

Computational Discourse

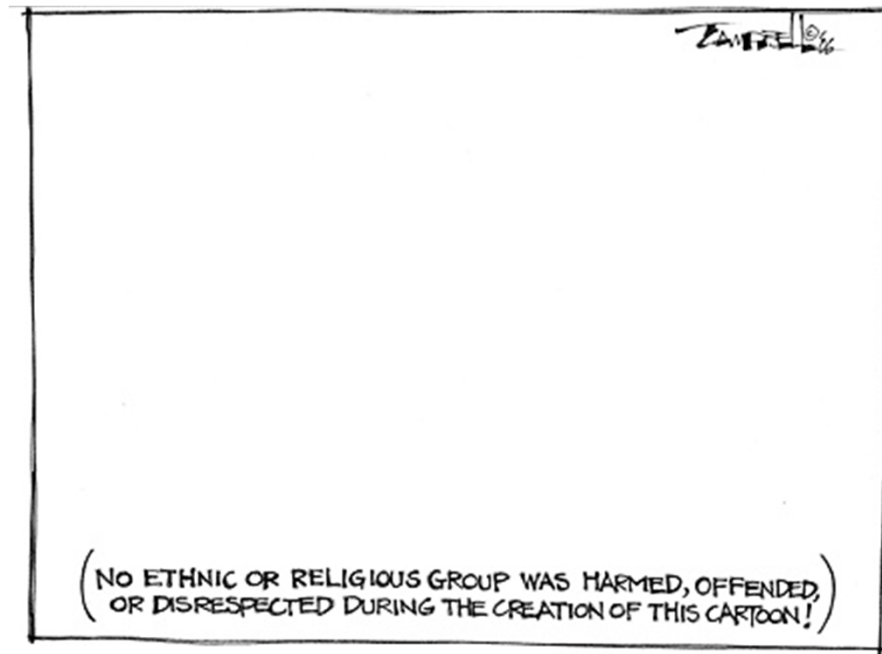
Section 21.3 – 21.5 in Textbook

hannes@ru.is

T-725-MALV

Discourse Context

- We interpret all meaning in some **context**.
Strictly speaking our context covers our entire existence up to the moment of interpretation!



Discourse Context

- To simplify let's divide the context into 3 categories:
 - **Cognitive Context** (Knowledge) Common sense world knowledge as well as special domain knowledge
 - **Situational Context** (Environment) The physical environment of the exchange, including participants and any objects
 - **Textual Context** (Co-Text) Whatever has been said so far in the exchange.

Discourse Model

- Every participant in the communication builds a **discourse model** that **keeps track** of the context and helps them with interpretation.
- For example, it can help them **interpret anaphoras**, where the text contains pointers into the context for staying concise.
- Anaphoras include personal ("she") or demonstrative **pronouns** ("that").

Anaphora Examples

- **"John lit a pipe. He pondered."**
Here "John" is an **antecedent** and "He" an **anaphor**. Keeping track of the textual context helped with interpretation here.
- **"This is madness!" (At construction site)**
Here "This" is an anaphor that refers (likely!) to the ongoing construction activity. Keeping track of the situation context helped with interpretation here.

Discourse Entity

- **What is discussed** in a discourse, requires some representation in the discourse model. This representation is called a **Discourse Entity** or a Discourse Referent
 - DEs are objects, concepts or whatever else the text can be about
 - **Referring Expressions** in the text are **pointers** to these **Referents**

Referring Expressions

- The simplest way to update our discourse model is to **keep track of all referring expressions** in the text.
- When we find a new (first-time) reference, we **create a new DE** in our model.
- Later references to the same thing connect to the **existing DE** in the model.
- All references to the same DE are said to **corefer**.

DEs and NPs

- If one only considers physical/concrete things, then the **referring expressions correspond to all the noun phrases** in the text.
- Then it is relatively easy to construct DEs inside the model from noun phrases.
- The **difficulty** is in realizing when a noun phrase is a **referring expression to an existing DE**, i.e. a coreference is found. There should only be one DE per actual object!

Finding Coreference

- **Grammatical** Cues of Coreference:
"I got a new car. It is gray."
- **Semantic** Cues of Coreference:
"I got a new dog. The creature is so loud!"
- **Tricky**: These can conflict - what's correct?
"Bílstjórinn ók jeppanum útaf. Hann missti bílprófið."
(The driver drove the jeep off the edge of the road.
he/it lost the license)

Types of Referring Expressions

- Indefinite Noun Phrases (new DE)

"In walked **a ghost**"

- Definite Noun Phrases (DE already exist)

"**The ghost** howled!"

- Pronouns (DE already exist, nearby) unless cataphora!

"Then **it** ran across the room, bumping into a chairs"

- Demonstratives (Either new or existing DE)

"**This poor thing** didn't see where it was going"

- Names (Either new or existing DE)

"You gotta love **Casper**..."

Information Status

- How different referential forms are used to provide new or old information is called **information status** or **information structure**
- **Givenness hierarchy (Gundel):**
in focus > activated > familiar > uniquely identifiable > referential > type identifiable

Information Status

- **Accessibility scale (Ariel):**

Full name < long definite description < short definite description < last name < first name < distal demonstrative < proximate demonstrative < NP < stressed pronoun < unstressed pronoun

- Accessibility scale correlates with length. Reference to less accessible DEs tend to be longer.

Tricky Referring Expressions

- **Inferrables**

"I almost bought a 1961 Ford today, but a door had a dent and the engine seemed noisy"

- **Generics**

"These days you have to be careful"

- **Non-referential uses**

"It is raining" (pleonastic case)

Filtering Potential Referents

- **Number** Agreement (e.g. she vs. they)
- **Person** Agreement (e.g. she vs. you)
- **Gender** Agreement (e.g. she vs. he)
- Binding Theory **Constraints**
(syntactic relationship, e.g. for reflexives vs. non-reflexives, himself vs. him)

Preferred Potential Referents

- **Recency** (closer is better)
- **Grammatical Role** (subject over object)
- **Repeated Mention** (the focus of the story)
- **Parallelism** (location in parallel sentence)
- **Verb Semantics** (cause may be object)
- **Selectional Restriction** (semantic knowledge)

Many Algorithms

- **The Hobbs Algorithm**

Preferences: Grammatical role, recency

- **Centering Algorithm**

Preferences: Grammatical role, recency, repeated mention

- **Supervised Machine Learning**

Preferences: Based on features trained