Attention, Intentions, and the Structure of Discourse

Barbara J. Grosz
Artificial Intelligence Center and Center for the Study of
Language and Information
SRI International

Candace L. Sidner
BBN Laboratories, Inc.

Presenter: Juan Ramos

Introduction

- Objective: build computational model of discourse
  - What makes discourse understandable?
  - How do we recognize discourse from a sequence of utterances?
- Generalize previous work to accommodate various discourse types
- Three constructs for discourse: linguistic structure, intentional structures, attentional states

Useful (Informal) Definitions (p. 176-177)

- Utterance: actual sequence of phrases and clauses that are spoken/written.
- Discourse: piece of language behavior that typically involves multiple utterances and multiple participants
- Discourse Segment (DS): informally, a "subtopic" within the full discourse
- Initiating Conversational Participant (ICP): entity that utters the first element of a DS. Unique for each DS
- Other Conversational Participant(s) (OCP): non-ICP entities that take part in a DC

Theoretical Basis

- Linguistic structure: arrangement of words, phrases into utterances, discourse structures
- Intentional structure: capture intentions of ICP and OCPs in discourse segments
- Attentional structure: capture information about world elements, their relationships, and their salience in discourse segments

Linguistic Structure

- Structure of utterances that make up a discourse. Aggregating utterances results in discourse segments
  - (Mann et al 75): humans classify DS quite well, with only relatively minor disagreements in DS boundaries
  - Apparent indications of DS boundaries: change in speech rate, intonation, pauses

Linguistic Structure

- Not strictly decompositional: can have DSs within other DSs, overlapping DSs
- Expressions give clues to DS structure ⇔ DS structure constrains linguistic choices
- Clue/cue words/phrases: *first, for example, before going on, that reminds me,* etc.

Intentional Structure

- Determine how the goal(s) and intention(s) of ICPs, OCPs in conversation affects discourse and DS structure
- Discourse purpose (DP): principal intention in the overall discourse. Eg "X is a bad candidate, so vote for Y"
- Discourse segment purpose (DSP): principal intention in each DS. Supports achievement of DP. Eg "X did this and that"

Intentional Structure

- DP and DSPs are (by necessity) intended by IPC to be noticed by OPC, either explicitly or implicitly
- Important relations
  - Dominance: if DSP1 helps to satisfy DSP2, then DSP2 dominates DSP1
  - Satisfaction-precedence: if DSP2 must be satisfied before DSP1, then DSP2 satisfaction-precedes DSP1
- Infinite number of DSPs, small finite relationships between DSPs

Attentional State

- Abstraction of discourse participants' attention focus
- Focus spaces: set of attentional state models represented in a stack data structure
- Focusing structure: collection of salient focus spaces in a discourse or DS. Altered via focusing process

Attentional State

- Contents of focus state: DSP, salient objects of a DS
- Representation: Stack data structure
  - Operations: push, pop focus spaces
  - Top focus space is most salient, other focus spaces still present but less salient

Examples

- Movies essay
- Task oriented dialogue

Implementation - Intention

- Issue: recognize DP/DSPs in a computational setting
- Key elements: utterance-level intentions, linguistic markers (eg cue phrases), world knowledge
- Emphasis on cue phrases, though these are often ambiguous: "But anyway, about what I was saying…"
- ICP can choose how much information about DSP to convey explicitly:
  - "I want you to arrange a trip for me to Palo Alto, with a two week duration and flying on TWA."

Implementation - Intention

- DP/DSPs not fully realized until end of corresponding DP/DSP
- OCP must "guess" DSP partially according to ongoing information
- See movies example

Implementation – Attentional states

- Use a stack data structure to model salient focus spaces, dominance and satisfaction-precedence relationships
- Limit search space for definite nouns and anaphora resolution (using added mechanisms, eg centering)

Interruptions

- P. 192: "John came by and left the groceries… *Stop that you kids!* And I put them away after he left."
- Possible definitions:
  - Strong: DP whose DSP is not dominated nor satisfaction-preceded by any prior DSP
  - Weak: DP whose DSP is not dominated nor satisfaction-preceded by immediately prior DSP

Interruptions

- True interruptions: "…selling the shoes I… oh, look at the time!"
  - Strong interruption
  - Cannot be associated with any prior focus space in focus stack
- Flashbacks: "Let me go back to what I said before…"
  - Interrupting DSP satisfies-precedes interrupted DSP
  - Interrupting DSP is dominated by some other DSP
  - Auxiliary stack saves focus states popped from main stack when making flashback

Interruptions

- Digressions: "Speaking of which, remember that conversation we had before?"
  - Strong interruption with a common entity in interrupted DS and the interruption itself
  - Normally treated as pushing a new focus space onto stack
- Semantic return: introduce a focus space used in a previous (long past) discourse
  - "Remember our discussion about Jack at the party? Well…" (p. 196)
  - Not fully a return to a previous space or its references
  - DSP of semantic return still adds to current DP

Cue Phrases

- Discourses don't usually state a change in DS explicitly or directly; ICP more likely to use cue phrases to signal change
- Key information for ICP to supply(p. 196):
  - Change of attention is imminent
  - If change returns to previous focus space (through various pop op's), or goes to new pace via push op.
  - How new intention relates to others
  - Relevant precedence relationships
  - What intention is entering into focus
- Cue phrases aid in identification of first four

Cue Phrases - Examples

- "For example": establish dominance relationships between a new DSP and a previous one
- "First, second, finally": establish satisfaction-precedence

Cue Phrases and Interruptions

- There are cue phrases unique to each kind of interruption
- True interruption: "Excuse me, wait a second…"
- Flashback: "Let me go back to, I forgot about…"
- Digression: "Speaking of which, now that you mention it…"

More Cue Phrases (p. 198)

- End of a DS: "The end, ok!, That's all, folks!"
- Announce a push operation to stack: "That reminds me…"
- Ambiguous: "now, next…"
- Cue phrases are helpful, but not required

Conclusions

- Authors propose basis for future investigation and discourse-processing systems
- Notable findings (p. 202):
  - DP must be shared by ICP and all OCPs
  - General ideas about "topic" correlate closely to DP/DSPs
  - One intentional structure can have different attentional structures

Relevance to Meaning Machines

- Clark's paper on joint projects
  - Cooperation and mutual understanding required from ICP/OCPs for successful project/discourse
  - Projects, like DS's, can have subgoals that build on and support main goal
  - Special cases of interruption: alter project ("wait a minute…"), decline project ("I don't follow you…"), withdraw from project ("I don't want to hear this anymore…")
  - Evidence gathering: allows for modification of DP (eg uttering a DS not originally considered) or confirmation of understanding of DP/DSP

Relevance to Meaning Machines

- There are other cases of discourses where participants might not share intentions (eg an argument). What would the resulting discourse/DPs/DSPs/focus stacks look like?