Modelling and Managing Dialogue

Approaches and Challenges

David Schlangen

Department of Linguistics
University of Potsdam
das@ling.uni-potsdam.de

Overview of talk

- Introduction: What is Dialogue Management?
- Challenges
  - *Spontaneous spoken* language
  - Context sensitivity
  - Interaction Management
- Approaches
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

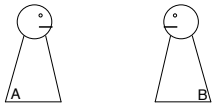
Introduction - What is DM?

- Dialogue management is the task of
  - enabling *understanding* of user utterances in context, and
  - *producing* in a natural way replies that
    - are contextually appropriate (coherent) and
    - drive the task forward.
- In particular, DM needs to maintain a representation of the dialogue context that contains all the required information.

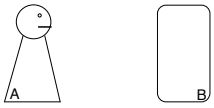
Introduction - What is DM?

- Dialogue management is the task of
  - enabling *understanding* of user utterances in context, and
  - *producing* in a natural way replies that
    - are contextually appropriate (coherent) and
    - drive the task forward.
- In particular, DM needs to maintain a representation of the dialogue context that contains all the required information.

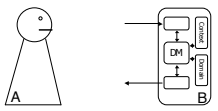
Introduction - What is DM?



Introduction - What is DM?

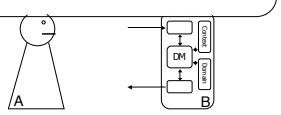


Introduction - What is DM?



Introduction - What is DM?

- but then what is not *dialogue management*?
  - speech recognition, parsing…
  - producing the *form* of the reply
  - managing domain knowledge
  - controlling technical devices



Part I - Challenges

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Challenges - *Spoken* Language

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

The syntax of spoken language differs from that of written language.

Spontaneous: hesitations, abortions, reformulations, self-repairs

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

A: Let's meet next week.
B: I'm busy after the 24th.

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

A: Let's meet next week.
B: I'm busy after the 24th.

October 2003

S	M	Tu	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

A: Let's meet next week.
B: I'm busy after the 24th.

October 2003

S	M	Tu	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

A: Let's meet next week.
B: I'm busy after the 24th.

October 2003

S	M	Tu	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

A: Let's meet next week.
B: I'm busy after the 24th.

October 2003

S	M	Tu	W	Th	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

-> utterances must be interpreted in context!

Challenges - Context Sensitivity

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue means more than the sum of its parts.

A: Did Peter come?
B: He was there, briefly.

A: Is this Peter's car?
B: Hm. The doors look weird.

-> utterances must be interpreted in context!

Challenges - Anaphora

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Anaphora as one device for achieving *cohesion*:

A: Did Peter come?
B: He was there, briefly.

A: Is this Peter's car?
B: Hm. The doors look weird.

Challenges - Fragments

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Fragments: utt. that are intentionally non-sentential, but convey *messages*.

A: Who came to the party?
B: Peter.

Challenges - Fragments

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Fragments: utt. that are intentionally non-sentential, but convey *messages*.

A: Who came to the party?
B: Peter *(came to the party)*.

Challenges - Fragments

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Fragments:

- frequent: around 10% in typical dialogue (Fernández & Ginzburg 2002, Schlangen 2003)
- not just answers, occur in all sorts of contexts

Challenges - Dialogue Acts

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue Acts: what is the function of the utt.?

Challenges - Dialogue Acts

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue Acts: what is the function of the utt.?

- depends on context.

A: Let's meet next week.
B: I'm busy after the 24th.
(Plan correction or plan elaboration?)

Challenges - Dialogue Acts

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue Acts: what is the function of the utt.?

- depends on context.

A: Let's meet next week.
B: I'm busy after the 24th.

(RUDI resolves such relations in dialogues from this domain.)

Challenges - Dialogue Acts

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue Acts: what is the function of the utt.?

- not just one function; several "layers"

answer

A: Who came to the party?
B: Peter.

Challenges - Dialogue Acts

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Dialogue Acts: what is the function of the utt.?

- not just one function; several "layers"

answer

A: Who came to the party?
B: Peter.

understood

Challenges - Gestures

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Non-verbal behaviour
- (gestures,
- facial expressions,
- eye gaze, etc.)

can support both conveying meaning and managing interaction.

Challenges - Interact. Manag.

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Autonomous agents have to coordinate their actions to reach common goal (to have a dialogue).

Challenges - Turn Taking

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Turn taking, observations to account for:

- overlaps are fairly rare in dialogue (less than 5%)
- pauses between turns are very short (around 200ms) --- shorter than motor-planning of new utterance!
- pauses can mean something ("significant silence")

Challenges - Initiative

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Initiative: who is driving the dialogue forward?

- distinguish between:
  - dialogue initiative
  - task initiative

e.g. "How may I help you?" -- DI, but not TI
"Is 4pm a good time for you?" DI and TI

- ideally, both kinds of initiative should be *mixed.*

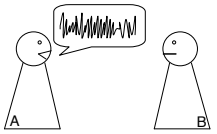
(PotBot: mixed initiative in structured dial. paradigm.)

Challenges - Grounding

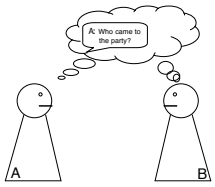
- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Grounding: ensuring that all participants mutually believe they have understood *what was said*. (Clark and Schaefer 1987; Clark 1996)

Grounding



Grounding



Part I - Challenges

- *Spontaneous spoken* language
  - syntax of utterances
  - disfluencies
- Context sensitivity
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

Part II - Models

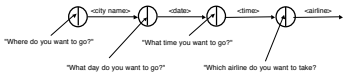
- Phenomena ("Challenges"):
  - Context sensitivity
  - Interaction Management
- Models ("Approaches"):
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

Part II - Models

- Criteria for comparison:
  - How well do models handle phenomena?
  - How do they decide what to say next?
- Models ("Approaches"):
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

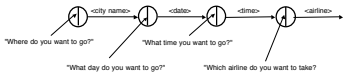
Structured dialogue approaches

- Example: booking a flight



Structured dialogue approaches

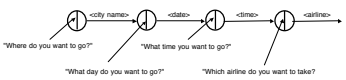
- Example: booking a flight



pro:
- good for ASR: no surprises
- easy to build (initially)

Structured dialogue approaches

- Example: booking a flight

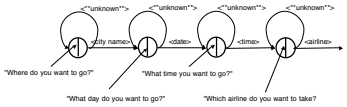


pro:
- good for ASR: no surprises
- easy to build (initially)

cons:
- very inflexible, order is fixed, no overansw.

Structured dialogue approaches

- Example: booking a flight

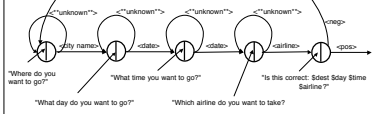


pro:
- good for ASR: no surprises
- easy to build (initially)

cons:
- very inflexible, order is fixed, no overansw.
- everything must be explicit (turn-taking, grounding, context sens., etc.)

Structured dialogue approaches

- Example: booking a flight

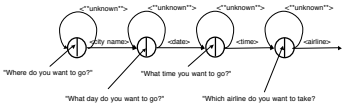


pro:
- good for ASR: no surprises
- easy to build (initially)

cons:
- very inflexible, order is fixed, no overansw.
- everything must be explicit (turn-taking, grounding, context sens., etc.)

Structured dialogue approaches

- Example: booking a flight



pro:
- good for ASR: no surprises
- easy to build (initially)

cons:
- very inflexible, order is fixed, no overansw.
- everything must be explicit (turn-taking, grounding, context sens., etc.)
- can become huge (banking sys: 1,500 states)

Structured dialogue approaches

- Various ways to make them more flexible:
  - forms, agendas...
  - .. topic structures (Stede & Schlangen 2004).
- But in general, it is difficult to separate *discourse knowledge* from *domain knowledge*. As a consequence:
  - it is difficult to integrate principled accounts of dialogue management;
  - system-designer must have detailed advanced knowledge about the dialogue that is to be expected, and hence suitable only for dialogues of limited complexity;
- Unappealing as model of human dialogue behaviour.

Structured dialogue approaches

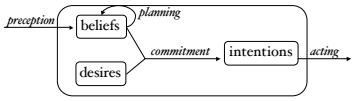
- How do they handle challenges ?
  - Must be hard-coded at each stage.
- How do they decide what to say next?
  - Follow path in predetermined graph.

Part II - Models

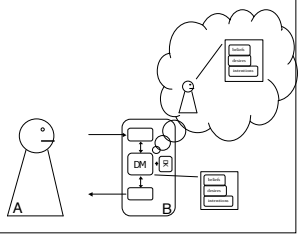
- Phenomena ("Challenges"):
  - Context sensitive interpretation
  - Interaction Management
- Models ("Approaches"):
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

Plan-based approaches

- History: developed out of AI theories of speech acts that integrated speech acts into reasoning about plans. (Cohen, Perreault, Allen 1980s)
- Take notion of *agent* seriously: model interpretation and deliberation with logics of *Beliefs, Desires, Intentions* (BDI)

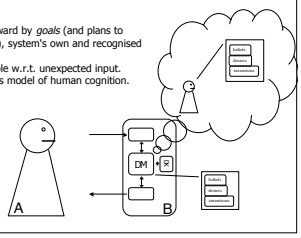


Plan-based approaches



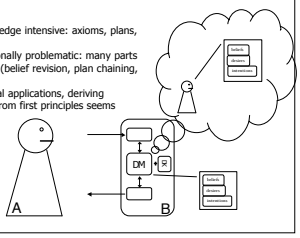
Plan-based approaches

Properties:
- Driven forward by *goals* (and plans to reach them), system's own and recognised from user
- More flexible w.r.t. unexpected input.
- Plausible as model of human cognition.



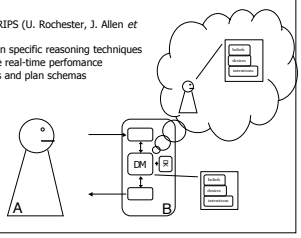
Plan-based approaches

Problems:
- Very knowledge intensive: axioms, plans, schemas
- Computationally problematic: many parts intracktable (belief revision, plan chaining, ...)
- For practical applications, deriving everything from first principles seems overkill.



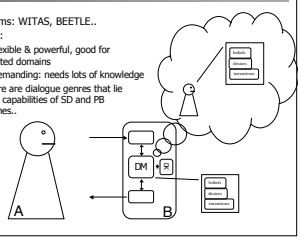
Plan-based approaches

TRAINS / TRIPS (U. Rochester, J. Allen *et al.*):
- use domain specific reasoning techniques to achieve real-time perfomance
- use scripts and plan schemas



Plan-based approaches, summary

Other systems: WITAS, BEETLE..
Conclusions:
- Very flexible & powerful, good for complicated domains
- Very demanding: needs lots of knowledge
- ... There are dialogue genres that lie between capabilities of SD and PB approaches..



Plan-based approaches, summary

- How do they handle challenges ?
  - Context: cognitive states.
  - Can have detailed recipies for how mutual beliefs are formed.
- How do they decide what to say next?
  - Detailed model of how beliefs (e.g. about what was said) and desires form intentions (e.g. to say something).

Part II - Models

- Phenomena ("Challenges"):
  - Context sensitive interpretation
  - Interaction Management
- Models ("Approaches"):
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

Information State Update, intro

- Developed within the EU-project "Trindi" (Larsson 2003; Traum and Larsson 2004), integrating ideas from many previous projects.
- Not an approach *per se:* more an abstraction that allows different approaches to be compared, a *framework.*

Information State Update, intro

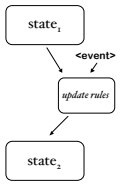
- Starting observation:
  All dialogue models have
  - a notion of the state the dialogue is in at a given point (that can include the dialogue history, the states of the DPs, etc.);
  - a notion of how the dialogue progresses from one state to the next (e.g., which events drive it forward, what are the conditions for which changes, etc.).

Information State Update, intro

- More formally, an ISU-theory consists of:
  - A formal representation of the ***Information State****;* i.e. a specification of its components (BDI, or common ground, or QUD..);
  - A set of ***Dialogue Moves*** that trigger updates (on any level of abstraction: surface moves, logical forms, speech acts);
  - A set of ***Update Rules*** that determine how observed moves change IS, or how changes in IS license moves to make. Rules inspect (parts of the) IS.

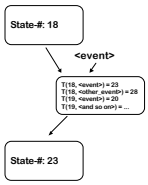
Information State Update, intro

- Schematic view of update process:



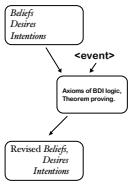
Information State Update, intro

- Is general enough to encode SD-style approaches:



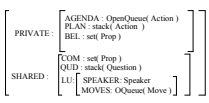
Information State Update, intro

- ... or BDI approach:
- IS is not finite, unlike state in SD approaches!



Information State Update, GODIS

- Example system: GODIS (Larsson 2002), travel booking system
- IS:



(based on Ginzburg's (1998) *Questions Under Discussion* approach)

Information State Update, GODIS

- Example dialogue plan in GODIS:
  - **findout(?x.transport(x))**
  - **findout(?x.dest-city(x))**
  - **findout(?x.depart-city(x))**
  - ....
  - **consultDB(?x.price(x))**
- update rules match these elementary actions to dialogue moves

Information State Update, GODIS

- Example update rule:
  - IntegrateAnswer
    "if the latest move was an assertion from the user that can be understood as a relevant answer to a question under discussion, then combine question and answer to a proposition, which is added to the common ground."

pre: { in($/SHARED/LU/MOVES, answer(A)); fst($/SHARED/QUD, Q); $DOMAIN:relevant(A, Q) }

eff: { ! DOMAIN: combine(Q, A, P); add(/SHARED/COM, P) }

- update rules use simple functions and relations that work on IS fields

Information State Update, GODIS

- Example mini-dialogue:

U: "price information please"
raises price issue:
- if user asks Q, push respond(Q) on AGENDA
- if respond(Q) on AGENDA and PLAN empty, find plan for Q and load to PLAN
- if findout(Q) first on PLAN, ask Q

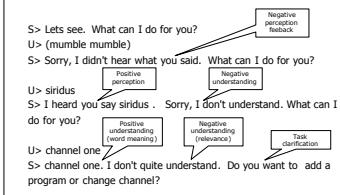
S: "where do you want to go?"

U: "Paris"
- if LM=answer(A) and A **about** Q, add P=Q(A) to SHARED.COM
- if P in SHARED.COM and Q topmost on QUD and P **resolves** Q, pop QUD
- if P in SHARED.COM and P **fulfils goal** of findout(Q) and findout(Q) on PLAN, pop PLAN

Information State Update, GODIS

- So far, nothing that form-filling couldn't handle…
- … but note clear separation of domain knowledge (plans) and conversational knowledge (update rules).
- GODIS adds principled way of dealing with grounding. (On dimension 1, level of action.)

Information State Update, GODIS



GODIS vs. RUDI

- GODIS:
  - practical dialogue system (connected to ASR & synthesis)
  - shallow processing (keyword spotting), simplified semantics
  - (fairly) robust
  - grounding: only level of action
- RUDI:
  - not practical system, "testbed" for theory of dialogue semantics and pragmatics
  - deep processing: "real" grammar, "real" LFs, "real" inference
  - detailed model of grounding
  - not very robust…

RUDI, intro

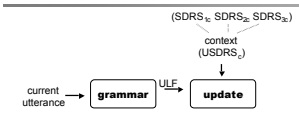
- not a (full) dial sys!
  overhearer that tracks conversation and
  - resolves bridging relations on temporal expressions… (Schlangen, Lascarides and Copestake 2001)
    A: Let's meet next week.
    B: How about Monday? RUDI: Monday of next week.
  - … and fragments (Schlangen and Lascarides 2002)
    A: On what day shall we meet?
    B: On Monday. RUDI: *We shall meet* on Monday.
  - and asks for clarification, if necessary (Schlangen 2004)
    A: Let's meet this weekend.
    B: How about 4 pm? RUDI: 4pm on Sat or Sun?

RUDI in a nutshell

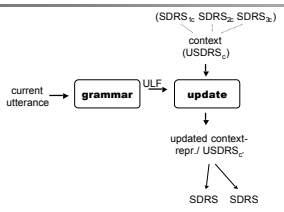
RUDI
- uses logical forms of new utterance and utterances in context, and determines
  - which utterance in context the current utterance is a reply to, and
  - what kind of reply it is (Dialogue Act);
- uses semantic constraints on reply-types (dialogue acts) to resolve missing information (fragments, bridging relations).

Logics of Conversation

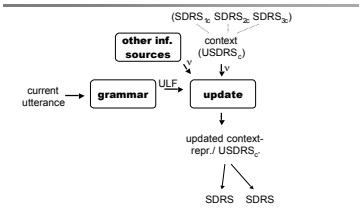
Logics of Conversation



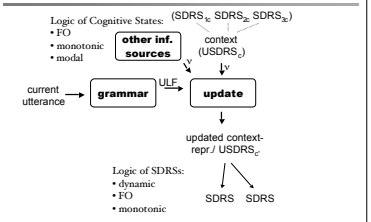
Logics of Conversation



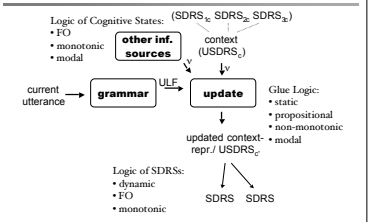
Logics of Conversation



Logics of Conversation



Logics of Conversation



RUDI, SDRT

- Main principles:
  - separate *logic of content* from *logic of information packaging*.
  - when computing coherence of discourse, always use "cheapest" information available: from syntax, lexical semx, semx, to WK & cognitive states.

2 ways of resolving underspec.

A: We should meet next week.
B: How about Friday?
day_of_week(x, Fri) ∧B(x,y) ∧ y=? ∧ B=?

2 ways of resolving underspec.

A: We should meet next week.
B: How about Friday?
day_of_week(x, Fri) ∧B(x,y) ∧ y=? ∧ B=?

`inferring Q-Elab` $?(\alpha,\beta) \wedge \beta :? > \text{Q-Elab}(\alpha,\beta)$

2 ways of resolving underspec.

A: We should meet next week.
B: How about Friday?
day_of_week(x, Fri) ∧B(x,y) ∧ y=? ∧ B=?

`inferring Q-Elab` $?(\alpha,\beta) \wedge \beta :? > \text{Q-Elab}(\alpha,\beta)$

`semantics of Q-Elab` β is a question s.t. all answers elaborate on plan to reach goal of α

2 ways of resolving underspec.

A: We should meet next week.
B: How about Friday?
day_of_week(x, Fri) ∧B(x,y) ∧ y=? ∧ B=?

`inferring Q-Elab` $?(\alpha,\beta) \wedge \beta :? > \text{Q-Elab}(\alpha,\beta)$

`semantics of Q-Elab` β is a question s.t. all answers elaborate on plan to reach goal of α
$\Rightarrow$ temp_inc(SARG_{α} , t_β)

RUDI, intro

- not a (full) dial sys!
  overhearer that tracks conversation and
  - resolves bridging relations on temporal expressions... (Schlangen, Lascarides and Copestake 2001)
    A: Let's meet next week.
    B: How about Monday? RUDI: Monday of next week.
  - ... and fragments (Schlangen and Lascarides 2002)
    A: On what day shall we meet?
    B: On Monday. RUDI: *We shall meet* on Monday.
  - and asks for clarification, if necessary (Schlangen 2004)
    A: Let's meet this weekend.
    B: How about 4 pm? RUDI: 4pm on Sat or Sun?

RUDI, summary

- RUDI:
  - not practical system, "testbed" for theory of dialogue semantics and pragmatics
    - test theory
    - improve practical systems?
  - deep processing: "real" grammar, "real" LFs, "real" inference
  - use different logics: "simple" logic to compute discourse structure, "heavy" logic to represent content.
  - detailed model of grounding
  - not very robust...

ISU, summary

- Information State Update framework is flexible enough to allow many different approaches to be encoded...
- ... and compared.
- Allows degrees of flexibility between SD and PB.

Summary

- Phenomena ("Challenges"):
  - Context sensitive interpretation
  - Interaction Management
- Models ("Approaches"):
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

The end

Thank you!

`http://www.ling.uni-potsdam.de/~das`