Pragmatics: Discourse Analysis

J&M’s Chapter 21
Discourse

• Natural languages consist of collocated and related sentences

• Such a group of sentences is referred to as discourse
Discourse Types

- Traditional distinctions:
  - Spoken/written discourse
  - Monologue/dialogue
- Dialogues consist of asking questions, giving answers, etc.
- New discourse types: SMS, chatting, e-mail...
**Reference Resolution**

*John* went to Bill‘s car dealership to check out a *BMW*. *He* looked at *it* for an hour.

- *John* and *he*, which denote a person named *John*, are called *referring expression*. *John* is their **referent**.

- Two referring expression that refer to the same entity are said to **co-refer** (e.g., *John* and *he*)
Reference Resolution

*John* went to Bill‘s car dealership to check out a *BMW*. *He* looked at *it* for an hour.

- *John* is called to be an *antecedent* of and *he*.

- Reference to an entity that has been *previously* introduced into the discourse is said to be *anaphora*, and

- the referring expression is said to be *anaphoric* (e.g., *he* and *it* are anaphoric)
Reference Resolution

Before *he* bought *it*, *John* checked over the *BMW* very carefully.

- Reference that is mentioned before its referent is said to be *cataphora*, and

- the referring expression is said to be *cataphoric* (e.g., *he* and *it*)
Discourse Model

- One cannot simply say *it* or *the BMW*, if the hearer has no prior knowledge of the subject

- The hearer’s mental model of the ongoing discourse is called the **discourse model**
Discourse Model

• There are two fundamental operations applied to the discourse model:

• When a referent is first mentioned in a discourse, a representation for it is evoked into the model.

• Upon subsequent mention, this representation is accessed from the model.
Reference Operations and Relationships

Refer (access)

Refer (evoke)

JOHN

Co-refer

HE
Referring Expressions

• Some of different types of referring expressions:
  – Indefinite noun phrases
  – Definite noun phrases
  – Pronouns
  – Demonstratives
  – One-anaphora
Indefinite noun phrases:

- Introduce entities that are *new* to the hearer

- usually marked with *a, an, some*, or even *this*

E.g.,

I saw *a BMW* today.

I saw *this awesome BMW* today.
Definite noun phrases

• Refer to an entity identifiable to the hearer, because:

• It has been mentioned in the discourse context and thus represented in the discourse model
  I saw a BMW today. The BMW was awesome.

• It is contained in the hearer’s set of beliefs about the world
  The Formula one is the most popular car race in Europe

• The uniqueness is implied by the description
  The fastest car in the Formula one was a Ferrari
Pronouns

- Are another form of definite references

- But, they require that the referent to have a high degree of salience

- They usually refer to entities that are no further than one or two sentences back,

- Whereas definite noun phrase can refer further back
Pronouns (Cont.)

- Usually refer to recent referents:

  E.g.,
  1. John went to Bob’s party, and parked next to a beautiful BMW.
  2. He went inside and talked to Bob for more than an hour.
  3. Bob told him that he recently got engaged.
  4. ?? He also said that he bought it yesterday.
  5. He also said that he bought the BMW yesterday.
Demonstratives

• Behave differently from simple definite pronouns such as *it*

• They can appear either alone, or as determiners

• John shows Bob a BMW and a Mazda. I like *this* more than *that*. 
One-anaphora

• Blends properties of definite and indefinite reference.

• May evoke a new entry into the discourse model, but

• It is dependent on an existing referent of this new entry

I saw more than 6 BMWs today. Now I want one (i.e., one of them).

• Them refers to a plural referent, and one selects a member from this set.
Complex Referring Expressions

- These referring expressions complicate the reference resolution problem
  - Inferrables
  - Discontinuous sets
  - Generics
Inferrables

• An inferrable does not refer to an entity that has been explicitly evoked in the text,

• Instead, it refers to one that is inferentially related to an evoked entity

I almost bought a BMW today, but a door had a dent and the engine seemed noisy.
Discontinuous Sets

• Plural references such as *they* and *them* may refer to sets of entities that are evoked together

John and Mary love their BMWs. *They* drive *them* all the time.
Discontinuous Sets

- Plural references may also refer to sets of discontinuous entities

John has a BMW, and Mary has a Mazda. They drive them all the time.
Generics

- Existence of a *generic reference* makes the problem of reference resolution even more complicated.

I saw more than 6 BMWs today. *They* are coolest cars.

- Here *they* refers to BMWs in general, and not the 6 BMWs mentioned in the first sentence.
Syntactic/Semantic Constraints on Co-reference

• We need a way of filtering the set of possible referents, using hard-and-fast constraints

• Some of such constraints include:
  – Number Agreement
  – Person and Case Agreement
  – Gender Agreement
  – Syntactic Constraints
  – Selectional Restrictions
Number Agreement

• Referring expressions and their referents must agree in number:

  John has a new BMW. *It* is red.
  John has three new BMWs. *They* are red.

* John has a new BMW. *They* are red.
* John has three new BMWs. *It* is red.
Person and Case Agreement

- English distinguishes between three forms of person: first, second, and third.

You and I have BMWs. *We* love *them*.
John and Mary have BMWs. *They* love *them*.

* John and Mary have BMWs. *We* love *them*.
* You and I have BMWs. *They* love *them*. 
Gender Agreement

• Referents also must agree with the gender specified by the referring expression.

• English third person pronouns distinguish between male, female, and non-personal

John has a BMW.

He is attractive. (he = John, not the BMW).

It is awesome (it = BMW, not John)
Syntactic Constraints

- Reference relations may also be constrained by the syntactic relationships.

- Reflexive pronoun co-refers with *the subject of the most immediate clause* that contains it, whereas a non-reflexive cannot co-refer with this subject.

  John bought *himself* a new BMW [himself=John]
  John bought *him* a new BMW [him≠John]
Syntactic Constraints

• The rule about reflexive pronouns applies only for the subject of the most immediate clause.

John said that Bill bought him a new BMW
[him≠Bill]

John said that Bill bought himself a new BMW
[himself=Bill]

He said that he bought John a new BMW
[He ≠ John; he≠John]
Selectional Restrictions

• The *selectional restrictions* that a verb impose on its arguments may be used for eliminating referents

John parked his BMW in the garage. He had driven *it* for hours.
Selectional Restrictions

- Selectional restrictions can be violated in the case of metaphor

John bought a new BMW. *It* drinks gasoline like you would not believe.
Preference in Pronoun Interpretation

• The majority of work on reference resolution algorithms has focused on *pronoun interpretation*

• Some of the preferences used in pronoun interpretation are:

• *Recency, Grammatical Role, Repeated Mention, Parallelism,* and *Verb Semantics*
Preferences in PI (*Recency*)

- Entities introduced in recent utterances are more *salient* than those introduced further back.

  John has a **BMW**. Bill has a **Mazda**. Mary likes to drive *it*. 
Preferences in PI (Grammatical Role)

• A salience hierarchy of entities can be formed by the grammatical position their mentions:

• Entities mentioned in subject position are more salient than those in object positions,

Bill went to the BMW dealership with John. He bought a BMW.
Bill and John went to the BMW dealership. He bought a BMW. [he=?]
Preferences in PI (Repeated Mention)

- Entities that have been focused on in the prior discourse are more likely to be focused on in subsequent discourse,
- and hence references to such entities are more likely to be pronoun

John needed a car to get to his new job. He decided that he wanted something sporty. Bill went to the BMW dealership with John. He bought a BMW.
Preferences in PI (Parallelism)

• Strong preferences are induced by parallelism effects.

Mary went with Sue to the BMW dealership. Sally went with her to the Mazda Dealership.

• Grammatical role ranks Mary as more salient than Sue, and there is no semantic reason that Mary cannot be the referent.
Preferences in PI (Verb Semantics)

• Certain verbs place a semantically-oriented emphasis on one of their argument positions

  John telephoned Bill. He lost the manual on BMW.
  John criticized Bill. He lost the manual on BMW.

• Implicit causality of a verb:
  Implicit cause of “criticizing” event is its object.
  Implicit cause of “telephoning” event is its subject.
Algorithms for Pronoun Resolution

• Lappin and Leass (1994) propose a simple algorithm that considers many of those preferences

• By a simple weighing scheme that integrates recency and syntactic based preferences

• The algorithm performs two types of operation:
  – discourse model update
  – pronoun resolution
**Lappin and Leass Algorithm for PR**

- When a noun phrase that evokes a new entity is encountered, a *saliency value* is computed for it.

- The salience value is calculated as the sum of the weights assigned by a set of salience factors.
<table>
<thead>
<tr>
<th>Salience Factors in Lappin &amp; Leass system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence recency</td>
</tr>
<tr>
<td>Subject emphasis</td>
</tr>
<tr>
<td>Existential emphasis</td>
</tr>
<tr>
<td>Direct object emphasis</td>
</tr>
<tr>
<td>Indirect object emphasis</td>
</tr>
<tr>
<td>Non-adverbial emphasis</td>
</tr>
<tr>
<td>Head noun emphasis</td>
</tr>
</tbody>
</table>
An Example

1. *A BMW* is parked in the lot. (subject)
2. There is *a BMW* parked in the lot. (existential predicate nominal)
3. John parked *a BMW* in the lot. (object)
4. John gave *his BMW* a bath. (indirect object)
5. Inside *his BMW*, John showed Susan his new CD player. (separated adverbial PP)
6. The owner’s manual for *a BMW* is on John’s desk. (non head noun)
Lappin and Leass Algorithm for PR

• Each time a new sentence is processed, weights of entities in the discourse model are cut into half.
• Referents mentioned in the current sentence get +100 for recency.
• Referents in separated adverbial PPs (i.e., those separated by punctuation) is penalized by adding +50 to other positions.
• Referents which are embedded in larger noun phrases are penalized by adding +80 to other referents.
Calculating Salience Value

• Several noun phrases may refer to the same referent, each having a different salience value

• Each referent is associated to an equivalence class that contains all the noun phrases that refer to it.

• The weight that a salience factor assigns to a referent is the highest of weights it assigns to the members of its equivalence class.
**Calculating Salience Value**

- The salience weight of a referent is calculated by summing the weight of relevant salience factors.
- The scope of a salience factor is a sentence:
  - If a potential referent is mentioned in the current sentence as well as the previous one, the sentence recency weight will be factored for each.
  - But, if the same referent is mentioned more than once in the same sentence, the weight will be counted only once.
- Thus multiple mentions of a referent in the prior discourse can potentially increase its salience.
Salience Factors in Lappin & Leass system

- There are two more salience weights:
  - Role Parallelism: 35
  - Cataphora: -175

- But, these two cannot be calculated independently of the pronoun, and thus cannot be calculated during the discourse model update.

- We will use the term *initial salience value* for the weight of a given referent before these factors are applied, and the term *final salience value* for after they have been applied.
**Resolving Pronouns Process**

1. Collect the potential referents (up to 4 sentence back).
2. Remove potential referents that don’t agree in number or gender with the pronoun.
3. Remove potential referents that don’t pass syntactic co-reference constraints.
4. Compute the total salience value of the referent by adding applicable values to the existing salience value computed during discourse model update.
5. Select the referent with the highest salience value. (in the case of tie, select the closest referent.)
An Example

1. John saw a beautiful BMW at the dealership.
2. He showed it to Bob.
3. He bought it.
An Example

1. John saw a beautiful BMW at the dealership.

<table>
<thead>
<tr>
<th></th>
<th>Rec</th>
<th>Subj</th>
<th>Exis</th>
<th>Obj</th>
<th>Ind-Obj</th>
<th>Non-Ad</th>
<th>Head</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>100</td>
<td>80</td>
<td></td>
<td></td>
<td>50</td>
<td>80</td>
<td></td>
<td>310</td>
</tr>
<tr>
<td>BMW</td>
<td>100</td>
<td></td>
<td>50</td>
<td></td>
<td>50</td>
<td>80</td>
<td></td>
<td>280</td>
</tr>
<tr>
<td>dealership</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>80</td>
<td></td>
<td>230</td>
</tr>
</tbody>
</table>
**An Example**

1. **John** saw **a beautiful BMW** at **the dealership**.

<table>
<thead>
<tr>
<th>Referent</th>
<th>Phrases (equivalence classes)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>{John}</td>
<td>310</td>
</tr>
<tr>
<td>BMW</td>
<td>{a beautiful BMW}</td>
<td>280</td>
</tr>
<tr>
<td>dealership</td>
<td>{the dealership}</td>
<td>230</td>
</tr>
</tbody>
</table>
An Example

1. John saw a beautiful BMW at the dealership.
2. He showed it to Bob. (Values are cut into half)

<table>
<thead>
<tr>
<th>Referent</th>
<th>Phrases (equivalence classes)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>{John}</td>
<td>155</td>
</tr>
<tr>
<td>BMW</td>
<td>{a beautiful BMW}</td>
<td>140</td>
</tr>
<tr>
<td>dealership</td>
<td>{the dealership}</td>
<td>115</td>
</tr>
</tbody>
</table>
An Example

1. John saw a beautiful BMW at the dealership.
2. He showed it to Bob.

(the only referent candidate for “He” is John)

<table>
<thead>
<tr>
<th>Referent</th>
<th>Phrases (equivalence classes)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>{John, he_1}</td>
<td>155+310 = 465</td>
</tr>
<tr>
<td>BMW</td>
<td>{a beautiful BMW}</td>
<td>140</td>
</tr>
<tr>
<td>dealership</td>
<td>{the dealership}</td>
<td>115</td>
</tr>
</tbody>
</table>
An Example

1. John saw a beautiful BMW at the dealership.
2. He showed it to Bob.

(there are two candidates for “it”, but parallelism support BMW)

<table>
<thead>
<tr>
<th>Referent</th>
<th>Phrases (equivalence classes)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>{John, he_1}</td>
<td>465</td>
</tr>
<tr>
<td>BMW</td>
<td>{a beautiful BMW, it_1}</td>
<td>140 + 280 = 420</td>
</tr>
<tr>
<td>dealership</td>
<td>{the dealership}</td>
<td>115</td>
</tr>
</tbody>
</table>
**An Example**

1. John saw a beautiful BMW at the dealership.
2. *He* showed *it* to **Bob**. (100+40+50+80 for Bob)

<table>
<thead>
<tr>
<th>Referent</th>
<th>Phrases (equivalence classes)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>{John, he$_1$}</td>
<td>465</td>
</tr>
<tr>
<td>BMW</td>
<td>{a beautiful BMW, it$_1$}</td>
<td>420</td>
</tr>
<tr>
<td>Bob</td>
<td>{Bob}</td>
<td>270</td>
</tr>
<tr>
<td>dealership</td>
<td>{the dealership}</td>
<td>115</td>
</tr>
</tbody>
</table>
An Example

1. John saw a beautiful BMW at the dealership.
2. *He* showed *it* to Bob.
3. *He* bought *it*. (The values are cut into half)

<table>
<thead>
<tr>
<th>Referent</th>
<th>Phrases (equivalence classes)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>{John, he₁}</td>
<td>232.5</td>
</tr>
<tr>
<td>BMW</td>
<td>{a beautiful BMW, it₁}</td>
<td>210</td>
</tr>
<tr>
<td>Bob</td>
<td>{Bob}</td>
<td>135</td>
</tr>
<tr>
<td>dealership</td>
<td>{the dealership}</td>
<td>57.5</td>
</tr>
</tbody>
</table>
Centering Theory

- Grosz et al. (1995) proposed a centering theory with an explicit representation of a discourse model.

- They claim that there is a single entity being “centered” on at any given point in the discourse.
Centering Algorithm Definitions

- There are two main representations tracked in the discourse model (let $U_n$ and $U_{n+1}$ be two adjacent utterances):
  1. The backward looking center of $U_n$, denoted as $C_b(U_n)$, represents the entity currently being focused on in the discourse after $U_n$ is interpreted.
  2. The forward looking centers of $U_n$, denoted as $C_f(U_n)$, forms an ordered list containing the entities mentioned in $U_n$, all of which could serve as the $C_b$ of the following utterance (i.e., $U_{n+1}$).
Centering Algorithm Definitions

- \( C_f(U_n) \) is ordered based on the grammar role hierarchy encoded by weights in Lappin and Leass algorithm:

  Subject > existential predicate nominal > object > indirect object > separated adverbial PP

- However, there are no numerical weights attached to the entities.
Centering Algorithm Rules

- There are two rules used by the centering algorithm:
  1. If any element of $C_f(U_n)$ is realized by a pronoun in utterance $U_{n+1}$, then $C_b(U_{n+1})$ must be realized as a pronoun, too.
  2. Transition states are ordered. *Continue* is preferred to *Retain* is preferred to *Smooth-Shift* is preferred to *Rough-Shift*. 
# Transition Types

<table>
<thead>
<tr>
<th>$C_b(U_{n+1}) = C_b(U_n)$ or undefined $C_b(U_n)$</th>
<th>$C_b(U_{n+1}) \neq C_b(U_n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_b(U_{n+1}) = C_p(U_{n+1})$</td>
<td>1) Continue</td>
</tr>
<tr>
<td>$C_b(U_{n+1}) \neq C_p(U_{n+1})$</td>
<td>2) Retain</td>
</tr>
</tbody>
</table>

$C_p$ (preferred center) is a short hand for the highest-ranked forward-looking center.
Centering Algorithm

1. Generate possible $C_b$-$C_f$ combinations for each possible set of reference assignments.

2. Filter by constraints, e.g., syntactic coreference constraints, selectional restrictions, centering rules and constraints.

3. Rank by transition orderings.
Centering Algorithm Example

John saw a beautiful BMW at the dealership ($U_1$).

*He* showed *it* to Bob ($U_2$).

*He* bought *it* ($U_3$).

- Using grammatical role hierarchy to order $C_f$ of $U_1$:
  $C_f(U_1)$: {John, BMW, Dealership}
  $C_p(U_1)$: John
  $C_b(U_1)$: undefined
Centering Algorithm Example

John saw a beautiful BMW at the dealership ($U_1$).
He showed it to Bob ($U_2$).
He bought it ($U_3$).

- Assuming “it” refers to the BMW:
  $C_f(U_2)$: \{John, BMW, Bob\}
  $C_p(U_2)$: John
  $C_b(U_2)$: John

Result: Continue ($C_p(U_2) = C_b(U_2)$; $C_b(U_1)$ is undefined)
Centering Algorithm Example

John saw a beautiful BMW at the dealership ($U_1$).
He showed it to Bob ($U_2$).
He bought it ($U_3$).

- Assuming “it” refers to the dealership:
  $C_f(U_2)$: {John, dealership, Bob}
  $C_p(U_2)$: John
  $C_b(U_2)$: John

Result: Continue ($C_p(U_2) = C_b(U_2)$; $C_b(U_1)$ is undefined)
Centering Algorithm Example

John saw a beautiful BMW at the dealership ($U_1$). He showed it to Bob ($U_2$). He bought it ($U_3$).

- Since both possibilities results in Continue, the algorithm doesn’t say which to accept.
- Let’s assume ties are broken in terms of the ordering on previous $C_f$ list.
- So, it refers to the BMW
Centering Algorithm Example

John saw a beautiful BMW at the dealership ($U_1$).
*He* showed *it* to Bob ($U_2$).
*He* bought *it* ($U_3$).

- Assuming “he“ refers to John:
  $C_f(U_3)$: {John, BMW}
  $C_p(U_3)$: John
  $C_b(U_3)$: John

Result: Continue ($C_p(U_3)=C_b(U_3)=C_b(U_2)$)
Centering Algorithm Example

John saw a beautiful BMW at the dealership \( (U_1) \).

*He* showed *it* to Bob \( (U_2) \).

*He* bought *it* \( (U_3) \).

- Assuming “he“ refers to Bob:

  \[ C_f(U_3): \{\text{Bob, BMW}\} \]

  \[ C_p(U_3): \text{Bob} \]

  \[ C_b(U_3): \text{Bob} \]

Result: Smooth-Shift \( (C_p(U_3) = C_b(U_3); C_b(U_3) \neq C_b(U_2)) \)