

Computational Dialogue Modelling

Raquel Fernández

Institute for Logic, Language & Computation
University of Amsterdam

www.illc.uva.nl/~raquel

Core Logic 2009

Computational Dialogue Modelling

- **Dialogue Modelling** is a fairly new research area at the interface of (computational) linguistics, artificial intelligence, computer science, psychology, neural science, philosophy of language, . . .
- It is concerned with designing formal systems that model aspects of **natural conversation**. Some general research questions are:
 - What kind of skills (linguistic and otherwise) are required to participate in conversation?
 - What kind of information does a participant need to keep track of?
 - What makes a dialogue coherent? How is dialogue structured?
 - How can we design artificial conversational agents that allow natural human-computer interaction?
- Roughly speaking, models of dialogue are considered '**computational**' if they are precise enough to be implemented in a computer, so that
 - they can be evaluated automatically
 - some of their properties can be verified automatically
 - practical tasks can be accomplished automatically

Outline of the Lecture

- Features of dialogue
 - Interaction
 - Coherence
- An abstract model
 - Dialogue as a game
 - Information state update
- Dialogue systems
 - Brief introduction
- Specific research topics
- Conclusions

Dialogue as a Form of Interaction

- Traditionally, (computational) linguistics has focused on analysing isolated sentences or written text.
- Dialogue is a form of **interaction** and hence brings in additional challenges.
- Crucially, it involves **multiple participants** and it **unfolds in time**.
- Participants are autonomous rational agents with their own intentions and interests. This shapes the interaction, introduces room for misunderstanding, and hence requires **coordination**.
- Timing matters: it also requires coordination – who speaks when. Furthermore, the spontaneity of speech often results in disfluencies that render utterances ‘ungrammatical’.

A Dialogue Transcript

From Levinson (1983) on Conversation Analysis (Schegloff 1972).

B: I ordered some paint from you uh a couple of weeks ago some vermilion
A: Yuh
B: And I wanted to order some more the name is Boyd
A: Yes // how many tubes would you like sir
B: U:hm (.) What's the price now eh with V.A.T. do you know eh
A: Er I'll just work that out for you =
B: = Thanks
(10.0)
A: Three pounds nineteen a tube sir
B: Three nineteen is it =
A: = Yeah
B: E::h (1.0) That's for the large tube isn't it
A: Well yeah it's the thirty-seven c.c.s.
B: Er, I'll tell you what I'll just eh eh ring you back I have to work
out how many I'll need. Sorry I did- wasn't sure of the price you see
A: Okay.

Levinson (1983) *Pragmatics*, Cambridge University Press.

Schegloff (1972) Sequencing in Conversational Openings. In *Directions in Sociolinguistics*, pp. 346–380.

Utterances, Dialogue Acts, and Coherence

- The minimal unit of analysis is the **utterance** (within a *turn*).
- Utterances are different from traditional sentences and can be defined as possibly disfluent and non-sentential **intentional units**.
- As other intentional behaviour, utterances can be analysed as *actions*, in particular as instantiating **dialogue act types**.
- Many different inventories of DA types (similar to *speech acts*): assert, request, accept, commit, acknowledge, hold, ...
- The utterances in a dialogue are somehow *connected* to form a **coherent discourse**:
 - new utterances relate to previous context;
 - the choice of an utterance constrains the future dialogue.
- In natural dialogues there are **regular patterns** of utterance types. Some DA types such as questions have *preferred* and *dispreferred* replies. Preferred replies are immediately relevant and expected.

Allen & Core (1997) *Draft of DAMSL: Dialogue Act Markup in Several Layers*. Discourse Research Initiative.

Grounding and Meta-communication

- During conversation, participants need to coordinate their interaction and make sure they understand each other.
- **Grounding** is the process by which participants reach mutual understanding (Clark & Schaefer 1989, Clark 1996).
- Participants need to **signal understanding** or else **request repair**.
- Grounding takes place at a meta-level (a *collateral track*):

communicative acts

B: I ordered some paint from you...

B: And I wanted to order...

A: Bill is around.

A: Bill...
...is around.

meta-communicative acts

A: Yuh

B: Bill Johnston?

A: Yes.

eh, I mean, John

- Modelling grounding is an important part of modelling dialogue.

Clark & Schaefer (1989) Contributing to discourse. *Cognitive Science*, 13:259–294.

Clark (1996) *Using Language*. Cambridge University Press.

Traum (1994) *A Computational Theory of Grounding in Natural Language Conversation*, PhD Thesis, Univ. Rochester.

A Sketch of a Formal Model

- The dynamics of dialogue can be modelled using a game metaphor, where participants (players) make moves that update an evolving *conversational scoreboard* (Lewis 1979).
- The scoreboard contains different types of information, including the *common ground* of the participants, and it is used to keep track of previous actions and to motivate future action.
- In abstract terms, a dialogue can be modelled as:
 - A set S of **dialogue states**, representing possible configurations of the conversational scoreboard;
 - A set M of **dialogue moves**, representing dialogue act types;
 - An **update function** $\delta : (S \times M) \rightarrow S$, that updates the conversational scoreboard given the current state of the dialogue and a new dialogue move.
 - m is a **coherent next move** at a state s iff $\delta(s, m)$ is defined.

Lewis (1979) Scorekeeping in a Language Game. *Journal of Philosophical Logic*, 8(1): 339–359.

Fernández & Endriss (2007) Abstract Models for Dialogue Protocols. *Jrnl. of Logic, Lang. & Info*, 16:121–140.

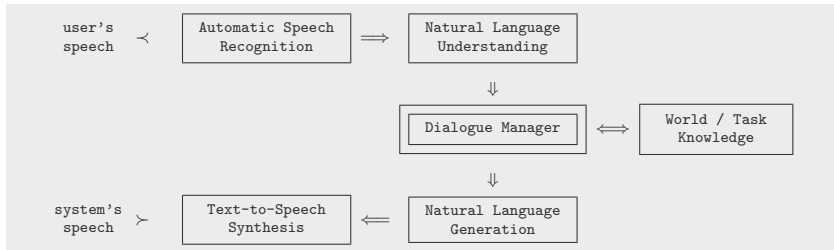
Information State Update Approach

- The previous abstract formulation contains the main ideas at the core of the **Information State Update (ISU)** approach, which currently is one of the most influential models.
- There are many tricky issues that need to be worked out in a detailed model:
 - What information do dialogue states keep track of?
 - Is there only one repository representing common ground? Is there a distinct informational state for each dialogue participant?
 - What is the exact specification of the update function?
 - What strategy can be used to choose a next dialogue move from a set of possible coherent next moves?
- The main goals behind the approach are to explain and predict dialogue phenomena, and to employ this knowledge to develop algorithms for use in human-computer interaction.

Traum & Larsson (2000) The Information State Approach to Dialogue Management. In *Current and New Directions in Discourse and Dialogue*, pp. 325–353.

Dialogue Systems (in brief)

- Spoken Dialogue Systems (SDS) require an **end-to-end** architecture, where all sub-systems of language processing are at play (in different degrees of sophistication).
- The main components of an SDS are the following:



- The **dialogue manager** is the core component of a dialogue system. It can be seen as the implemented version of particular computational models of dialogue.

Additional issues...

... that complicate the picture:

- Timing and **turn-taking**: not only what to say next, but when to say it. When to start speaking, when can backchannels (*uh-huh*) be inserted, etc.
- **Multi-party dialogue** (more than two participants): same update function? turn-taking? who's being addressed? how can an agent decide whether s/he is being addressed?
- **Multimodality**: usually speech is accompanied by other modalities, such as gaze, head nods, gestures. These can be grounding clues, add extra meaning or/and complement speech. Handling multimodality in SDSs requires multimodal *fusion* (for understanding) and possibly *fission* (for generation).

Some Specific Research Topics

- Modelling the meaning and the update effect of **disfluencies**: differences and similarities between self- and other-repair.
- Modelling the concept of **addressing**, in particular, specifying how the pronoun “you” gets resolved in multi-party dialogue.
- Modelling the process of interpreting **referring expressions** to objects in the context: how they are incrementally constructed and resolved by the dialogue participants. [↷ papers available for research reports]

Ginzburg, Fernández & Schlangen (2007) Unifying Self- and Other-Repair. In *Proc. 11th SemDial Workshop (Decalog)*, pp. 57-63.

Frampton et al. (2009) Who is You? Combining Linguistic and Gaze Features to Resolve Second-Person References in Dialogue. In *Proc. 12th Conf. of the European Chapter of the Association for Computational Linguistics (EACL)*.

Fernández, Lucht & Schlangen (2007) Referring under Restricted Interactivity Conditions. In *Proc. 8th SIGdial Workshop on Discourse and Dialogue*, pp. 136-139.

Summing up

- Dialogue modelling is an interdisciplinary research area that aims at modelling aspects of conversational interaction.
- As a field of study, it is amenable to empirical, formal, and computational analyses.
- This is only a brief sketch of the topic. If you want pointers to further references on dialogue modelling, please contact me.
- Some resources:
 - **SIGdial**: Conference of the Special Interest Group on Discourse and Dialogue of the Association for Computational Linguistics [<http://www.sigdial.org>]
 - **SemDial**: Workshop Series on the Semantics and Pragmatics of Dialogue [<http://www.illc.uva.nl/semDial>]
 - **Dialogue & Discourse**: new international online journal [<http://www.dialogue-and-discourse.org>]
- MSc Project - January 2010 (with Galit Sassoon):
Empirically motivated logical representations in lexical semantics