Dialogue Management
Hauptseminar

Ivana Kruijff-Korbayová
korbay@coli.uni-sb.de
http://www.coli.uni-sb.de/˜korbay/Lehre/DM-beschreibung.html

Sommersemester 2002

Introduction

Today Session Plan:

- Formal matters
- Introduction

Formal Matters

- Participants
- Referat
- Hausarbeit
- Session-question
- Evaluation
- Seminar plan

Introduction

Outline:[1]

- Dialogue management
- Features of natural dialogue
- Difficulties of modelling dialogue

[1] This lecture draws on some material from the "Interactive Natural Language Systems" lectures taught within the Language Technology programme at the Computing Department of Macquarie University, Sydney, by Dale and Schwitter (*http://www.comp.mq.edu.au/courses/comp349/*).

Dialogue management

- Dialogue management systems simulate the process of a dialogue
  - who can speak, when, for how long, what about
  - influence of context on interpretation
  - ability to make a contribution (agents with aims and objectives)
  - reasoning (planning and making inferences using general knowledge)
- Dialogue modelling is necessary for any type of dialogue
- Pleasant and easy interaction is an essential design aspect
- Dialogue systems have specific requirements,
  for example, need for error recovery when system breaks

Features of Natural Dialogue

- Dialogue between humans shows a number of specific characteristics
- Although humans simplify their language in human-computer dialogue, many of these characteristics remain:
  - overall dialogue structure (opening, body, closing)
  - turn-taking (some discipline in who speaks, when and how long)
  - mixed initiative (who is in control of the dialogue)
  - collaboration (trying to achieve some goal together)
  - contextual interpretation
  - error recovery (handling mistakes and misunderstandings)
- These characteristics make dialogue modelling difficult

Example of a Dialogue

. . .

(1) U: I need to travel in May.

(2) S: And, what day in May did you want to travel?

(3) U: OK, uh, I need to be there for a meeting that's from the 12th to the 15th.

(4) S: And you're flying into what city?

(5) U: Seattle.

(6) S: And what time would you like to leave Pittsburgh?

(7) U: Uh hmm I don't think there's many options for non-stop.

(8) S: Right. There's three non-stops today.

(9) U: What are they?

(10) S: The first departs PGH at 10am and arrives Seattle at 1205 their time. The second flight departs PGH at 555pm, arrives Seattle at 8pm. And the last flight departs at 815pm arrives Seattle at 1028pm.

(11) U: OK, I'll take the 5ish flight on the night before on the 11th.

(12) S: On the 11th? OK. Departing at 555pm arrives Seattle at 8pm, US Air flight 115.

(13) U: OK.

Difficulties of Dialogue Modelling

- turn-taking
- conversational fillers (back-channel)
- adjacency pairs and insertion sequences
- indirectness
- ellipsis
- anaphoric reference

Turn-Taking

- A conversation is made up of turns: participant A says something, then participant B does, then A again, then B . . .
- Though it is generally not obvious when a turn in natural dialogue is finished, turn-taking appears fluid (the amount of speech overlap in normal conversation is less than 5%)
- How do participants know when to contribute a turn?
  - Conversational analysis argues that turn-taking is governed by a set of rules
  - Turn-taking rules are assumed to apply at transition-relevance points — places where the dialogue/utterance structure allows speaker shift to occur (typically but not only at utterance boundaries)

Conversational Fillers

- Fillers, i.e., short phrases such as *a-ha, yeah, mm, hmm, right* are often used in conversation
- They do not seem to fulfill an act in themselves, so why are they commonplace?
- It is not obvious what makes dialogue participants to use fillers.
  - Fillers appear to add cohesion to dialogue.
  - Fillers re-assure the speaker that the hearer is paying attention.
  - Fillers provide acknowledgment of (at least) basic "understanding" (e.g., I heard you, I know what (words) you're saying)
- Fillers do not take over the turn, they are back-channel indications that the speaker can continue her turn (also called continuers).

Adjacency Pairs

- Adjacency pairs are manifestations/instances of discourse structure beyond the sentence level.
- An adjacency pair is a sequence of two utterances that are:
  - adjacent (but cf. insertion sequences)
  - produced by different speakers
  - ordered: first and second part
  - typed: a particular first part requires a particular (range of) second part(s)
- Examples are question-answer, greeting-greeting, offer-acceptance, request-grant, apology-minimization, compliment-displayer . . . .
- Silence after the first part of an adjacency pair is a significant silence.

Insertion Sequences

- Adjacency pairs are intuitively sensible structures, however, they do not adequately model insertion sequences, i.e., cases where within an adjacency pair, another adjacency pair is embedded (cf. pumping lemma):

(14) A. Where are you going? Q_1
B. Why? Q_2
A. I thought I'd come with you. A_2
I'm going to the supermarket. A_1

- There is no closed set of insertion sequences
  - another example: (7-11) above (beyond pumping: (Q_1 (Q_2-A_2; Q_3-A_3) A_1))
  - another case: misapprehension; clarification, termination

Indirectness

- Indirectness means that participants can make sense of a dialogue beyond what is explicitly said, e.g., (3) above.
- The interpretation often goes beyond the literal meaning of what is said (cf. Grice's notion of conveyed/communicated meaning)
- A wide range of cognitive skills and world knowledge comes into play when interpreting a dialogue such as

(15) A. That's the telephone.
B. I'm in the bath.
A. OK

- Considering only the literal meaning, this dialogue is incoherent.
- But, given the real world knowledge of the participants, what is actually communicated can be paraphrased as follows:

(15) A. That's the telephone.
A requests B to perform an action (answer the phone)
B. I'm in the bath.
B states a reason (being in bath) why he cannot comply with the request, thereby not granting the request.
A. OK.
A accepts the reason and (presumably) undertakes to perform action (answer phone)

- Dialogue is interpreted as social interaction ($\rightarrow$ speech acts)

Ellipsis

- Ellipsis occurs in sentences where for reasons of economy and cohesion a part of structure is omitted.
- The missing structure can normally be recovered from the dialogue context:

(16) A. What time is Twelfth Night paying tonight?
B. 8:10 pm.
A. And Hamlet?

(cf. also (5))

- Keeping track of the context is essential to coherent dialogue.
- Without modelling ellipsis, dialogue can appear unnatural.

Anaphoric Reference

- Anaphors are expressions which refer back to previously mentioned entities or concepts (e.g., 'they' (9), 'the 11th' (12), 'the first', 'the second', 'the last' (10), 'the 5ish' (11) (Forward-referring expressions are called cataphors.)
- Anaphors reduce redundancy and increase cohesion (as ellipsis).
- Deixis is the linguistic term for expressions which need to be interpreted in the situational context (outside the linguistic context of the dialogue), e.g., 'now/then', 'here/there', 'I/you', 'this/that'
- Anaphors and deixis can be augmented by gestures (e.g., pointing, nodding).
- Therefore, the context must be sufficiently modelled to resolve these.

Error Recovery

- Many things can go wrong in dialogue: repairs (11), mistakes, misunderstandings, . . .
- In dialogue system, additional errors can happen due to the (inevitable) limitations of the system, or due to a failure (esp. in spoken dialogue systems).
- Therefore, the dialogue manager of a system should be able to
  - detect errors
  - have repair strategies to recover from errors (clarification, correction, . . .) or at least fail gracefully
- When attempting to correct an error, more errors can arise.

Summary

We have addressed essential properties of natural dialogues which make their modelling difficult:

- dialogue structure and turn-taking, conversational fillers (back-channel acknowledgments)
- adjacency pairs and insertion sequences
- indirectness and the need for reasoning with general knowledge
- contextual reference (ellipsis, anaphoric reference, deixis) and the need for a sufficient model of both the linguistic and extra-linguistic context
- recovery from errors

Outlook

Next time, we will make an overview of some existing dialogue systems, the technology they employ and the kind of dialogues they handle, including demonstrations when possible. For example. . .

- Eliza
- TRAINS, TRIPPS
- Philips Train Timetable system
- SRI-Autoroute
- Verbmobil
- Trindi and Siridus systems (MIDAS, PTT and EDIS, GoDIS)
- D'Homme

and then, it will be your turn to start talking!