Computation, Information, and Intelligence (COMS/ENGRI/INFO/COGST 172), Fall 2005
11/14/05: **Lecture 33 aid — Discourse structure**

Agenda: Grosz and Sidner's (1986) theory of discourse, introduced in "Attention, Intentions, and the Structure of Discourse", *Computational Linguistics*, pp. 175–204.

Follow-ups: We gave an example last time in which recall was higher than precision. Here is an example in which the reverse is true: the correct segmentation is "this|is|a|sentence" and the system outputs "this|isasentence".

I. Reference to previous segments Adapted from Sidner (1979). Imagine that the speaker is a distinguished professor of philosophy.

Version (A):

1. A while ago, I ordered a rare manuscript from an overseas antique dealer.
2. It contained the early writings of the "Bohemian" philosophers,
3. a group of radicals from Prague
4. whose theory of poetics is fiendishly difficult to understand.
5. I only got it after several months of consulting another source.

Version (B):

1. A while ago, I ordered a rare manuscript from an overseas antique dealer.
2. It contained the early writings of the "Bohemian" philosophers,
3. a group of radicals from Prague
4. whose theory of poetics is fiendishly difficult to understand.
5. But anyway, I only got it after several months of consulting another source.

Version (C):

1. A while ago, I ordered a rare manuscript from an overseas antique dealer.
2. It contained the early writings of the "Bohemian" philosophers,
3. a group of radicals from Prague
4. whose theory of poetics is fiendishly difficult to understand.
5. Anyway, I only got the book after several months of consulting another source.
6. Their theory of ethics is even more demanding. (??)

II. Incoherent (yet plausible) discourse Grosz and Sidner cite this as appearing in Polanyi and Scha, "forthcoming".

1. John came by and left the groceries.
2. Stop that you kids.
3. And I put them away after he left.

III. The importance of recognizing intentions From Grishman (1986), pg. 157.

A: Do you know when the train to Boston leaves?
B: Yes.
A: I want to know when the train to Boston leaves.
B: I understand.

(OVER)

IV. The three components of the Grosz and Sidner theory

1. *Linguistic:* contiguous groupings of the sentences (or utterances) into *discourse segments*. These are (implicitly) "defined" by the fact of their correspondence to a single *discourse segment purpose* (DSP), which is defined as the one intention that the speaker wants the other conversational participant(s) to recognize.

2. *Intentional:* relations between DSPs:

   - *domination*: satisfaction of one DSP furthers another DSP;
   - *sat-precedence*: satisfaction of one DSP must precede the satisfaction of another DSP.

   The intentional structure defined by these two relations can be represented by a partially-ordered tree. *Cue phrases* aid the recognition of this structure.

3. *Attentional:* groupings of discourse entities into *focus spaces*, indicating a *salience ordering*. The focus spaces for active discourse segments are maintained on the *focus stack*.

V. Example of intentional structure From Jurafsky and Martin (2000), pp. 745–747. C: caller; A: agent. Note that because this example does not address the attentional structure of the discourse, it does *not* constitute a full analysis.

```
C1: ...I need to travel in May.
A1: And, what day in May did you want to travel?
C2: OK uh I need to be there for a meeting that's from the 12th to the
    15th.
A2: And you're flying into what city?
C3: Seattle.
A3: And what time would you like to leave Pittsburgh?
C4: Uh hmm I don't think there's many options for non-stop.
A4: Right. There's three non-stops today.
C5: What are they?
A5: The first one departs PGH at 10:00am arrives Seattle at 12:05 their
    time. The second flight departs PGH at 5:55pm, arrives Seattle at
    8pm. And the last flight departs PGH at 8:15pm arrives Seattle at
    10:28pm.
C6: OK I'll take the 5ish flight on the night before on the 11th.
A6: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S.
    Air flight 115.
C7: OK.
```

DS_1 =C1-C7	DSP_1	C wants A (to want) to find a good flight for C.
DS_2 =A1-C2	DSP_2	A wants C (to want) to give a departure date.
DS_3 =A2-C3	DSP_3	A wants C (to want) to give a destination.
DS_4 =A3	DSP_4	A wants C (to want) to give a departure time.
DS_5 =C4-C7	DSP_5	C wants A (to want) to find a nonstop flight.