.

3.B Electrify Everything

The increased use of electrical energy to support human needs is a megatrend that will influence all market sectors. Optimally managing key energy assets requires coordination of the wide and disparate sources, sinks, and energy storage that operate over multiple spatial and temporal scales. Co-design, distributed decision-making, computation over disparate networks, and accommodating adversarial actors are only some examples of many in which the control community can contribute.

3.C Engineering Biology

Engineering biology's methods of designing, building, and testing engineered biological systems is critically relevant to improving human health, the economy, and the environment. The role of control systems is valuable in understanding and designing the emerging large-scale systems in this field that can operate in the presence of uncertainty.

3.D Robots in the Real World

In order to realize the full potential of robotics operating in the real world and have them achieve truly disruptive impact, they have to be made effective, resilient, and safe. This may require a focus on the higher levels in the control systems hierarchy, such as perception, situational awareness, autonomy, AI, and the use of digital twins.

Chapter 4 Emerging Methodologies

4.A Learning and Data-Driven Control

Applications drive the fruitful interaction between machine learning and data-driven control. To fully exploit synergies between the two disciplines, the foundations of systems and control need to be revisited with a focus on the intersection between dynamics, scalability, learning, and architecture.

4.B Safety-Critical Systems

Ensuring intelligent and autonomous systems operate safely is critical to establishing human trust in these system and enabling their widespread deployment. A formal representation of safety is needed to provide a paradigm for synthesizing, testing, and guaranteeing safe behavior that can be implemented in real-world systems.

4.C Resilient Cyber-Physical Systems

New complex CPS are being created to increase performance for systems in mobility, energy, water, and other vital infrastructures. Designing these systems to be resilient to cyberattacks, extreme weather events, and other adverse events enable infrastructures to quickly detect and recover from these disruptions.

4.D Cyber-Physical-Human Systems

CPHS have diverse applications in healthcare, logistics, production, and infrastructure systems. Understanding the challenges of designing control systems to better interact with humans requires research into the derivation of control policies that are safe, robust, adaptive, and trustworthy as well as the computational modeling of humans as basis for model-based control.

4.E Control Architecture

Societal-scale control systems are composed of a variety of interacting elements, such as sensors, actuators, computers, communication devices, algorithms, software, and human-machine interfaces. Control architecture design principles play an important role in putting control technologies into real-world societal systems.

Chapter 5 Technology Validation and Translation

The industrial ecosystem needs systems and control. Next-generation research validation can bridge the gap between academia and industry and enable the development of tech transfer and control technologies and products with high societal impact.

Chapter 6 Education, Training, and Retraining

Because new mathematical models and methodologies are required for the control design of systems that dominate 21st-century technology, control systems curricula must stay current and applicable.

On the university level, introductory control systems courses should emphasize and discuss the broader applicability of control. These courses should be introduced earlier in the curriculum and teach fundamental control concepts to a broader array of students. Modular organization of teaching material may be a more effective way to impart the basics of control systems.

Chapter 7 Ethics, Regulatory Issues, and Interoperability

Problems related to societal issues such as ethics, fairness, social responsibility, and justice are becoming more and more important as the digital way of life is becoming more and more common across the world population. Control systems has much to offer in this context through incorporation of elements that quantify these issues even as it focuses on analytical foundations of the underlying technology. New research is needed to explore this fascinating area as to how quantitative thinking, modeling, and analysis can be carried out for these people- and sociocentric challenges.

Epilogue

Epilogue

This road map is intended to guide a shift in how control systems can meet grand societal-scale challenges. The societal drivers, technological trends, emerging methodologies, and technology validation and translation explored in this document have captured this role. We identify paths that can be taken for control systems to effectively address the challenges facing global society.

This document is a product of a two-year effort by several community members. The topics explored in all chapters reflect our current understanding of some of the dominant directions our community may need to move in. While we have made every effort to avoid major gaps and omissions, there are bound to be certain methods and action items that we have missed. We encourage the community to spotlight these gaps in appropriate forums.

As mentioned elsewhere in the document, this road map is intended to provide strategic directions for the community, and elements of this road map may serve to form the core of the IEEE control systems community's strategic plan. As we progress into the 21st century and the needs of society change, we expect new road maps to suitably evolve, address new methods, and embrace new tools and trends.

Acknowledgments

Acknowledgments

As the saying goes, it takes a village to raise a child—and it has taken a global village to put this document together. We would like to take this opportunity to thank several individuals who provided time and effort to help with this process.

First and foremost, we would like to thank the IEEE Control Systems Society for supporting this initiative. The process started during Anuradha Annaswamy's presidency in 2020 and took shape during Thomas Parisini's presidency in 2021–2022, with crucial support from the CSS Executive Committee and the Board of Governors. For this, we are truly grateful. We would like to thank the National Science Foundation for providing travel support for the panelists and participants to attend the 2022 Workshop. The deliberations therein were crucial for the creation of this volume. We would like thank Digital Futures for providing us support for boarding and lodging during the 2022 Workshop. Finally, we would like to thank IFAC for the support in publishing this roadmap.

Second, we would like to thank the many members of the community who assisted us in the creation of this roadmap: authors, panelists, reviewers, scribes, illustrators, and all other collaborators. These individuals came together and participated in one or both workshops held in 2021 and 2022. Table 8.1 lists authors by chapter and reviewers in alphabetical order.

The first workshop, the 2021 IEEE CSS Workshop on Control for Societal-Scale Challenges (https://sites.google.com/view/ieee-css-societal-challenges21/home), was held June 4–5, 2021, in the form of a series of virtual, open-to-the-public panels and had more than 900 registrants from over 50 different countries. Table 8.2 lists the panel leaders, panelists, and all authors that contributed to the open call to the community.

The second workshop, the 2022 IEEE CSS Workshop on Control for Societal-Scale Challenges (https://sites.google.com/view/ieee-css-societal-challenges22/home), was an in-person event, held in Stockholm June 17–18, 2022. Experts across different disciplines in control theory participated in moderated panel sessions, where the themes of the roadmap were presented, followed by group discussions, where the attendees discussed the various topics. The invitees were carefully selected for their leadership and expertise in topics that are directly related to the roadmap, and their contributions significantly furthered its creation. Table 8.3 lists the panel chairs, reviewers, scribes, and participants.

Table 8.1: Authors and reviewers of the Control for Societal-Scale Challenges Roadmap: 2030 (in alphabetical order).

CHAPTER	FULL NAME	AFFILIATION
CHAPTER 1		
	Anuradha M. Annaswamy	Massachusetts Institute of Technology
	Karl H. Johansson	KTH Royal Institute of Technology
	George J. Pappas	University of Pennsylvania
CHAPTER 2		
Section 2.A	Saurabh Amin	Massachusetts Institute of Technology
	Margret Bauer	HAW Hamburg
	Aranya Chakrabortty	North Carolina State University, US National Science Foundation
	Fabrizio Dabbene	Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy
	Amritam Das	KTH Royal Institute of Technology
	Masayuki Fujita	Tokyo Institute of Technology
	Mario Garcia-Sanz	Case Western Reserve University
	Dennice Gayme	Johns Hopkins University
	Gabriela Hug	ETH Zürich
	Marija Ilić	Massachusetts Institute of Technology
	Pramod Khargonekar	University of California, Irvine
	Iven Mareels	IBM
	Kevin Moore	Colorado School of Mines
	Lucy Y. Pao	University of Colorado Boulder
	Akshay Rajhans	Mathworks
	Tariq Samad	University of Minnesota
	Jakob Stoustrup	Aalborg University
	Junaid Zafar	Independent Engineering Consultant
Section 2.B	Eyal Dassau	Harvard University
	Francis J. Doyle III	Harvard University
	Sarah Fay	Massachusetts Institute of Technology
	Anette Hosoi	Massachusetts Institute of Technology
	Dalton Jones	Massachusetts Institute of Technology
	Tommaso Menara	University of California, San Diego
	Timothy O'Leary	University of Cambridge
	Marcia K. O'Malley	Rice University
	Philip E. Paré	Purdue University
	Fabio Pasqualetti	University of California, Riverside
	Sri Sarma	Johns Hopkins University
	Rodolphe Sepulchre	University of Cambridge

CHAPTER	FULL NAME	AFFILIATION
Section 2.C	Anuradha M. Annaswamy	Massachusetts Institute of Technology
	Carlos Canudas de Wit	CNRS/GIPSA-Lab
	Masayuki Fujita	Tokyo Institute of Technology
	Hideaki Ishii	Tokyo Institute of Technology
	Karl H. Johansson	KTH Royal Institute of Technology
	Dan Work	Vanderbilt University
Section 2.D	Anuradha M. Annaswamy	Massachusetts Institute of Technology
	Christos G. Cassandras	Boston University
	Robert Shorten	Imperial College London
Section 2.E	Andrew Alleyne	University of Minnesota
	Saurabh Amin	Massachusetts Institute of Technology
	Giacomo Como	Politecnico di Torino
	Johan Eker	Lund University/Ericsson Research
	Frederick Leve	US Air Force Office of Scientific Research
	Nikolai Matni	University of Pennsylvania
	George J. Pappas	University of Pennsylvania
	Thomas Parisini	Imperial College London & University of Trieste
	Henrik Sandberg	KTH Royal Institute of Technology
	Jacquelien Scherpen	University of Groningen
	Emma Tegling	Lund University
	Dawn Tilbury	University of Michigan
Appendix 2.A	Sarah Fay	Massachusetts Institute of Technology
	Anette Hosoi	Massachusetts Institute of Technology
	Dalton Jones	Massachusetts Institute of Technology
Appendix 2.B	Tommaso Menara	University of California, San Diego
	Timothy O'Leary	University of Cambridge
	Fabio Pasqualetti	University of California, Riverside
	Sri Sarma	Johns Hopkins University
	Rodolphe Sepulchre	University of Cambridge

CHAPTER	FULL NAME	AFFILIATION
CHAPTER 3		
Section 3.A	James Anderson	Columbia University
	John Baras	University of Maryland
	Mykel J. Kochenderfer	Stanford University
	Luca Schenato	University of Padova
Section 3.B	Andrew Alleyne	University of Minnesota
	Aranya Chakrabortty	North Carolina State University, US National Science Foundation
	Lucy Y. Pao	University of Colorado Boulder
	Thomas Parisini	Imperial College London & University of Trieste
	Jakob Stoustrup	Aalborg University
Section 3.C	Neda Bagheri	University of Washington
	Mustafa Khammash	ETH Zürich
	Richard Murray	California Institute of Technology
Section 3.D	Dimos V. Dimarogonas	KTH Royal Institute of Technology
	Tor Arne Johansen	Norwegian University of Science and Technology
	Karl H. Johansson	KTH Royal Institute of Technology
	Dawn Tilbury	University of Michigan
CHAPTER 4		
Section 4.A	Florian Dörfler	ETH Zürich
	Na Li	Harvard University
	Nikolai Matni	University of Pennsylvania
	Lucy Y. Pao	University of Colorado Boulder
	Gireeja Ranade	University of California, Berkeley
	Anders Rantzer	Lund University
	Angela Schoellig	TU Munich & University of Toronto
	Claire J. Tomlin	University of California, Berkeley
	Melanie Zeilinger	ETH Zürich
Section 4.B	Aaron D. Ames	California Institute of Technology
	Mykel J. Kochenderfer	Stanford University
	Koushil Sreenath	University of California, Berkeley
	Paulo Tabuada	University of California, Los Angeles
	Claire J. Tomlin	University of California, Berkeley
Section 4.C	Carlos Canudas de Wit	CNRS/GIPSA-Lab
	Sandra Hirche	TU Munich
	Bruno Sinopoli	Washington University in Saint Louis
	Dan Work	Vanderbilt University

CHAPTER	FULL NAME	AFFILIATION
Section 4.D	Aaron D. Ames	California Institute of Technology
	Françoise Lamnabhi-Lagarrigue	CNRS
	Angela Fontan	KTH Royal Institute of Technology
	Sandra Hirche	TU Munich
	Tariq Samad	University of Minnesota
Section 4.E	Aaron D. Ames	California Institute of Technology
	Marie Csete	CEO of OkuloVision
	John C. Doyle	California Institute of Technology
	Na Li	Harvard University
	Steven Low	California Institute of Technology
	Nikolai Matni	University of Pennsylvania
	Anders Rantzer	Lund University
CHAPTER 5		
	Stefano Di Cairano	Mitsubishi Electric Research Laboratories
	Johan Eker	Lund University/Ericsson Research
	Sonia Martinez	University of California, San Diego
	Thomas Parisini	Imperial College London & University of Trieste
	Tariq Samad	University of Minnesota
CHAPTER 6		
	Panos J. Antsaklis	University of Notre Dame
	Christos G. Cassandras	Boston University
	Florian Dörfler	ETH Zürich
	João P. Hespanha	University of California, Santa Barbara
	Gireeja Ranade	University of California, Berkeley
	J. Anthony Rossiter	University of Sheffield
	Steven Brunton	University of Washington, Seattle
CHAPTER 7		
	Anuradha M. Annaswamy	Massachusetts Institute of Technology
	Munther Dahleh	Massachusetts Institute of Technology
	Matin Jafarian	Delft University of Technology
	Françoise Lamnabhi-Lagarrigue	CNRS
REVIEWERS		
	Thierry-Marie Guerra	Université Polytechnique Hauts-de-France
	Mariana Netto	Université Gustave Eiffel
	Benoît Lung	University of Lorraine
	Tiago Roux Oliveira	University of Rio de Janeiro
	Wei Ren	University of California, Riverside
	Antonio Visioli	University of Brescia

Table 8.2: Panel leaders and panelists of the 2021 IEEE CSS Workshop on Control for Societal-Scale Challenges (https://sites.google.com/view/ieee-css-societal-challenges21/home) held virtually on June 4–5, 2021.

ROLE	FULL NAME	AFFILIATION
SAFETY CRITICAL AUTONOMOUS SYSTEMS WITH ML		
Panel Leaders	Angela Schoellig	TU Munich & University of Toronto
	Claire J. Tomlin	University of California, Berkeley
Panelists	Dimos V. Dimarogonas	KTH Royal Institute of Technology
	Anca Dragan	University of California, Berkeley
	Jonathan P. How	Massachusetts Institute of Technology
	Mykel J. Kochenderfer	Stanford University
	Andreas Krause	ETH Zürich
	Sandeep Neema	Vanderbilt University
	Koushil Sreenath	University of California, Berkeley
	Paulo Tabuada	University of California, Los Angeles
	Melanie Zeilinger	ETH Zürich
RESILIENT INFRASTRUCTURE SYSTEMS WITH AI AND IOT		
Panel Leaders	Carlos Canudas de Wit	CNRS/GIPSA-Lab
	Dan Work	Vanderbilt University
Panelists	Saurabh Amin	Massachusetts Institute of Technology
	Bassam Bamieh	University of California, Santa Barbara
	Giacomo Como	Politecnico di Torino
	Mario di Bernardo	University of Naples Federico II
	Marco Pavone	Stanford University
	Lillian Ratliff	University of Washington
	Henrik Sandberg	KTH Royal Institute of Technology
	Jeff Shamma	University of Illinois Urbana-Champaign
EDUCATION AND TRAINING		
Panel Leaders	Christos G. Cassandras	Boston University
	João P. Hespanha	University of California, Santa Barbara
Panelists	Frank Allgöwer	University of Stuttgart
	Panos J. Antsaklis	University of Notre Dame
	Florian Dörfler	ETH Zürich
	Magnus Egerstedt	Georgia Institute of Technology
	Hideaki Ishii	Tokyo Institute of Technology
	Samuel Qing-Shan Jia	Tsinghua University
	Françoise Lamnabhi-Lagarrigue	CNRS
	Richard Murray	California Institute of Technology
	J. Anthony Rossiter	University of Sheffield
	Dawn Tilbury	University of Michigan

ROLE	FULL NAME	AFFILIATION
DECISION-MAKING WITH REAL-TIME AND DISTRIBUTED DATA		
Panel Leaders	Na Li	Harvard University
	Anders Rantzer	Lund University
Panelists	Alexandre Bayen	University of California, Berkeley
	Francesco Bullo	University of California, Santa Barbara
	Munther Dahleh	Massachusetts Institute of Technology
	John C. Doyle	California Institute of Technology
	Ali Jadbabaie	Massachusetts Institute of Technology
	Anthony Kuh	University of Hawaii
	Steven Low	California Institute of Technology
	Nikolai Matni	University of Pennsylvania
	Angelia Nedich	Arizona State University
	Jacquelien Scherpen	University of Groningen
CONTROL WITH HUMAN-IN-THE-LOOP		
Panel Leaders	Aaron D. Ames	California Institute of Technology
	Sandra Hirche	TU Munich
Panelists	Etienne Burdet	Imperial College London
	Joel Burdick	California Institute of Technology
	Ming Cao	University of Groningen
	Francis J. Doyle III	Harvard University
	Elisa Franco	University of California, Los Angeles
	Dana Kulić	Monash University
	Naomi E. Leonard	Princeton University
	Katja Mombaur	Heidelberg University
	Marcia K. O'Malley	Rice University
	Dorsa Sadigh	Stanford University
CONTROL FOR CLIMATE CHANGE MITIGATION AND ADAPTATION		
Panel Leaders	Pramod Khargonekar	University of California, Irvine
	Tariq Samad	University of Minnesota
Panelists	Aranya Chakrabortty	North Carolina State University, US National Science Foundation
	Fabrizio Dabbene	Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy
	Masayuki Fujita	Tokyo Institute of Technology
	Mario Garcia-Sanz	Case Western Reserve University
	Dennice Gayme	Johns Hopkins University
	Gabriela Hug	ETH Zürich
	Marija Ilić	Massachusetts Institute of Technology

ROLE	FULL NAME	AFFILIATION
	Iven Mareels	IBM
	Kevin Moore	Colorado School of Mines
	Lucy Y. Pao	University of Colorado Boulder
	Akshay Rajhans	Mathworks
	Jakob Stoustrup	Aalborg University
CONTRIBUTORS TO OPEN CALL		
	Wen-Hua Chen	Loughborough University
	Jorge Cortes	University of California, San Diego
	Ashwin Dani	University of Connecticut
	Jyotirmoy V. Deshmukh	University of Southern California
	Stefan Engblom	Uppsala University
	Shuo Han	University of Illinois Chicago
	Naira Hovakimyan	University of Illinois Urbana-Champaign
	Clara Ionescu	Ghent University
	Cedric Langbort	University of Illinois Urbana-Champaign
	Mohammad Mamduhi	ETH Zürich
	Sergio Pequito	Uppsala University
	Nicanor Quijano	Los Andes University

Table 8.3: Panel chairs, reviewers, scribes, and participants of the 2022 IEEE CSS Workshop on Control for Societal-Scale Challenges (https://sites.google.com/view/ieee-css-societal-challenges22/home) held in Stockholm, Sweden, on June 17–18, 2022.

ROLE	FULL NAME	AFFILIATION
SESSION 1.1: SOCIETAL DRIVERS		
1.1.A. CLIMATE CHANGE MITIGATION AND ADAPTATION		
Chair	Tariq Samad	University of Minnesota
Reviewer	Dennice Gayme	Johns Hopkins University
Scribe	Aranya Chakrabortty	North Carolina State University, US National Science Foundation
1.1.B. QUALITY OF LIFE		
Chair	Munther Dahleh	Massachusetts Institute of Technology
Reviewer	Philip E. Paré	Purdue University
1.1.C. SMART SOCIETY		
Chair	Hideaki Ishii	Tokyo Institute of Technology
Reviewer	Dan Work	Vanderbilt University
1.1.D. SHARING ECONOMY		
Chair	Christos G. Cassandras	Boston University
Reviewer	Toru Namerikawa	Keio University
1.1.E. GLOBAL SECURITY		
Chair	Saurabh Amin	Massachusetts Institute of Technology
Reviewer	Henrik Sandberg	KTH Royal Institute of Technology
Scribes	Jacquelien Scherpen	University of Groningen
	Dawn Tilbury	University of Michigan
SESSION 1.2: TECHNOLOGICAL TRENDS		
1.2.A. ENGINEERING BIOLOGY		
Chair	Mustafa Khammash	ETH Zürich
Reviewer	Antonis Papachristodoulou	University of Oxford
1.2.B. ROBOTICS AND AUTOMATION		
Chair	Dimos V. Dimarogonas	KTH Royal Institute of Technology
Reviewer	Frederick Leve	US Air Force Office of Scientific Research
1.2.C. ELECTRIFY EVERYTHING		
Chair	Jakob Stoustrup	Aalborg University
Reviewer	Andrew Alleyne	University of Minnesota
Scribes	Anthony Kuh	University of Hawaii
	Lucy Y. Pao	University of Colorado Boulder

ROLE	FULL NAME	AFFILIATION
1.2.C. ELECTRIFY EVERYTHING		
Chair	John Baras	University of Maryland
Reviewer	Mykel J. Kochenderfer	Stanford University
Scribes	Luca Schenato	University of Padova
	James Anderson	Columbia University
SESSION 2.1: METHODOLOGICAL CHALLENGES		
2.1.A. LEARNING AND DATA-DRIVEN CONTROL		
Chair	Anders Rantzer	Lund University
Reviewer	Florian Dörfler	ETH Zürich
2.1.B. SAFETY-CRITICAL SYSTEMS		
Chair	Aaron D. Ames	California Institute of Technology
Reviewer	Kevin Wise	The Boeing Company
2.1.C. RESILIENCE, SECURITY, PRIVACY		
Chair	Dan Work	Vanderbilt University
Reviewer	Bruno Sinopoli	Washington University in Saint Louis
2.1.D. CYBER-PHYSICAL-HUMAN SYSTEMS		
Chair (Plenary)	Aaron D. Ames	California Institute of Technology
Chair (Breakout)	Tariq Samad	University of Minnesota
Reviewer	Françoise Lamnabhi-Lagarrigue	CNRS
2.1.E. ARCHITECTURE AND CONTROL		
Chair	Nikolai Matni	University of Pennsylvania
Reviewer	Frank Allgöwer	University of Stuttgart
SESSION 2.2: VALIDATION, TECHNOLOGY TRANSFER, AND EDUCATION		
2.2.A. TECHNOLOGY VALIDATION		
Chair	Johan Eker	Lund University/Ericsson Research
Reviewer	Aranya Chakrabortty	North Carolina State University, US National Science Foundation
2.2.B. EDUCATION		
Chair	Christos G. Cassandras	Boston University
Reviewer	Claudio De Persis	University of Groningen
2.1.C. TECHNOLOGY TRANSITION		
Chair	Thomas Parisini	Imperial College London & University of Trieste
Reviewer	John Lygeros	ETH Zürich

ROLE	FULL NAME	AFFILIATION
2.1.D. ETHICS AND FAIRNESS		
Chair	Anuradha M. Annaswamy	Massachusetts Institute of Technology
Reviewer	Jing Sun	University of Michigan
SESSION 2.3: NEXT STEPS		
2.3.A PUBLICATION		
Chair	Anuradha M. Annaswamy	Massachusetts Institute of Technology
2.3.B OUTREACH AND MEDIA		
Chair	Karl H. Johansson	KTH Royal Institute of Technology
2.3.C IMPACT		
Chair	George J. Pappas	University of Pennsylvania
OTHER PARTICIPANTS		
	Hamsa Balakrishnan	Massachusetts Institute of Technology
	Paul Bogdan	University of Southern California
	Laurent Burlion	Rutgers University NJ
	Dimitar Filev	Ford Research
	Wilhelmus P.M.H. Heemels	Eindhoven University of Technology
	Linnea Sundling	Digital Futures
	Jana Tumova	KTH Royal Institute of Technology

Glossary

Glossary

adaptation The process by which a control system makes self-adjustments, in the form of parameter estimation, and uses the estimates to make relevant decisions

agrivoltaics The simultaneous use of land areas for agriculture and photovoltaic power generation

architecture A framework that interconnects one or more elements or components so as to exhibit advanced functionalities

artificial intelligence The development of machines that mimic human intelligence (e.g., languages, perception, decision)

artificial pancreas An artificial system that mimics how a healthy pancreas controls blood glucose

artificial photosynthesis A chemical process that mimics the natural process of photosynthesis

automated anesthesia A system in which a controller adjusts the flow of an anesthetic drug to achieve the desired brain state

autonomy The process by which self-governing decisions are made in a system based on data, self-monitoring, and feedback

autotuning The process by which parameters are self-adjusted based on a criterion to improve performance; typically used for tuning PID gains

benchmarks Examples or cases used to evaluate and compare methods and best practices

bioengineering The application of engineering principles to biological and biomedical applications

biological networks Interconnections of biological entities such as genetic or metabolic networks

biomolecular control systems Control systems that regulate biological systems

carbon capture The trapping of carbon dioxide emissions and their storage in deep underground geological formations

carbon markets Systems in which carbon credits are bought and sold

cardiac-assist devices Electromechanical devices for assisting circulation

climate change Long-term shifts in temperatures and weather patterns, mainly driven over the last 200 years by human activities

climate change adaptation and mitigation How to protect people and places by making them less vulnerable to the impacts of climate change and how to reduce or prevent emission of greenhouse gases

control co-design Integrating control design with other system design aspects (hardware, software, etc.)

control engineering The design and analysis of feedback systems that ensure safety, stability, and robustness properties

control systems Analysis and synthesis of the process by which a physical system is controlled using data obtained using measurements, computations, and other exogenous information to lead to system properties such as stability, optimality, robustness, and tracking

control systems theory The theoretical framework that provides analytical underpinnings for a physical system that is controlled using data obtained using measurements, computations, and other exogenous information; this framework allows the derivation of theoretical guarantees of desired system properties such as stability, optimality, robustness, and tracking

controller/architecture co-design The joint design of a control system and the infrastructure (e.g., communication, actuation, sensing, and computation) needed to implement it

cooperative control Control of a system with multiple agents aiming for a common goal critical infrastructure systems

critical infrastructures The basic physical and organizational structures and facilities critical for the operation of a society or enterprise

critical infrastructure systems Systems essential for the functioning of a society or economy

cyber-physical-human systems An interconnection of physical systems, computational devices, and humans, with varied and complex interactions

decarbonization The process of reducing carbon dioxide emissions

decision support systems Information systems that support decision-making activities

deep brain stimulation A neurosurgical procedure sending electrical impulses through implanted electrodes to specific targets in the brain

demand response The opportunity for consumers to shift their electricity usage during peak periods in response to time-varying rates

distributed control Control algorithms distributed across systems without any central coordination

distributed learning Learning over distributed or networked systems where data are not centralized

distributed optimization Optimization performed by a set of interconnected agents updating their variables to optimize a global cost

ecosystem engagement The process of getting stakeholders and networks of peers to interact and share in value creation

education modules A compilation of lectures with supporting tools of videos and quizzes that is a comprehensive exposition of a particular method, concept, or architecture

electricity markets Systems that facilitate the participants to sell and buy, and thereby exchange, electricity

energy efficient computing Handling limiting energy resources in computing systems, from large data centers to small embedded systems

energy efficient machine learning Handling limiting energy resources when training machine learning models

environmental monitoring The process of observing the environment and characterizing its quality

exploration and exploitation Exploration is the process of collecting more data and improving the underlying control systems; exploitation is the process of making a decision based on the existing data with the expectation that it may be optimal

fairness Ensuring impartial access and impact of engineering systems or approaches

geoengineering The engineering of carbon dioxide removal and solar radiation management applied at a planetary scale (also called climate engineering)

hierarchical control A type of layering in control architectures in which “higher” layers (e.g., planning) send commands to “lower” layers (e.g., control)

human-in-the-loop Systems that interact with humans in a feedback loop

human-in-the-loop control A control system where the feedback includes a combination of sensors, computation, communication, control, and humans, where humans can provide any of the functionalities in the combination

human-robot interaction Robots interacting with or operating in the proximity of humans

innovation and technology transfer The process of bringing research results to the market or general society

interdependent infrastructures A feature of modern infrastructures linked by multiple connections, feedback loops, and other dependencies (e.g., interconnection of the power grid with transport and information networks)

irrigation management The management of water distribution resources for agricultural crops

layering A ubiquitous design pattern in complex engineered systems that decomposes a system's behavior into well-defined modules (e.g., planning and control)

learning The process by which uncertainties in a system are learned so as to improve the overall performance

machine learning Approaches that develop prediction models from input–output data

mechanism design The process by which incentives or other mechanisms are designed so as to meet a certain objective

medical internet of things Medical devices and applications that connect to healthcare IT systems

metabolic engineering Engineering approaches for optimizing genetic and regulatory processes within cells

microgrids A local electric grid able to act in an island mode independent of the main grid

modularity of biological components Engineering approaches for understanding properties of compositions of biological components given properties of the individual components

multi-rate control A control system in which different layers operate at different timescales

multiscale modeling When multiple models at different scales are used to simultaneously de- scribe a system

negative emissions technologies Technologies for the absorption and storage of greenhouse gases

network resilience The ability of a network to maintain an acceptable level of service despite faults and other challenges to normal operation

neural networks Brain-inspired mathematical networks of neuron models

optimization (of models/parameters) Mathematical and computational approaches for choosing optimal parameters given constraints

perception-based planning and control Planning and control approaches that heavily rely on vision and perception as a sensor

personalized medicine The use of an individual's genetic profile for prevention, diagnosis, and treatment of disease

physics-informed learning Machine learning approaches that are informed or constrained by physical laws

quality of life An individual's perception of their position in life in the context of the culture and value systems in which they live

reinforcement learning A sequential learning approach for a reward-maximizing agent interacting with an unknown environment

renewable fuels Energy resources that replenish naturally over time

resilience The property of a system to withstand large anomalies and to restore itself to at least a reduced level of functionality

restoration The property of a system to return as closely possible to its original state following an anomaly

risk assessment Identifying and analyzing the potential for negative events and judging tolerance for such events

robotics The design and programming of robots for personal or automation tasks

robust AI Learning that is robust to various data perturbations

robust control Methodology for controlling systems with uncertain models

robustness The property of a system by which it can withstand perturbations without any adaptation

safety A fundamental property of a system wherein there is a high emphasis on satisfactory behavior with stringent specifications

safety control Control methodology that ensures safety of system behavior

safety-critical A property of a system where any failure may result in a significant impact such as fatalities and injuries

safety-preserving resilience The property of a system to withstand large anomalies and return as closely possible to its original state while ensuring system safety

sharing economy Systems built around the sharing of resources by renting and borrowing rather than buying and owning

smart agriculture The integration of digital technologies and robots for sustainable agricultural productivity

smart mobility Intelligent networks integrating various modes of transport and technology

smart power grid An electric power grid with smart sensing, distribution, and adaptation in response to varying energy supply and demand

smart society A human-centered society that aims to solve societal problems through systems that highly integrate cyber and physical space

smart water The integration of advanced technologies for the management and operation of water distribution networks

social contract An implicit agreement among the members of a society defining the rights and duties of each member

socially responsible automation Automation that incorporates social considerations in design (e.g., safety, privacy, fairness)

societal-scale cyber-physical systems Systems that integrate cyber and physical parts through computation and communication to add new capabilities to physical systems of great societal importance

stochastic modeling Mathematical and computational biological models that capture probabilistic behavior

synthetic biology Engineering approaches for building functional systems from biological substrates

systems biology Engineering approaches for understanding biological systems

technological convergence The process of originally unrelated technologies becoming more closely related or even unified as they develop

testbeds Platforms for testing new scientific theories, methods, tools, and technologies

transactive control Distributed control strategy based on the engagement of self-interested agents interacting through a dynamic market mechanism, such as demand response for power systems

transdisciplinary research Research that integrates knowledge across academic disciplines and with nonacademic stakeholders to address societal challenges

uncertainty modeling Modeling what we do not know about the underlying systems

vehicle-to-grid A system enabling electric vehicles to sell energy to the electric grid

Index

Index