Discourse Obligations in Dialogue Processing

Traum and Allen 1994

Anubha Kothari
Meaning Machines, 10/13/04

Main Question

Why and how should discourse obligations be incorporated into models of social interaction?
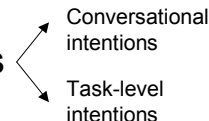
The Intentional Story

- True of most computational models of discourse
  - An agent has certain goals & forms intentions
  - Acts on those intentions & produces utterances
  - Hearer reconstructs a model of speaker's intentions
  - She adopts speaker's goals and acts on that basis
- Focus on recognition of speaker intentions and forming a joint plan/intention
- Requires a strong degree of cooperativity

Problems with this story

- Consider question-answering:
  - Hearer must adopt speaker's goal of finding out the answer and then produce a reply of some sort
  - Why should the hearer say anything at all when she doesn't know the answer or when she isn't predisposed to adopting the other's goals (e.g. politicians in debates!)?
  - What if the hearer may not be interested in forming a shared plan but complies anyway (e.g. when a stranger asks her "Do you have the time?")
- Automatic goal adoption doesn't accommodate cases in which one might want to violate rules of cooperation, esp. when they are in conflict with one's personal goals.
- So what goals motivate conversational cooperation?

Intentional/plan-based attempts at solving this problem:

- Intentional analysis at two levels
  - Conversational intentions
  - Task-level intentions

- Social intentional constructs like 'Joint intentions' or 'Shared plans'
- Cooperative dialogue games (with fixed set of responses) at discourse level

  All of these muddle up intentions with straightforward social conventions, or conflate personal and higher-level goals.
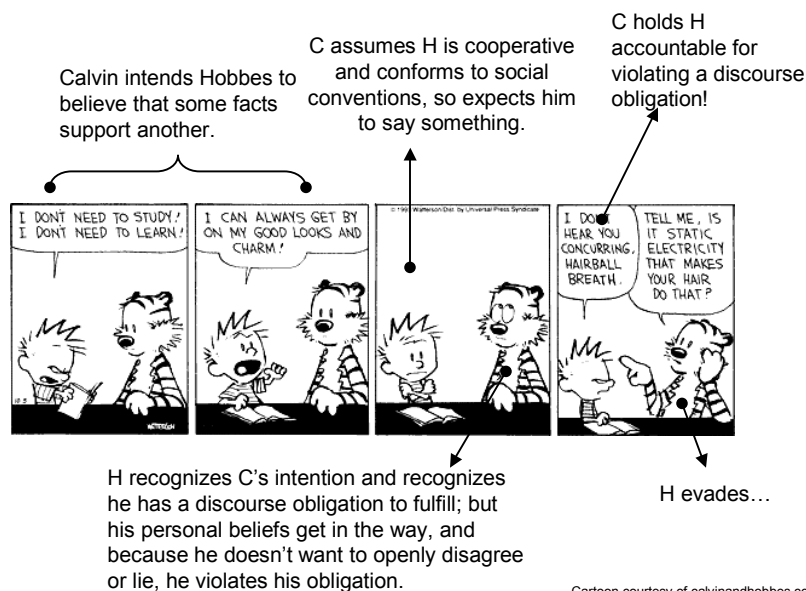
Alternative: Model based on Obligations

- Quite intuitively, "much of one's behavior arises from a sense of obligation to behave within limits set by the society that the agent is part of."
- One feels obliged to respond b/c of social convention.
- So include obligations in our model!
- But crucially, "obligations are independent of shared plans and intention recognition."
- People have a complex set of motivations for action, but there is a clear, sensible separation between obligations and intentions/domain-level goals. The former cannot be confused for or reduced to the latter.

Obligations and Discourse Obligations

- An action is *obligatory* if it is not permissible not to do it, w.r.t. a set of rules.
- But just because it is obligatory doesn't mean the agent will do it => it can be violated.
- Many kinds of obligations (e.g. to fulfill a promise, to address a request, etc.)
- *Discourse obligation* = the obligation to say something.

The Basic Ideas Illustrated

Calvin intends Hobbes to believe that some facts support another.

C assumes H is cooperative and conforms to social conventions, so expects him to say something.

C holds H accountable for violating a discourse obligation!

H recognizes C's intention and recognizes he has a discourse obligation to fulfill; but his personal beliefs get in the way, and because he doesn't want to openly disagree or lie, he violates his obligation.

H evades…

Cartoon courtesy of calvinandhobbes.com

Prioritization Problems

- More modularized, less complicated model of discourse behavior
- But how does the agent plan actions now?
  - Do obligations always get first priority when they conflict with prior intentions?
  - Or are personal goals most important, so much so that conflicting obligations are violated?
- "An agent may have many obligations and many different goals, and planning involves a complex tradeoff between these different factors."
- A multiple-constraint satisfaction problem

How should robots behave?

Should they be allowed to violate obligations?

Must we enforce that they always adopt the goals of the human who speaks to them?

Do they have personal goals, and if so, when do those take priority?

What kind of reasoning and prioritization scheme do we want to endow them with?

"The ethics of AI servitude"

Asimov's Three Laws of Robotics (from *I, Robot*):

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey orders given to it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

The TRAINS Dialogue System

- Intelligent assistant that helps create and execute a transportation plan; uses NL to converse
- Has a 'dialogue manager' that explicitly uses discourse obligations
  - maintains flow of conversation and ensures conversational goals are met
  - interfaces with plan reasoner and executor
  - two stacks of pending discourse obligations (one for user, one for discourse actor/agent)
  - attend to obligations before attending to other aspects of the discourse; pop obligation when completely done

Discourse Actor Algorithm

"When deciding what to do next, the agent first considers obligations and decides how to update the intentional structure (add new goals or intentions) based on these obligations. Obligations might also lead directly to immediate action. If there are no obligations, then the agent will consider its intentions and perform any actions which it can to satisfy these intentions. If there are no performable intentions, then the system will deliberate on its overall goals and perhaps adopt some new intentions."

Loop reflecting prioritization

while conversation is not finished
    if system has obligations
    then address obligations
    else if system has turn
    then if system has intended conversation acts
        then call generator to produce NL utterances
        else if some material is ungrounded
        then address grounding situation
        else if some proposal is not accepted
        then consider proposals
        else if high-level goals are unsatisfied
        then address goals
        else release turn
            or attempt to end conversation
    else if no one has turn
    then take turn
    else if long pause
    then take turn

Sample Dialog

Utt. #	Speaker: Utterance
1	U: We better ship a boxcar of oranges to Bath by 8 AM.
2	S: Okay.
3-3=6	U: So we need to get a boxcar to Corning where there are oranges.
3-7	U: There are oranges at Corning.
3-8	U: Right?
4	S: Right.
5-1	U: So we need an engine to move the boxcar.
5-2:	U: Right?
6	S: Right.
……..	
15-2=4	U: And move it from Dansville to Corning.
15-5=7	U: Load up some oranges into the boxcar.
15-8=10	U: And then move it on to Bath.
16	S: Okay.
17	U: How does that sound?
18-3	S: That's no problem.

Observations & Comments

- Relaxed, unaggressive turn-taking style
- Processing time-lag in addition to perceived slowness?
- Reactive approach => the system does not need complicated planning (like in strong plan-based models)
- Can handle varying degrees of initiative on part of the user
- Flexible architecture; can be set to varying levels of cooperativity

Questions

- Can we make the prioritization scheme more flexible or dynamic – so the actor can switch from an obligation-driven process to goal-driven when it needs to?
- How to improve time-lag?
- What other aspects of social interaction might be useful in dialogue processing?