Natural Language Processing

Lecture 21

Today's Outline

Review

- representing and interpreting dialog acts
- dialogue structure and coherence

Conversational Agents

Generation

Review: Try Conversational Agents!

Academic

- CMU Movieline: 412 268 1185
- CMU Communicator (travel): 412 268-1084 or -5144
- MIT Jupiter (weather): 1-888-573-8255

Commercial

- Nuance: 1-888-NUANCE-8 (demo)
- Speechworks: United Flight Info 1.800.824.6200
- Tellme: 1 800 555 tell

More Conversational Agents!

TrindiKit

http://www.ling.gu.se/projekt/trindi/trindikit/

- GoDiS and the Dialogue Move Engine
  - http://www.ling.gu.se/~sl/DME/
- Multiple Inference-based Dialogue Analysis System
  - http://www.coli.uni-sb.de/~bos/trindi/midas.html

Review: Dialogue Acts

DAMSL: A recent computational, expanded, hierarchical dialogue act tagging scheme

- forward looking level
  - draws from Searle/Austin speech acts
- backward looking level
  - draws from grounding, adjacency pairs, ...

A continuum of dialogue act tagging algorithms

- plan inference model
- multiple sources of knowledge (e.g., words, prosody, structure) provide cues

Leads to <u>intentional</u> approaches to discourse structure and coherence

What is NLG?

Beyond canned text

Beyond template filling (TOOT, Eliza, etc.)

Mapping meaning to text

- inverse of NLU
- meaning representations can vary widely

NLG vs. NLU

- in some sense, NLG is "easier" than NLU
- same issues, but often different mechanisms

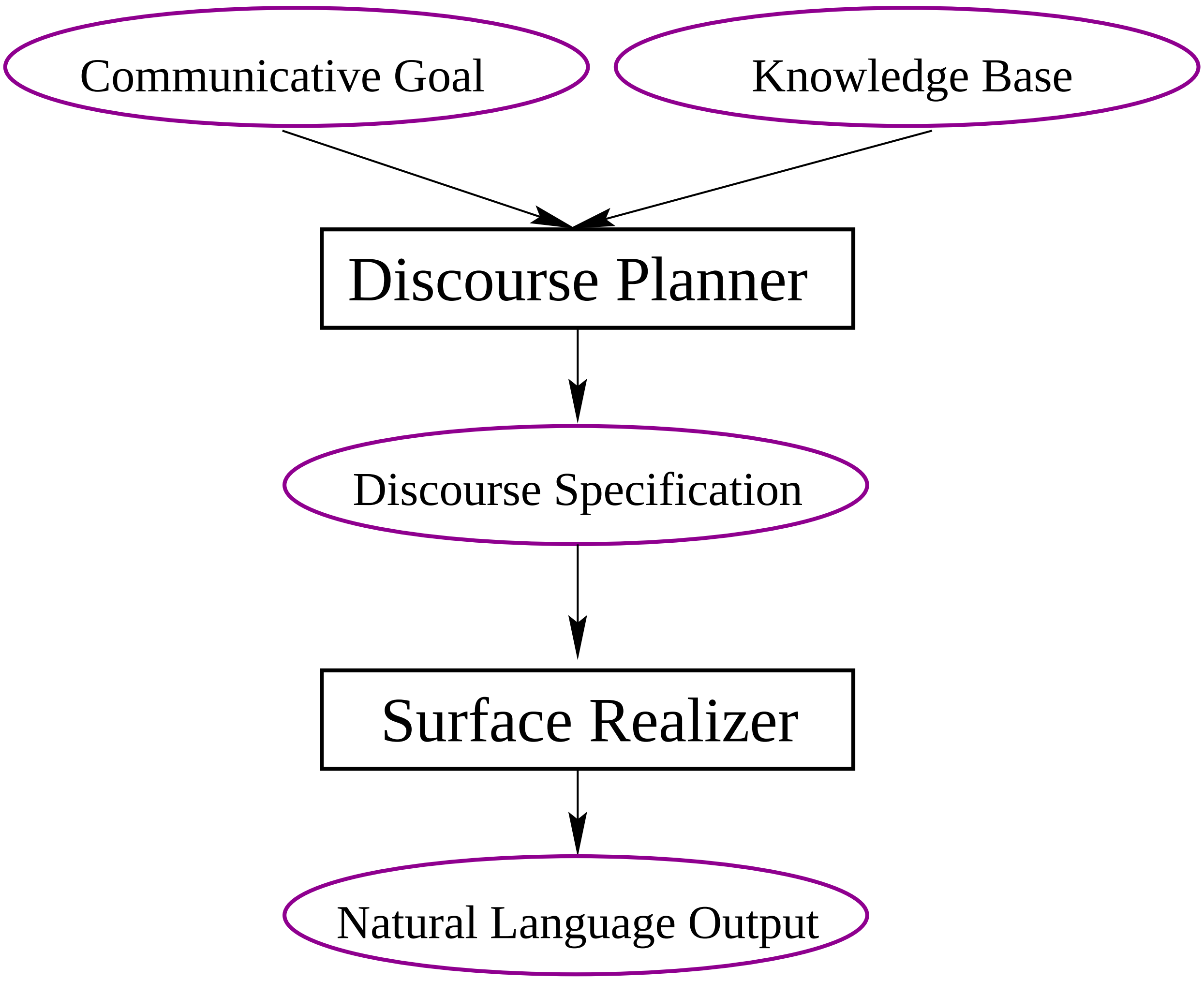
Choice

NLG is primarily concerned with choice (rather than disambiguation as in NLU)

Stages

- content selection
- lexical selection
- sentence structure
  - aggregation
  - referring expressions
- discourse structure

An Architecture



Components in the Architecture

Discourse planner

- plans the structure of the discourse
- approaches include text schemata and rhetorical relation planning

Surface realizer

- generates single sentences
- approaches include systemic grammar and functional unification grammar

Surface Realization

Language is viewed as a resource for expressing meaning in context (Halliday, 1985)

Systemic grammar

Layers

- mood (interpersonal issues)
- transitivity (like propositional content / semantics)
- theme (discourse issues)
- NOTE: mood and theme represent concern for context

<u>Functional</u> as opposed to syntactic viewpoint

Functional Unification Grammar

Based on Kay's (83) formalism

Partial information, declarative, uniform, compact

Generation grammars as feature structures with alternation

Same framework used for all stages

- syntactic realization
- lexicalization
- text planning

Functional vs. Structural Analysis

John eats an apple.

Functional

- actor(John), affected(apple), process(eat)

Structural

- NP VP NP

Partial vs. Complete Specification

$$\begin{bmatrix} action & eat \\ actor & John \\ object & apple \end{bmatrix}$$

Voice

- *An apple is eaten by John.*

Tense

- *John ate an apple.*

Mode

- *Did John eat an apple?*

Modality

- *John must eat an apple.*

Unification

Target sentence

Input functional description (FD)

Grammar

Unification process

Linearization process

Sample Input

$$\begin{bmatrix} CAT & S \\ ACTOR & \begin{bmatrix} HEAD & \begin{bmatrix} LEX & SYSTEM \end{bmatrix} \end{bmatrix} \\ PROCESS & \begin{bmatrix} HEAD & \begin{bmatrix} LEX & SAVE \end{bmatrix} \\ TENSE & FUTURE \end{bmatrix} \\ GOAL & \begin{bmatrix} HEAD & \begin{bmatrix} LEX & DOCUMENT \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Sample Grammar

$$
\begin{bmatrix}
\text{alt} \left\{
\begin{array}{l}
\begin{bmatrix}
\text{cat} & \text{s} \\
\text{actor} & \begin{bmatrix} \text{cat} & \text{np} \end{bmatrix} \\
\text{process} & \begin{bmatrix} \text{cat} & \text{vp} \\ \text{number} & \{\text{actor number}\} \end{bmatrix} \\
\text{goal} & \begin{bmatrix} \text{cat} & \text{np} \end{bmatrix} \\
\text{pattern} & \text{(actor process goal)}
\end{bmatrix} \\[2ex]
\begin{bmatrix}
\text{cat} & \text{np} \\
\text{head} & \begin{bmatrix} \text{cat} & \text{noun} \\ \text{number} & \{\uparrow\uparrow \text{number}\} \end{bmatrix} \\
\text{determiner} & \begin{bmatrix} \text{cat} & \text{article} \\ \text{lex} & \text{“the”} \end{bmatrix} \\
\text{pattern} & \text{(determiner head)}
\end{bmatrix} \\[2ex]
\begin{bmatrix}
\text{cat} & \text{vp} \\
\text{alt} & \left\{
\begin{array}{l}
\begin{bmatrix}
\text{tense} & \text{present} \\
\text{head} & \begin{bmatrix} \text{cat} & \text{verb} \\ \text{number} & \{\uparrow\uparrow \text{number}\} \end{bmatrix} \\
\text{pattern} & \text{(head)}
\end{bmatrix} \\[2ex]
\begin{bmatrix}
\text{tense} & \text{future} \\
\text{auxiliary} & \begin{bmatrix} \text{cat} & \text{modal} \\ \text{lex} & \text{“will”} \end{bmatrix} \\
\text{head} & \begin{bmatrix} \text{cat} & \text{verb} \\ \text{ending} & \text{root} \end{bmatrix} \\
\text{pattern} & \text{(auxiliary head)}
\end{bmatrix}
\end{array}
\right\}
\end{bmatrix}
\end{array}
\right\}
\end{bmatrix}
$$

After 1st Unification

$$\begin{bmatrix} CAT & S \\ ACTOR & \begin{bmatrix} CAT & NP \\ HEAD & \begin{bmatrix} LEX & SYSTEM \end{bmatrix} \end{bmatrix} \\ PROCESS & \begin{bmatrix} CAT & VP \\ NUMBER & \begin{bmatrix} ACTOR\ NUMBER \end{bmatrix} \\ HEAD & \begin{bmatrix} LEX & SAVE \end{bmatrix} \\ TENSE & FUTURE \end{bmatrix} \\ GOAL & \begin{bmatrix} CAT & NP \\ HEAD & \begin{bmatrix} LEX & DOCUMENT \end{bmatrix} \end{bmatrix} \\ PATTERN & (ACTOR\ \ PROCESS\ \ GOAL) \end{bmatrix}$$

The Fully Unified FD

The system will save the document.

$$
\left[\begin{array}{ll}
\text{cat} & \text{s} \\
\text{actor} & \left[\begin{array}{ll}
\text{cat} & \text{np} \\
\text{head} & \left[\begin{array}{ll}
\text{cat} & \text{noun} \\
\text{lex} & \text{system} \\
\text{number} & \left\{\uparrow\uparrow \text{number}\right\}
\end{array}\right] \\
\text{determiner} & \left[\begin{array}{ll}
\text{cat} & \text{article} \\
\text{lex} & \text{``the''}
\end{array}\right] \\
\text{pattern} & \text{(determiner head)}
\end{array}\right] \\
\text{process} & \left[\begin{array}{ll}
\text{cat} & \text{vp} \\
\text{number} & \left\{\text{actor number}\right\} \\
\text{head} & \left[\begin{array}{ll}
\text{cat} & \text{verb} \\
\text{lex} & \text{save} \\
\text{ending} & \text{root}
\end{array}\right] \\
\text{auxiliary} & \left[\begin{array}{ll}
\text{cat} & \text{modal} \\
\text{lex} & \text{``will''}
\end{array}\right] \\
\text{tense} & \text{future} \\
\text{pattern} & \text{(auxiliary root)}
\end{array}\right] \\
\text{goal} & \left[\begin{array}{ll}
\text{cat} & \text{np} \\
\text{head} & \left[\begin{array}{ll}
\text{cat} & \text{noun} \\
\text{lex} & \text{document} \\
\text{number} & \left\{\uparrow\uparrow \text{number}\right\}
\end{array}\right] \\
\text{determiner} & \left[\begin{array}{ll}
\text{cat} & \text{article} \\
\text{lex} & \text{``the''}
\end{array}\right] \\
\text{pattern} & \text{(determiner head)}
\end{array}\right] \\
\text{pattern} & \text{(actor process goal)}
\end{array}\right]
$$

Evaluation

Similarity of the output with a corpus

Human panel evaluates according to some criteria (like Turing Test)

Usability of the larger system

Not Covered

Discourse Planning

- text schemata
- rhetorical relations / rhetorical structure theory

Other Topics

- microplanning
- lexical selection
- speech

Recent Statistical Work

For Next Time

Study for exam!

Acknowledgements

Some of this material was taken from the course web pages of Drago Radev.