

**Modelling and  
Managing Dialogue  
Approaches and Challenges**

---

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

**Introduction**

---

**Two dialogue models / systems:**

- RUDI, implementation of theory of dialogue semantics / pragmatics (SDRT)  
[ joint work with Alex Lascarides (Edinburgh), Ann Copestake (Stanford / Cambridge) ]
- PotBot, a tourism information system  
[ joint work with Manfred Stede (Potsdam) ]

**Introduction**

---

**Classifying work on Dialogue**

- Motivation. *Why?*
- Phenomena. *What?*
- Approach. *How?*

**Introduction**

---

**Classifying work on Dialogue**

- Motivation:
  - AI / cognitive science perspective: construe and test models of human behaviour
  - NLP perspective: get a job done

**Introduction**

---

**Classifying work on Dialogue**

- Motivation:
  - AI / cognitive science perspective: construe and test models of human behaviour
    - RUDI
  - NLP perspective: get a job done
    - PotBot, "information seeking chat"

**Introduction**

---

**Classifying work on Dialogue**

- Phenomena covered:
  - Spontaneous spoken language
    - Syntax of utterances; disfluencies
  - Context sensitive interpretation
    - Anaphora; fragments; dialogue acts; gestures
  - Interaction Management
    - Turn taking; initiative; grounding

**Introduction**

---

**Classifying work on Dialogue**

- Phenomena covered:
  - Spontaneous spoken language
    - Syntax of utterances; disfluencies
  - Context sensitive interpretation
    - Anaphora; fragments; dialogue acts; gestures
  - Interaction Management
    - Turn taking; initiative; grounding

**Introduction**

---

**Classifying work on Dialogue**

- Approach taken:
  - Structured dialogue (FSA, forms, etc.)
  - Plan-based systems
  - Information State Update

**Introduction**

---

**Classifying work on Dialogue**

- Approach taken:
  - Structured dialogue (FSA, forms, etc.)
  - Plan-based systems
  - Information State Update

## Introduction

### Some general points...

- There is not one correct approach to modelling dialogue: depends on what you want to do!
- However, using a common *framework* is good practice, as it makes models comparable. → *Information State Update* model.
- Perspectives (AI and NLP) are complementary, can learn from each other..

## Overview of talk

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

## Part I - Phenomena

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

## Phenomena - *Spoken* Language

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

## Phenomena - *Spoken* Language

- Spontaneous spoken* language
    - syntax of utterances
    - disfluencies
  - Context sensitive interpretation
    - 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

## Phenomena - Context Sensitivity

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

## Phenomena - Context Sensitivity

- Spontaneous spoken* language
    - syntax of utterances
    - disfluencies
  - Context sensitive interpretation
    - anaphora;
    - fragments;
    - dialogue acts;
    - gestures;
  - Interaction Management
    - turn taking;
    - initiative;
    - grounding
- Dialogue means more than the sum of its parts.

## Phenomena - Context Sensitivity

- Spontaneous spoken* language
    - syntax of utterances
    - disfluencies
  - Context sensitive interpretation
    - 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.

## Phenomena - Context Sensitivity

- Spontaneous spoken* language
    - syntax of utterances
    - disfluencies
  - Context sensitive interpretation
    - 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 |    |

### Phenomena - Context Sensitivity

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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.

January 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!

### Phenomena - Context Sensitivity

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

Dialogue means more than the sum of its parts.

A: Is this interesting?  
 B: Yes, very.

A: Is this boring?  
 B: Yes, very.

-> utterances must be interpreted in context!

### Phenomena - Context Sensitivity

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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!

### Phenomena - Anaphora

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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.

### Phenomena - Anaphora

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

Anaphora as one device for achieving cohesion:

Bridging relations (Clark 1977)

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

### Phenomena - Anaphora

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

Anaphora as one device for achieving cohesion:

Bridging relations (Clark 1977)

A: Let's meet next week.  
 B: I'm busy after the 15th.

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

### Phenomena - Fragments

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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.

(RUDI resolves such fragments in dialogues from that domain.)

### Phenomena - Fragments

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

Fragments:

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

### Phenomena - Fragments

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

Fragments, two kinds:

• resolution via identity:

A: On what day shall we meet?  
 B: On Monday (shall we meet).

• resolution via inference:

A: Peter has left already.  
 B: Exams.  
 (= Peter has left because he has to take / supervise / mark / etc. exams)

### Phenomena - Dialogue Acts

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

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

### Phenomena - Dialogue Acts

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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?)

### Phenomena - Dialogue Acts

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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.)

### Phenomena - Dialogue Acts

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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"

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

### Phenomena - Dialogue Acts

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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"

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

understood

### Phenomena - Gestures

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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.

### Phenomena - Interact. Manag.

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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).

### Phenomena - Turn Taking

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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")

### Phenomena - Initiative

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

Initiative: who is driving the dialogue forward?

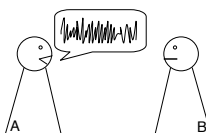
- distinguish between:
  - dialogue initiative
  - task initiative
- "How may I help you?" --- DI, but not TI
- "Is 4pm a good time for you?" --- DI and TI
- ideally, both kinds of initiative should be mixed.
- (PotBot: mixed initiative in structured dial. paradigm.)

## Phenomena - Grounding

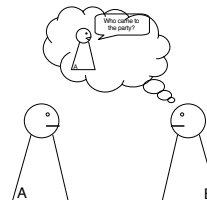
- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - anaphora;
  - fragments;
  - dialogue acts;
  - gestures;
- Interaction Management
  - turn taking;
  - initiative;
  - grounding

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

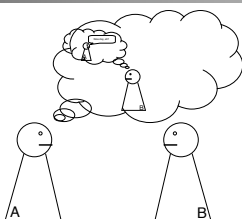
## Grounding



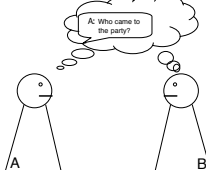
## Grounding



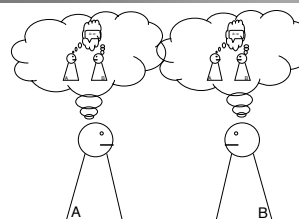
## Grounding



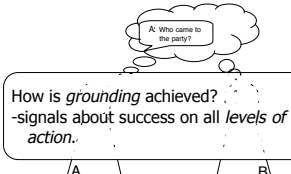
## Grounding



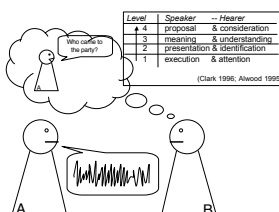
## Grounding



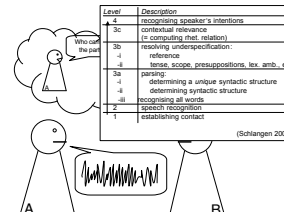
## Grounding



## Grounding - Levels of action



## Grounding - Levels of action



### Grounding - Levels of action

Level	Description
4	recognizing speaker's intentions
3c	contextual relevance
3b	resolving underspecification:
-i	reference
-ii	tense, scope, presuppositions, lex. amb., etc.
3a	planning:
-i	determining a unique syntactic structure
-ii	determining syntactic structure
-iii	recognizing all words
-iv	speech recognition
2	maintaining context
1	

Signal success - how?  
 continued attention, relevant next contribution,  
 acknowledgement, demonstration, display...  
 ... and by indicating a *problem*.

### Grounding - Clarification Requests

Level	Speaker - Hearer
4	proposal & consideration X
3	meaning & understanding X
2	presentation & identification V
1	execution & attention V



### Grounding - Clarification Requests

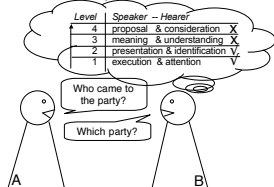
- frequent: around 5% of utterances in task-oriented dialogues (Purver et al. 2001, Rodriguez & Schlangen 2004)

- multi-dimensional classification of problems leading to CRs, from (Schlangen 2004):

- Level of problem
- Extent
- Severity

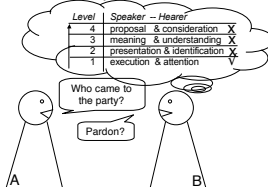
### Clarification Requests

Dimension 1: Level of problem



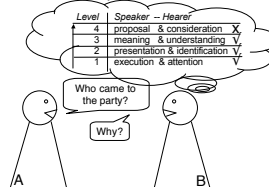
### Clarification Requests

Dimension 1: Level of problem



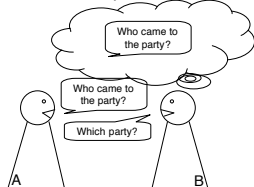
### Clarification Requests

Dimension 1: Level of problem



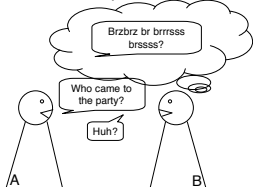
### Clarification Requests

Dimension 2: Extent of problem



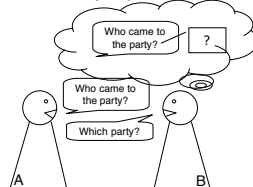
### Clarification Requests

Dimension 2: Extent of problem



### Clarification Requests

Dimension 3: Severity of problem



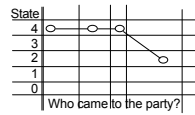
## Clarification Requests

### Dimension 3: Severity of problem



## Dimension "Extent"

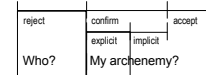
- e.g. "which party?" vs. "huh?"
- (Clark & Schaefer 1987):



## Dimension "Severity"

- is hypothesis maintained or not? ("Peter?" vs. "Who?")
- *quality* of hypothesis / confidence in it
- ... old news to people working on SDS: confidence score of speech recognition

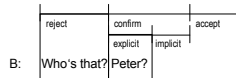
A: I just met your archenemy.



## Dimension "Severity"

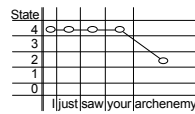
- *quality* of hypothesis / confidence in it needed at *all* levels of processing! (here: reference resolution)

A: I just met your archenemy.



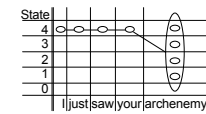
## Interdependencies

- not always in discrete state..



## Interdependencies

- snowballing: wrong hypothesis at one level will lead to dubious hyps. at higher levels.



## Grounding - Summary

- DPs must reach mutual understanding about what was said / meant (to a degree sufficient for current purposes).
- This means they are obliged to indicate problems (& obliged to help to repair them)
- 3 Aspects:
  - Level of understanding at which problem occurred
  - Extent of the problem
  - Severity of the problem

## Part I - Phenomena

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

## Summary challenges



## Summary challenges



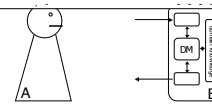
## Summary challenges

- interpretation side:
  - non-standard grammar, disfluencies, fragments
  - (context sensitive recognition of) dialogue acts
- interaction side:
  - managing turn-taking
  - managing initiative
  - managing grounding
- global structure:
  - managing progress of task, *phases*



## Summary challenges

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



## Part II - Models

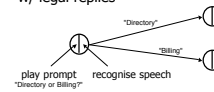
- Phenomena ("Challenges"):
  - *Spontaneous spoken* language
  - Context sensitive interpretation
- Interaction Management
- Models ("Approaches"):
  - structured dialogue approaches
  - plan-based approaches
  - information state update-based approaches
- Summary

## Structured dialogue approaches

- Characteristics: define *legal* dialogues in advance, user "chooses" one of those
  - NLP perspective: get job done...
  - mostly used for dialogues of low complexity..
- several varieties:
  - finite state automata / graphs
  - forms, agendas, *topic structures*
- (vary in strictness of definition of legal dialogs.)

## Structured dialogue approaches

- in (FSM-based) *spoken dialogue systems*: nodes consist of questions (prompts) and arcs are labeled w/ legal replies



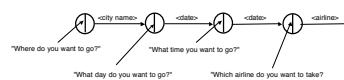
## Structured dialogue approaches

- Example: booking a flight



## Structured dialogue approaches

- Example: booking a flight



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

## Structured dialogue approaches

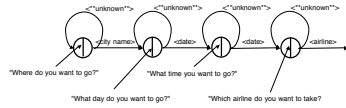
- Example: booking a flight



- pro:
- good for ASR: no surprises
  - easy to build (initially)
- cons:
- very inflexible, order is fixed, no overansw.

## Structured dialogue approaches

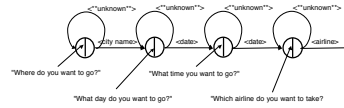
### Example: booking a flight



- 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.)
  - can become huge (banking sys: 1,500 states)

## Structured dialogue approaches

- various ways to make approach more flexible. E.g. forms:

flight_booking	
destination	
origin	
date	
(airline)	
(class)	
(non-stop)	

## Structured dialogue approaches

- forms: first stab at separating (potentially re-usable) *conversational competence* from *domain competence*
- can be extended further:
  - allow sub-routines (e.g. for grounding)
  - use structured parts as elements in planning framework / plan *dialogue games*
- but then loses a lot of the appealing clarity... becomes "hacking"
- unappealing as model of human dialogue behaviour.

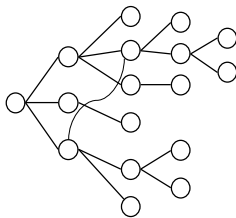
## Structured dialogue approaches

- good where result ("product") is a data-structure; where task is structured by bits of information that have to be *acquired*.
- if well designed, very robust.
- can be extended: let task be structured by information that is to be *conveyed*.
  - --> *PotBot*, the Potsdam Tourism Information system. (Stede & Schlangen 2004)

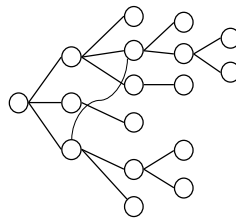
## Example: PotBot

- setting:
  - Models dialogue sub-genre "information-seeking *chat*": explorative, no objective measure of task completion.
- Paradigm example: find out whether place is potential travel destination.
- main ideas of approach:
  - use structure of topic to structure dialogue..
  - ... but only if *system* has initiative; if *user* has initiative, arbitrary jumps are allowed.
  - additionally: learn about users preferences, let that influence decisions as well.

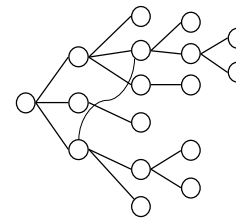
## A Topic Map / Ontology



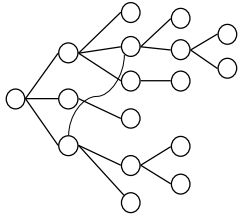
## Initial state: a priori prominence



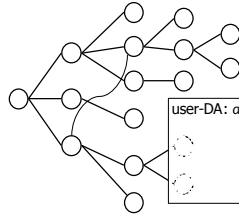
## System gives info about node



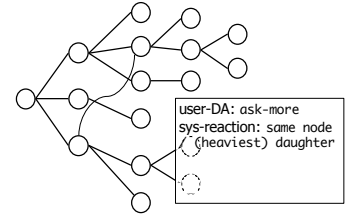
The weight-update mechanism



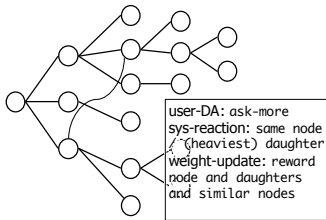
The weight-update mechanism



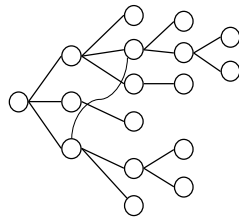
The weight-update mechanism



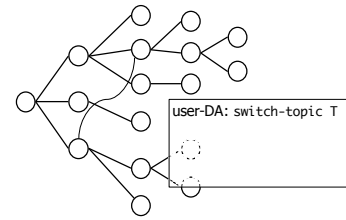
The weight-update mechanism



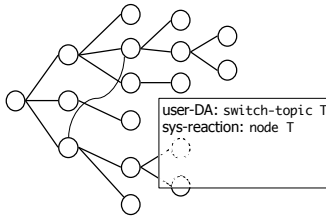
The weight-update mechanism



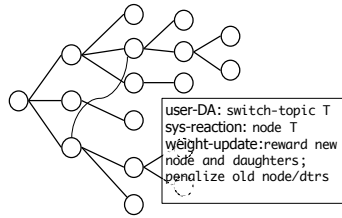
The weight-update mechanism



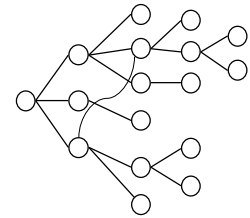
The weight-update mechanism



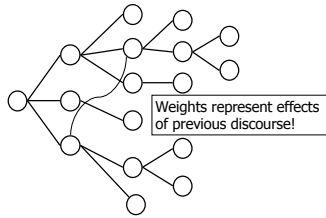
The weight-update mechanism



The weight-update mechanism



### The weight-update mechanism



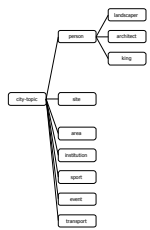
### The City-Tourism Ontology

city-topic

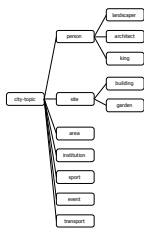
### The City-Tourism Ontology



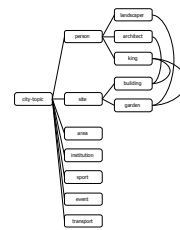
### The City-Tourism Ontology



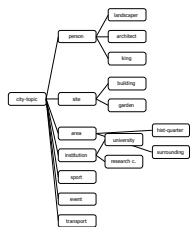
### The City-Tourism Ontology



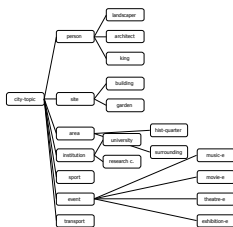
### The City-Tourism Ontology



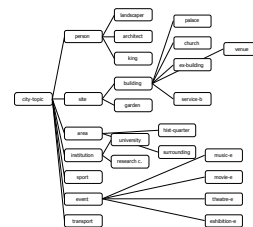
### The City-Tourism Ontology



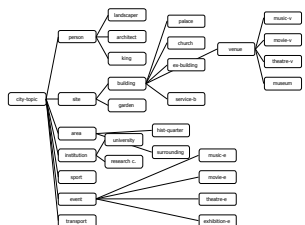
### The City-Tourism Ontology



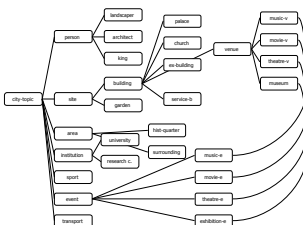
### The City-Tourism Ontology



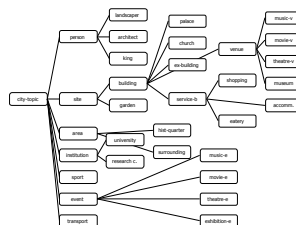
### The City-Tourism Ontology



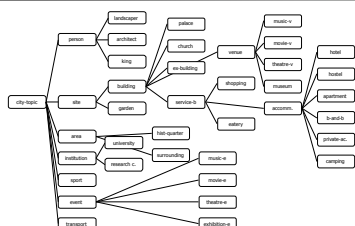
### The City-Tourism Ontology



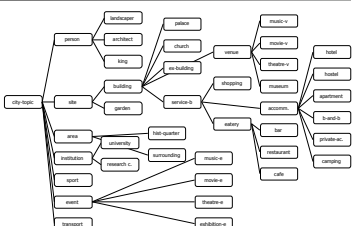
### The City-Tourism Ontology



### The City-Tourism Ontology



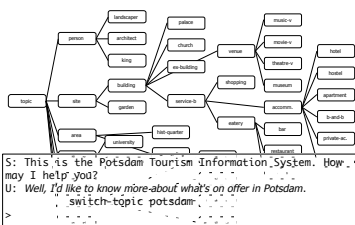
### The City-Tourism Ontology



### Ontology - Summary

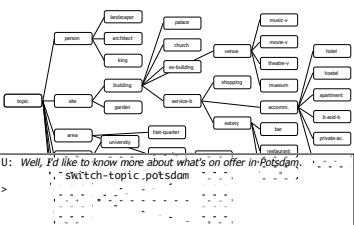
- Graph here: structured information about topic.
- Not stringent ontology, sub-topic relation semantically weak.
- Associated with instances of concepts are (canned) system utterances.
  - (Brand new: information about attributes like birthdays etc. is generated on the fly, using McRoy's YAG system)

### An Example



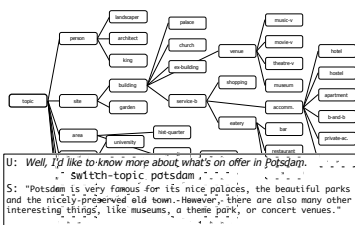
S: This is the Potsdam Tourism Information System. How may I help you?  
 U: Well, I'd like to know more about what's on offer in Potsdam.  
 . . . switch-topic potsdam . . .  
 >

### An Example



U: Well, I'd like to know more about what's on offer in Potsdam.  
 . . . switch-topic potsdam . . .  
 >

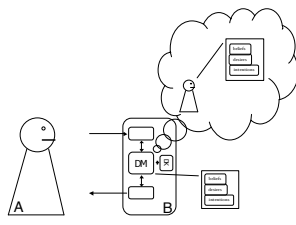
### An Example



U: Well, I'd like to know more about what's on offer in Potsdam.  
 . . . switch-topic potsdam . . .  
 S: "Potsdam is very famous for its nice palaces, the beautiful parks and the nicely-preserved old town. However, there are also many other interesting things, like museums, a theme park, or concert venues."  
 . . .

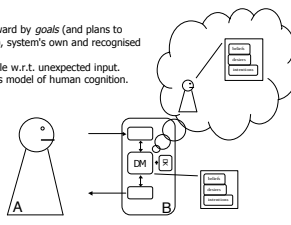


### Plan-based approaches



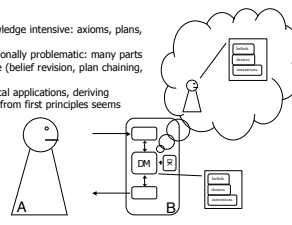
### Plan-based approaches

- 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.



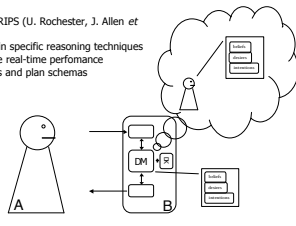
### Plan-based approaches

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



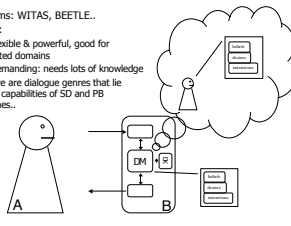
### Plan-based approaches

- TRAINS / TRIPS (U. Rochester, J. Allen et al.):
- use domain specific reasoning techniques to achieve real-time performance
  - 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.



### Part II - Models

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

### Information State Update, intro

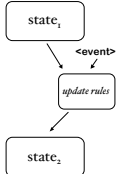
- Developed within the EU-project "Trindi" (Larsson 2003; Traum and Larsson 2004), integrating ideas from many previous projects.
- Not an approach *per se*: more an abstraction that allows different approaches to be compared, a *framework*.
- Main idea:
  - Dialogue models all have
    - a notion of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

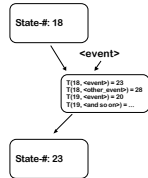
### Information State Update, intro

- Schematic view of update process:



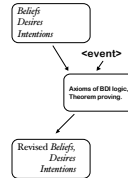
### Information State Update, intro

- 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, sys.

- Systems that have been implemented in this framework:
  - GODIS (Larsson 2001). Uses Ginzburg's "Questions under Discussion" structure; focus on shared vs. private information.
  - MIDAS (Bos 2000). Uses DRTs as part of IS. Theorem provers for updates.
  - EDIS (Mattheson, Poesio, Traum 2000). Focus on grounding and obligations.
  - ... RUDI (not actually in TRINDI-kit, but similar in spirit)

### 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, intro

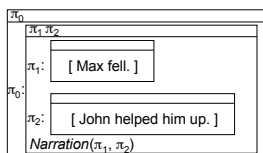
- works in domain of scheduling dialogues:
  - corpus available (VerbMobil)
  - grammar available (Stanford ERG, large scale HPSG)
- nicely restricted domain:
  - simple goal: agree on a time
  - simple plan: zoom in on time
  - utterances are either about good or about bad times.
  - finite number of bridging relations, conventionalised.

### RUDI, SDRT

- is a (partial) implementation of SDRT (Asher 1993, Asher & Lascarides 2003)
  - dynamic semantics + (AI-style) pragmatics
  - DRT + rhetorical relations (Hobbs 1985, Mann & Thompson 1987)
  - computes pragmatically preferred interpretation of discourse
  - central notion: *coherence*
  - models "pragmatic competence": more than what grammar outputs, less than full belief revision

### RUDI, SDRT

Logic of Content: FOL  
 "Max fell. John helped him up."



### RUDI, SDRT

Logic of Information Packaging: CSE

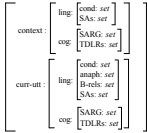
$$(?(\alpha, \beta) \wedge [some\ info..]) > R(\alpha, \beta)$$

### RUDI, SDRT

- Technical ingredients:
  - underspecified logical forms (descriptions of the form of logical formulae)
  - non-monotonic logics
- RUDI: resolving underspecification using discourse information
- Main hypothesis: resolving underspecification is a by-product of establishing *discourse coherence* (= computing rhetorical relations / speech acts)

## RUDI, information state

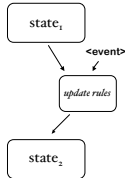
- RUDI's information state:



- dialogue moves: surface speech acts, Logical Forms (underspecified).

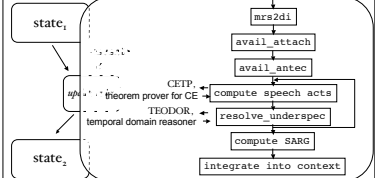
## RUDI, overview

- ... Using these LFs and the representation of the context, RUDI computes *update*, in several stages:



## RUDI, overview

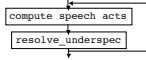
- ... Using these LFs and the representation of the context, RUDI computes *update*, in several stages:



## RUDI, 2 ways of resolving US

- RUDI resolves underspecification by...

- first trying to infer rhetorical relation, and using semantic constraints on the relation to resolve underspecification...
- ... if that fails, RUDI "guesses" resolution (i.e., compute specification of ULF), then infers relation using this additional information.



→ Information can flow in both directions..

## 2 ways of resolving underspec.

A: We should meet next week.

B: How about Friday?

$$\text{day\_of\_week}(x, \text{Fri}) \wedge B(x,y) \wedge y=? \wedge B=?$$

## 2 ways of resolving underspec.

A: We should meet next week.

B: How about Friday?

$$\text{day\_of\_week}(x, \text{Fri}) \wedge B(x,y) \wedge y=? \wedge B=?$$

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

## 2 ways of resolving underspec.

A: We should meet next week.

B: How about Friday?

$$\text{day\_of\_week}(x, \text{Fri}) \wedge B(x,y) \wedge y=? \wedge B=?$$

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

semantics of Q-Elab  $\beta$  is a question s.t. all answers elaborate on plan to reach goal of  $\alpha$

## 2 ways of resolving underspec.

A: We should meet next week.

B: How about Friday?

$$\text{day\_of\_week}(x, \text{Fri}) \wedge B(x,y) \wedge y=? \wedge B=?$$

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

semantics of Q-Elab  $\beta$  is a question s.t. all answers elaborate on plan to reach goal of  $\alpha$   
 $\Rightarrow \text{temp\_inc}(\text{SARG}_\alpha, t_\beta)$

## 2 ways of resolving underspec.

A: We should meet next week.

B: How about Friday?

$$\text{day\_of\_week}(x, \text{Fri}) \wedge B(x,y) \wedge y=? \wedge B=?$$

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

semantics of Q-Elab  $\beta$  is a question s.t. all answers elaborate on plan to reach goal of  $\alpha$   
 $\Rightarrow \text{temp\_inc}(\text{SARG}_\alpha, t_\beta)$

Inferring DS  $\Rightarrow$  Resolving US

## 2 ways of resolving underspec.

- A: We should meet next week.  
 B: I'm busy from the 24th until the 3rd.
- rejection or elaboration?

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	

## 2 ways of resolving underspec.

- A: We should meet next week.  
 B: I'm busy from the 24th until the 3rd.
- rejection or elaboration?

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	

Resolving US  $\Rightarrow$  Inferring DS

## Speech acts / Rhet. Rels in RUDI

- $Q\_Elab(\alpha, \beta)$ 
  - $\beta$  is a question; any possible answer elaborates a plan for achieving a SARG of  $\alpha$

## Speech acts / Rhet. Rels in RUDI

- $Q\_Elab(\alpha, \beta)$ 
  - $\beta$  is a question; any possible answer elaborates a plan for achieving a SARG of  $\alpha$
- $IQAP(\alpha, \beta)$ 
  - $\alpha$  is a question; questioner can infer direct answer from  $\beta$

## Speech acts / Rhet. Rels in RUDI

- $Q\_Elab(\alpha, \beta)$ 
  - $\beta$  is a question; any possible answer elaborates a plan for achieving a SARG of  $\alpha$
- $IQAP(\alpha, \beta)$ 
  - $\alpha$  is a question; questioner can infer direct answer from  $\beta$
- $Plan-Correction(\alpha, \beta)$ 
  - $\beta$  is proposition; speaker of  $\beta$  rejects a SARG of  $\alpha$

## Speech acts / Rhet. Rels in RUDI

- $Q\_Elab(\alpha, \beta)$ 
  - $\beta$  is a question; any possible answer elaborates a plan for achieving a SARG of  $\alpha$
- $IQAP(\alpha, \beta)$ 
  - $\alpha$  is a question; questioner can infer direct answer from  $\beta$
- $Plan-Correction(\alpha, \beta)$ 
  - $\beta$  is proposition; speaker of  $\beta$  rejects a SARG of  $\alpha$
- $Plan-Elaboration(\alpha, \beta)$ 
  - $\beta$  is proposition; elaborates plan to achieve SARG of  $\alpha$

## Speech acts / Rhet. Rels in RUDI

- $Q\_Elab(\alpha, \beta)$ 
  - $\neg(\alpha, \beta) \wedge \beta \text{?} > Q\_Elab(\alpha, \beta)$
  - all answers elaborate plan for achieving SARG of  $\alpha$ :  
 $Q\_Elab(\alpha, \beta) \rightarrow temp\_include(SARG_{\alpha, \beta})$

## Speech acts / Rhet. Rels in RUDI

- $Q\_Elab(\alpha, \beta)$ 
  - $\neg(\alpha, \beta) \wedge \beta \text{?} > Q\_Elab(\alpha, \beta)$
  - all answers elaborate plan for achieving SARG of  $\alpha$ :  
 $Q\_Elab(\alpha, \beta) \rightarrow temp\_include(SARG_{\alpha, \beta})$
- $IQAP(\alpha, \beta)$ 
  - $\neg(\alpha, \beta) \wedge \alpha \text{?} > IQAP(\alpha, \beta)$
  - questioner can infer direct answer to  $\alpha$  from  $\beta$ :  
 $IQAP(\alpha, \beta) \rightarrow temp\_overlap(SARG_{\alpha, \beta})$

## Speech acts / Rhet. Rels in RUDI

- $Plan-Correction(\alpha, \beta)$ 
  - $\beta$  is proposition; speaker of  $\beta$  rejects a SARG of  $\alpha$
  - $\neg(\alpha, \beta) \wedge \beta \wedge temp\_include(t_p, SARG_{\alpha, \beta}) > P-Corr(\alpha, \beta)$

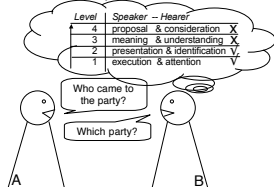
A: We should meet next week.  
 B: I'm busy from the 24th until the 3rd.

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	



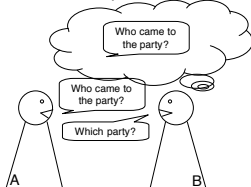
### Clarification Requests revisited

Dimension 1: Level of problem



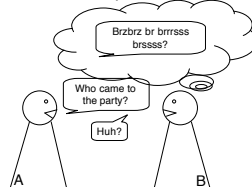
### Clarification Requests revisited

Dimension 2: Extent of problem



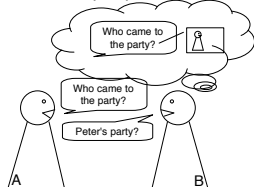
### Clarification Requests revisited

Dimension 2: Extent of problem



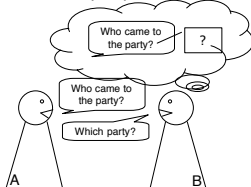
### Clarification Requests revisited

Dimension 3: Severity of problem



### Clarification Requests revisited

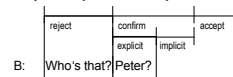
Dimension 3: Severity of problem



### Dimension "Severity"

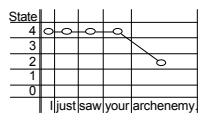
- quality of hypothesis / confidence in it needed at *all* levels of processing! (here: reference resolution)

A: I just met your archenemy.



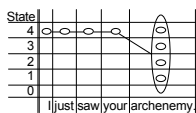
### Interdependencies

- not always in discrete state..



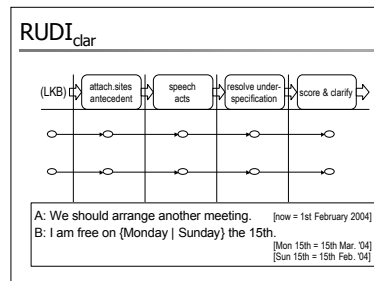
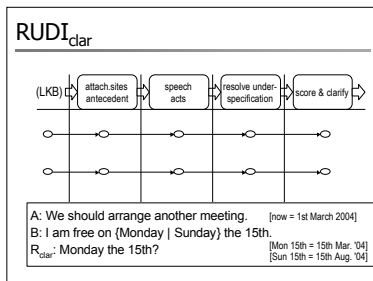
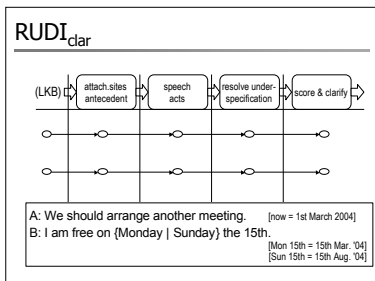
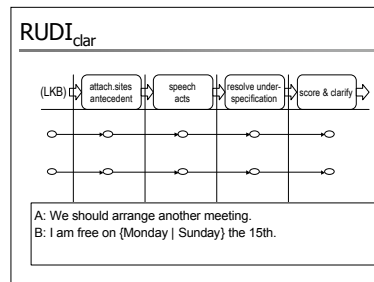
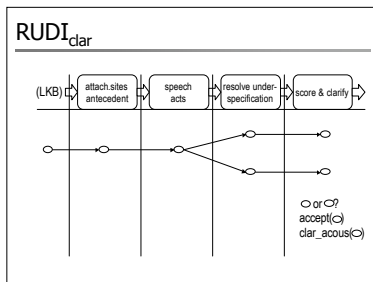
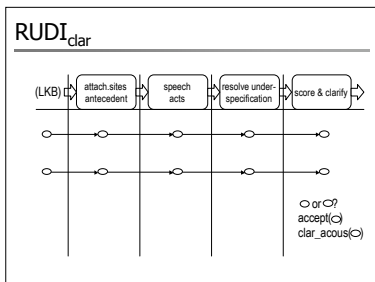
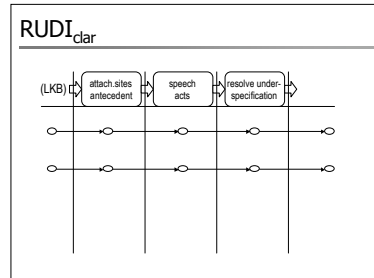
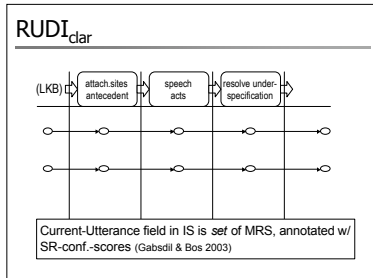
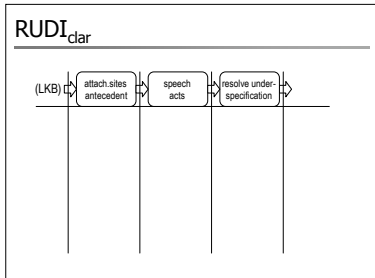
### Interdependencies

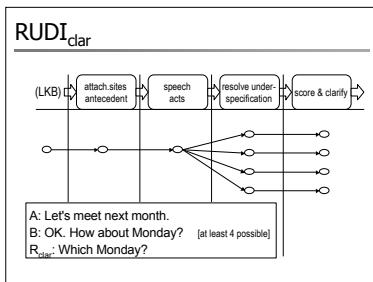
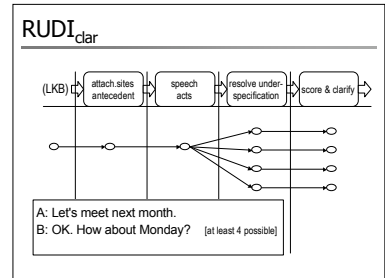
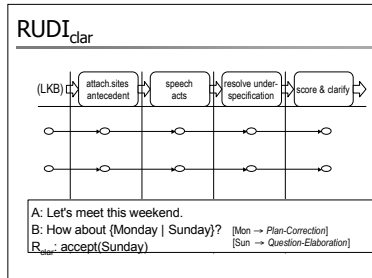
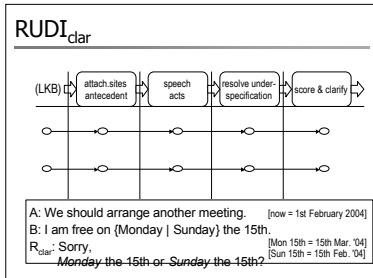
- snowballing: wrong hypothesis at one level will lead to dubious hyps. at higher levels.



### Interdependencies

- SR confidence + NLU features: (Duff, Gates & LuperFoy 1996), (Walker, Wright & Langkilde 2000), (Gabsdii 2004)
- generalised to other levels: (Paek & Horvitz 1999)
- these approaches: probabilistic. RUDI is symbolic.





- RUDI, summary**
- Main points:
    - Axioms in non-monotonic logic for inferring SAs / rhetorical relations;
    - Resolution of underspecification interdependent w/ computation of SAs.
  - Contextual aspects of interpretation:
    - bridging relations, resolving temporal expressions...
    - fragments.
  - When RUDI has problems understanding what it overhears, it can ask for help, using a fine-grained model of understanding problems.
  - ... using principled theories of pragmatics in dialogue systems seems not impossible...

- 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.

- Part II - Models**
- Phenomena ("Challenges"):
    - Spontaneous spoken language
    - Context sensitive interpretation
    - Interaction Management
  - Models ("Approaches"):
    - structured dialogue approaches
    - plan-based approaches
    - information state update-based approaches
  - Summary

- Summary**
- Two dialogue models / systems:**
- RUDI, implementation of theory of dialogue semantics / pragmatics (SDRT) [ joint work with Alex Lascarides (Edinburgh), Ann Copestake (Stanford / Cambridge) ]
  - PotBot, a tourism information system [ joint work with Manfred Stede (Potsdam) ]

- Summary**
- Classifying work on Dialogue**
- Motivation. *Why?*
  - Phenomena. *What?*
  - Approach. *How?*

## Summary

### Classifying work on Dialogue

- Motivation:
  - AI / cognitive science perspective: construe and test models of human behaviour
    - RUDI
  - NLP perspective: get a job done
    - PotBot, "information seeking chat"

## Summary

### Classifying work on Dialogue

- Phenomena covered:
  - Spontaneous spoken language
    - Syntax of utterances; disfluencies
  - Context sensitive interpretation
    - Anaphora; fragments; dialogue acts; gestures
  - Interaction Management
    - Turn taking; initiative; grounding

## Summary

### Classifying work on Dialogue

- Approach taken:
  - Structured dialogue (FSA, forms, etc.)
  - Plan-based systems
  - Information State Update

## Summary

### Some general points...

- There is not one correct approach to modelling dialogue: depends on what you want to do!
- However, using a common *framework* is good practice, as it makes models comparable. → *Information State Update* model.
- Perspectives (AI and NLP) are complementary, can learn from each other..

## The end

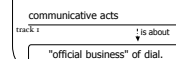
Thank you!

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

## Phenomena - Interact. Manag.

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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).



## Phenomena - Interact. Manag.

- Spontaneous spoken language
  - syntax of utterances
  - disfluencies
- Context sensitive interpretation
  - 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).

