What we can learn from Dialogue Systems that don't work

Dialogue Systems as Cognitive Models

David Schlangen
University of Potsdam, Germany

1 Introduction

There are different lines of research in the field of spo
ken dialog. Some researchers attempt at understanding
and possibly replicating, the mechanisms of human
dialog through linguistically motivated studies on hu-
man-human corpora. Others are interested in general
design principles that, once applied
are, in a way, orthogonal and complementary. The fo
cus of the first is on understanding human communica
tion, the second on designing the interface for usable
machines, and the third on building those usable ma-
chines. The topic of this paper is concerned with the

(Pieraccini & Huerta, 2005)

logue" research.) I distinguish two possible answers – the "engineering" (or "interface") view and the "simulation" view – representing the most extreme positions taken in response to this question. I then review an ar-

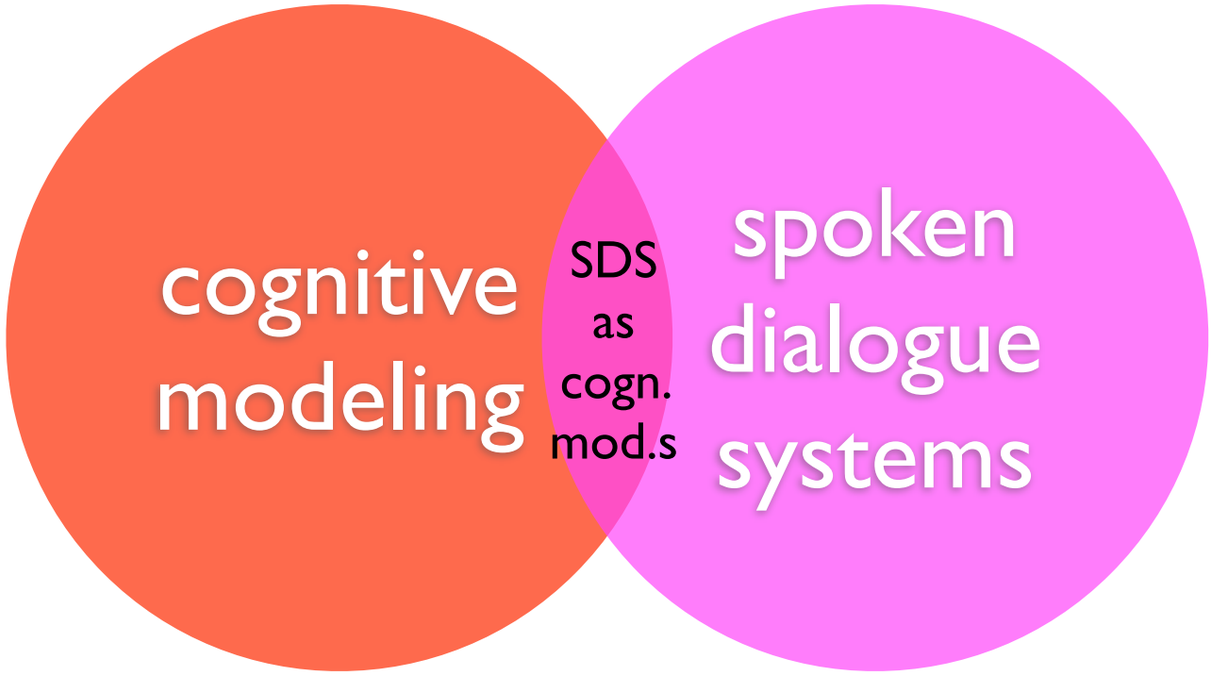
(Larsson, 2005)

2. Interfaces and interlocutors

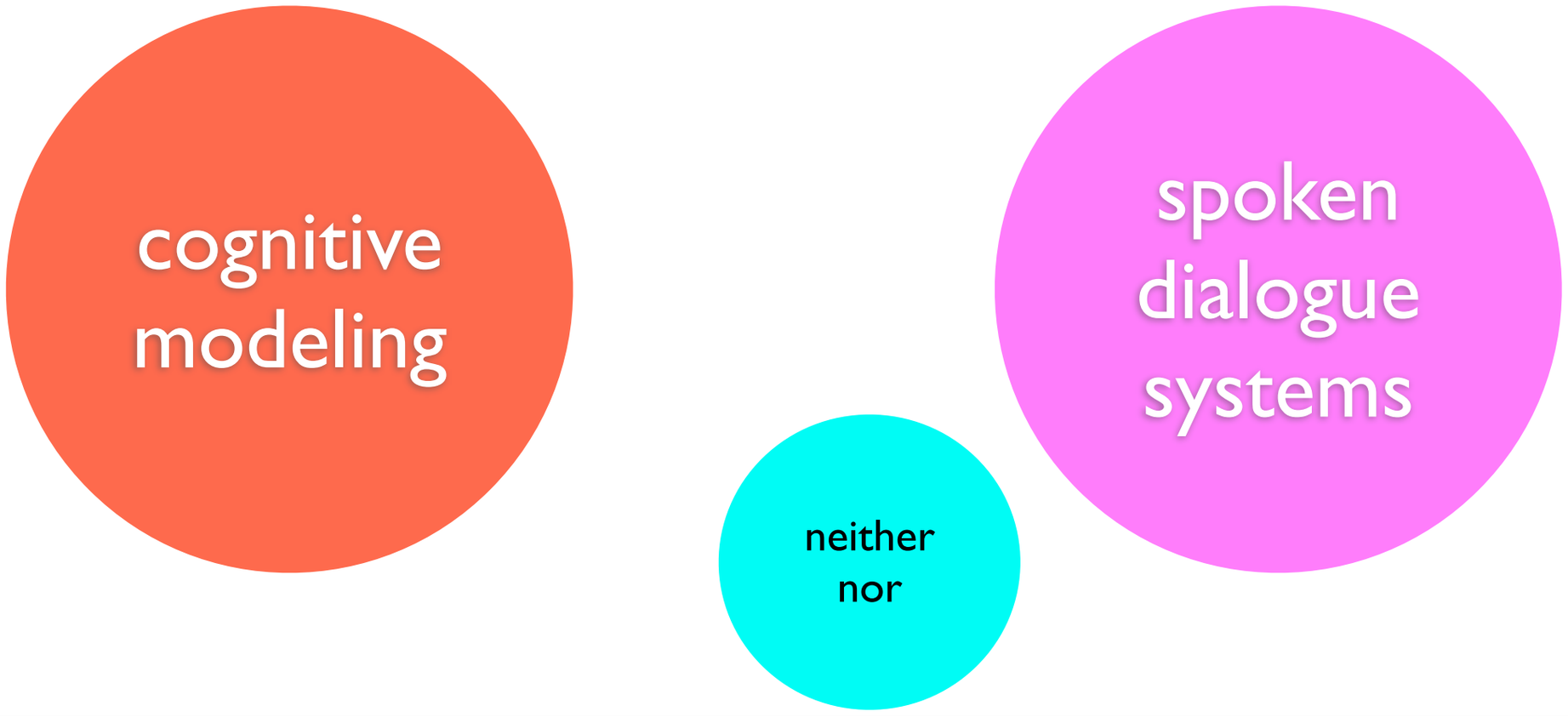
The two metaphors we will discuss here canbe described as follows. In ***the interface metaphor***, the spoken dialogue system is perceived as a machine interface – often, but not always, a computer interface. If the user in the first dialogue example below perceives the system this way, then saying "Lund" is equivalent to choosing Lund in a web page form or something similar. The spoken dialogue system is seen as an alternate means of interfacing with a computer. In ***the human metaphor***, the computer is perceived as a human-like

(Edlund, Heldner & Gustafson, 2006)

- is that a viable research goal?
- if so, what are some ways to approach it?


cognitive
modeling
SDS
as
cogn.
mod.s
spoken
dialogue
systems

or


cognitive
modeling
neither
nor
spoken
dialogue
systems
?

Overview

- Motivation
- Commercial, applied, and pure research on SDS
- Cognitive Modeling, and Dialogue
- SDS as Cognitive Models: Practical Considerations
  - Set Up, Evaluation
  - One objection: isn't this AI complete?
- More Objections
- Conclusions

	task	*method*	*human-likeness*
commercial systems	Comes first. Goal: Find efficient and effective solution.	Whatever fits specification.	If it doesn't hurt...
applied research	Free to choose. (But often inspired by potential for commercial use.)	Comes first. Explore methods that promise improvements.	If it doesn't hurt...
"pure" research	Free to choose. But how?	Possibly determined by theory... ... definitely influenced by what's possible.	Comes first. Explore what makes humans human-like.

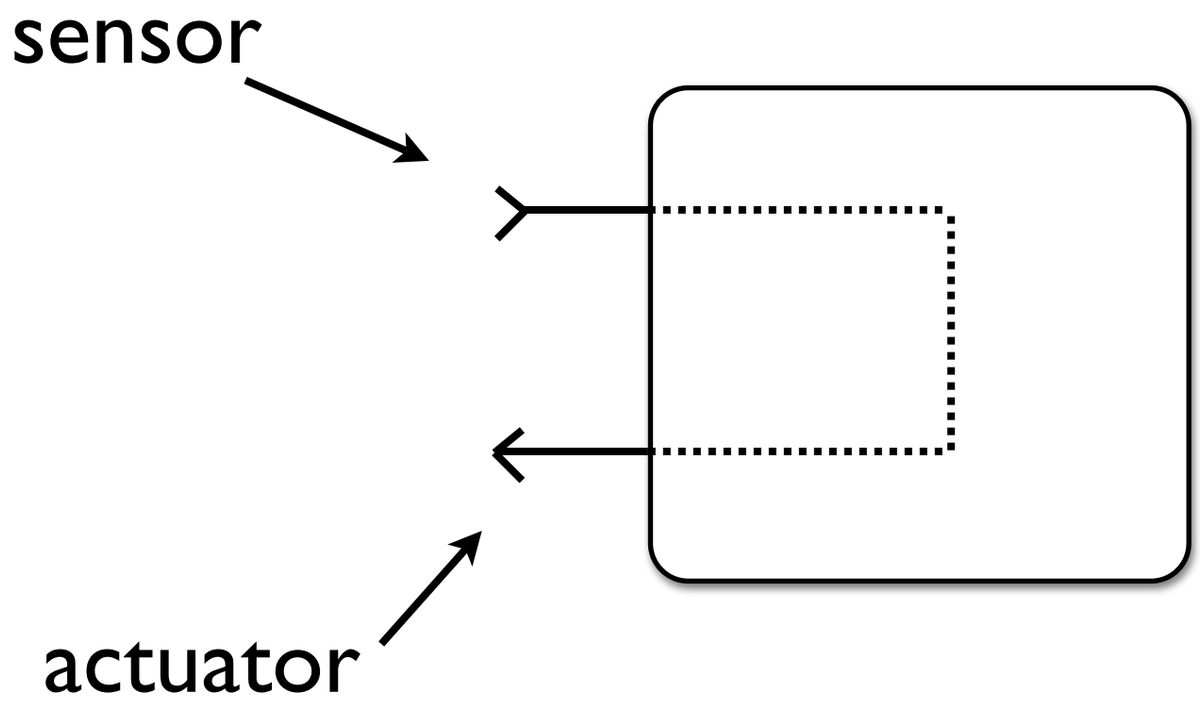
Overview

- Commercial, applied, and pure research on SDS
- Cognitive Modeling, and Dialogue
- SDS as Cognitive Models: Practical Considerations
  - one objection: isn't this AI complete?
- More Objections
- Conclusions

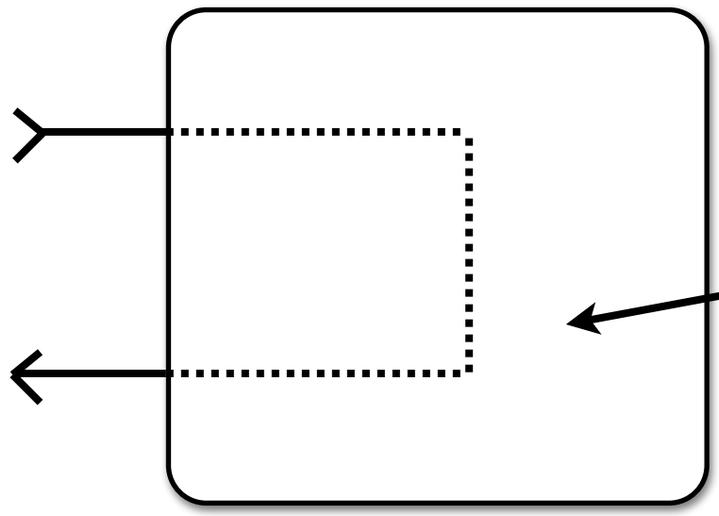
Background assumptions

- Participating in dialogue as *information processing task*
- Cognitive model = model of (human) information processor
- Dialogue System = Conversational Agent

Background assumptions



Background assumptions



cognitive model: specifies what's going on here

external constraints:

- functionality
- chronometric data
- eye movements
- error patterns
- learning rate, skill transfer

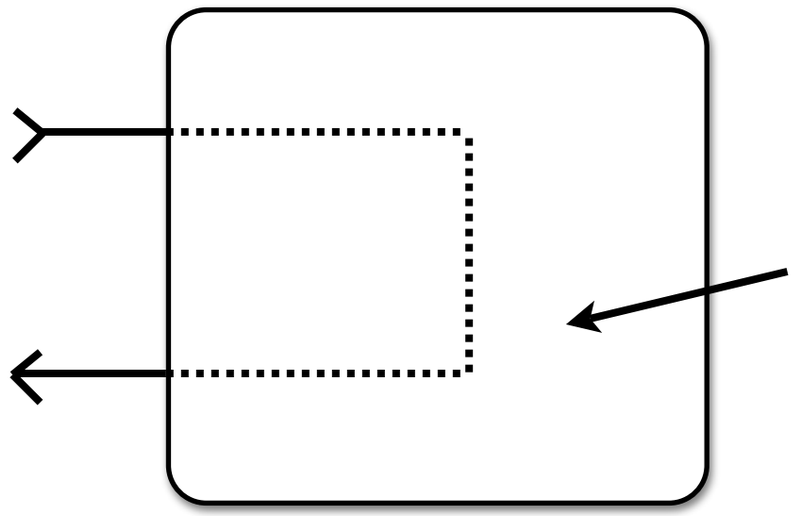
internal constraints:

- assume agent is *rational*, maximes expected return
- use independently motivated constraints (e.g. working memory)

status of model parts:

- computational model: structure of solution
- reprs. model
- implement. model

Dialogue Agents



- speech recognition
- parsing
- general reasoning, decision making, memory access, learning, ...
- language generation
- ....

external constraints:

- functionality
- chronometric data
- eye movements
- error patterns
- learning rate, skill transfer

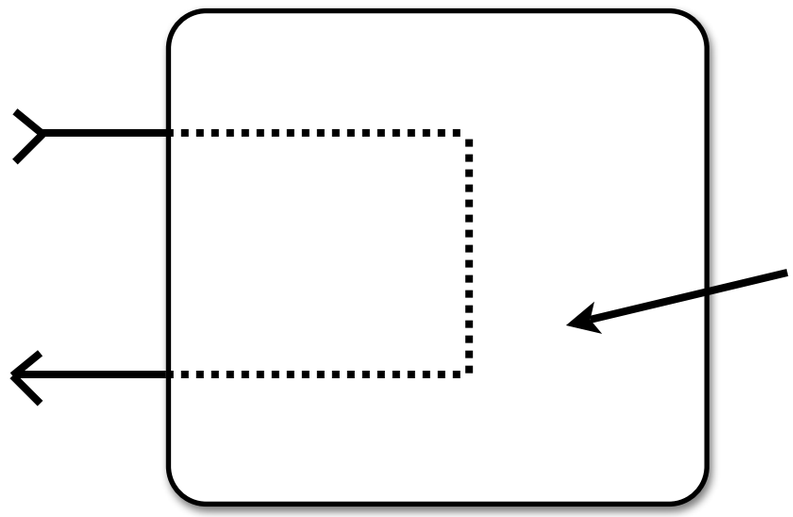
internal constraints:

- assume agent is *rational*, maximes expected return
- use independently motivated constraints (e.g. working memory)

status of model parts:

- computational model: structure of problem & solution
- reprs. model
- implement. model

Dialogue Agents



- speech recognition
- parsing
- general reasoning, decision making, memory access, learning, ...
- language generation
- ....

why model all this together?

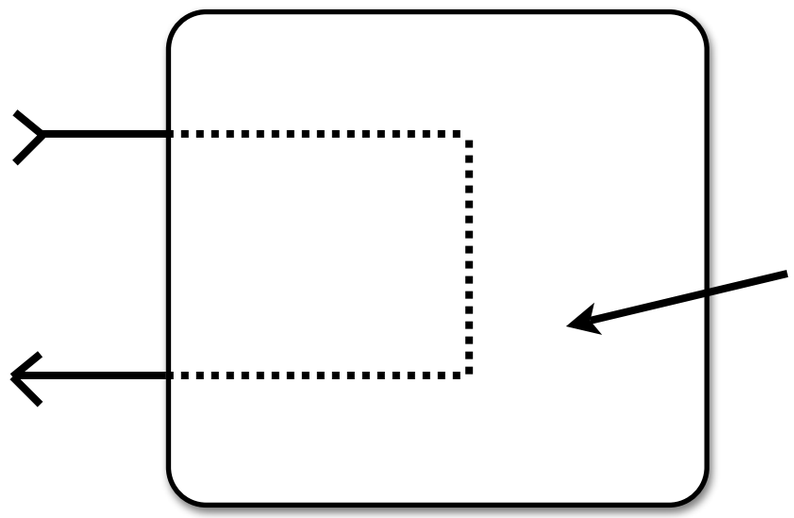
→ interactions between "modules"

why *implement* this?

→ real-time evaluation in ecologically valid context

SDS as *situated, implemented, computational whole-agent models*

Dialogue Agents



- speech recognition
- parsing
- general reasoning, decision making, memory access, learning, ...
- language generation
- ....

why model all this together?

→ interactions between "modules"

Model everything in same detail?
Which skills can be dissociated? (Dynamics of Interaction / Dynamics of Meaning)

Intermediate Summary

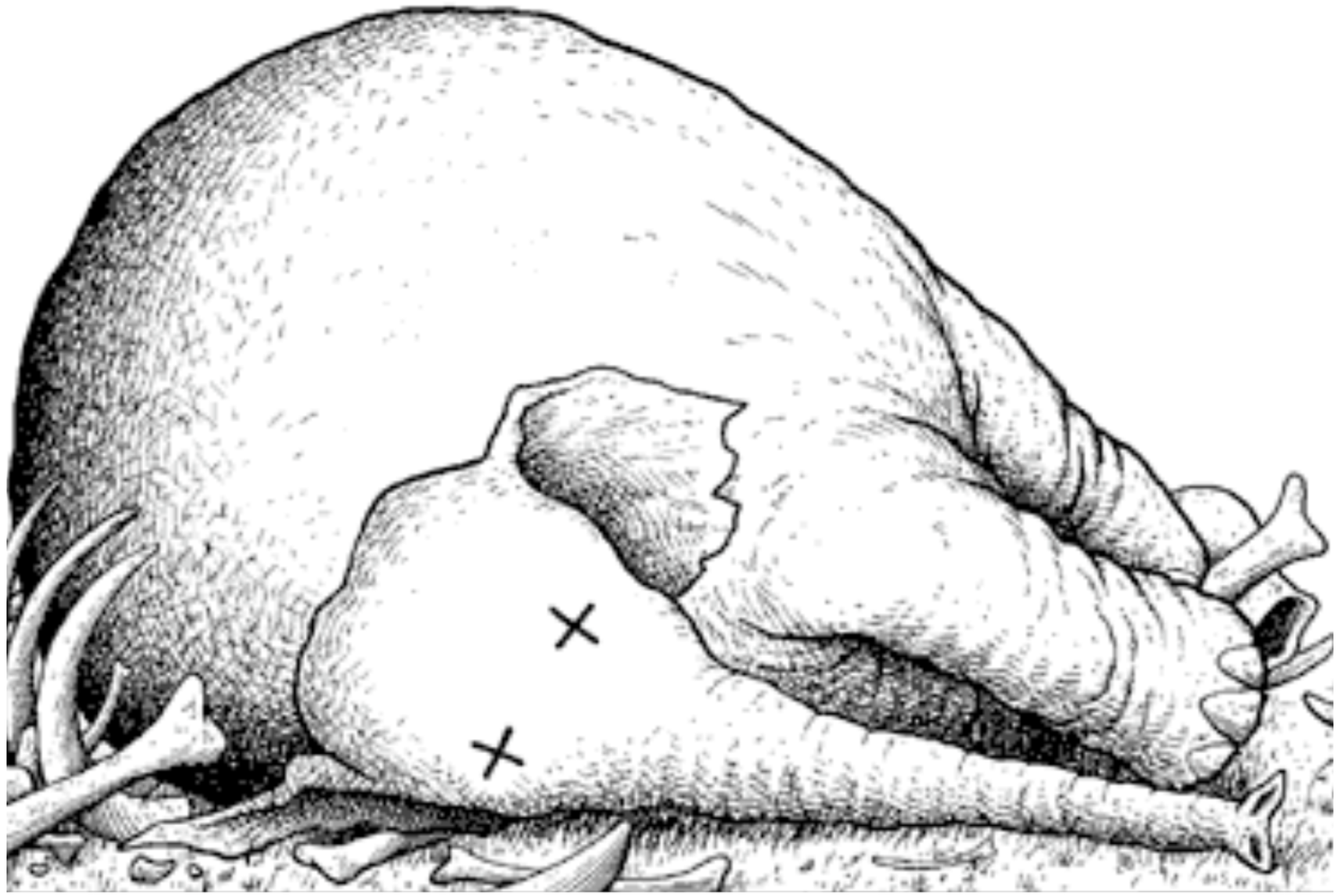
- Agent-based modeling forces one to be precise, and to try to answer interesting questions.
- SDS as *situated, implemented, computational whole-agent models*

Overview

- Motivation
- Commercial, applied, and pure research on SDS
- Cognitive Modeling, and Dialogue
- SDS as Cognitive Models: Practical Considerations
  - Set up, Evaluation
  - One objection: isn't this AI complete?
- More Objections
- Conclusions

the elephant in the room....

ASR!!



One solution: forget ASR, use typed input.
Strategy if focus on dynamics of content.
Downside: loss of situatedness, ecological validity

Other solution: restrict domain.
Strategy if focus on dynamics of interaction.
Downside: loss of semantic flexibility?

Methodology

for given phenomenon:

1. derive initial model from theory
2. devise setting that restricts phenomenon as little as possible, and rest as much as possible
3. collect human-human data
4. improve model
5. implement
6. evaluate

Methodology

for given phenomenon:

1. derive initial model from theory
2. devise setting that restricts pheno-menon as little as possible, and rest as much as possible
3. collect human-human data
4. improve model
5. implement
6. evaluate

e.g.

- entrainment / alignment
- turn-taking, feedback,
- repair
- adaptation, learning of personal ontologies
- argumentation
- coherence
- ...

Methodology

for given phenomenon:

1. derive initial model from theory
2. devise setting that restricts phenomenon as little as possible, and rest as much as possible
3. collect human-human data
4. improve model
5. implement
6. evaluate

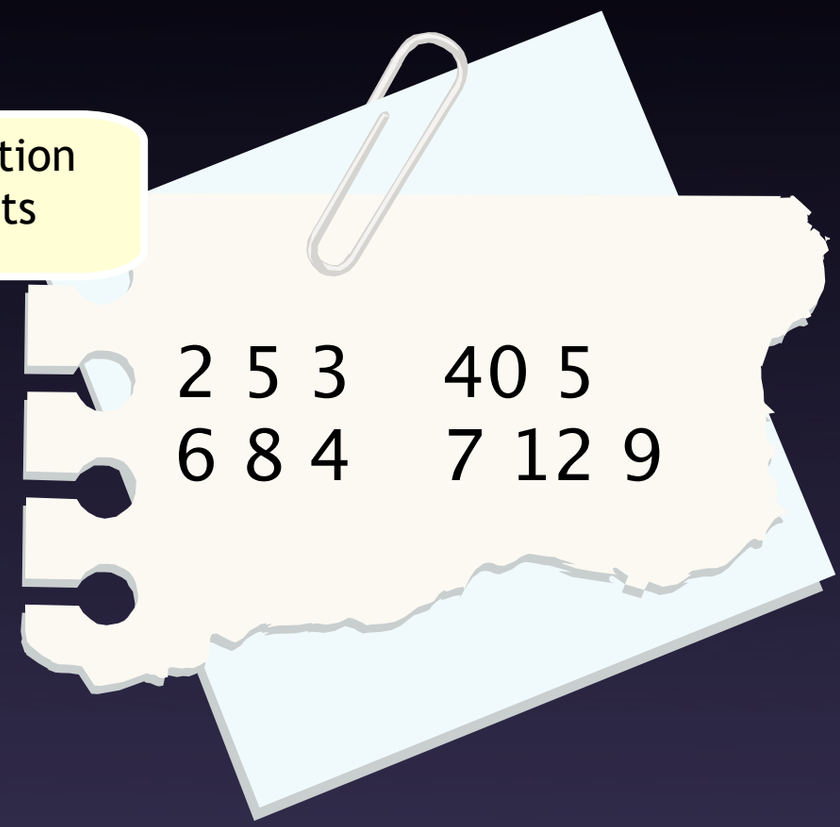
- micro-domains (Edlund et al. 2008)

Example dialogue

U.01: two five three -
S.02: okay
U.03: forty five -
S.04: sorry, do you mean four five?
U.05: no, forty … five
S.06: okay, four oh five
U.07: yeah, and six eight four -
S.08: okay
U.09: seven twelve nine.
S.10: so, from the start, two five three -
U.11: yeah
S.12: four oh five -
S.13: six six four -
U.14: no, six eight four
S.15: okay, six eight four -
S.16: and then seven one two nine.
U.17: that's right

"Syntactic ambiguity"

Clarification requests



Cue phrases

Misrecognition and Correction

(Skantze & Schlangen 2009, EACL)

Methodology

for given phenomenon:

1. derive initial model from theory
2. devise setting that restricts phenomenon as little as possible, and rest as much as possible
3. collect human-human data
4. improve model
5. implement
6. evaluate

- micro-domains (Edlund et al. 2008)
- oracles: system knows what user tries to do
- WOz

Methodology

for given phenomenon:

1. derive initial model from theory
2. devise setting that restricts phenomenon as little as possible, and rest as much as possible
3. collect human-human data
4. improve model
5. implement
6. evaluate

- record human-machine dialogues, and compare objective measure to corpus
- subjective evaluation; perhaps with variants of system
- Particularised Turing Test: can system be distinguished by judge from human working under same constraints?

Overview

- Motivation
- Commercial, applied, and pure research on SDS
- Cognitive Modeling, and Dialogue
- SDS as Cognitive Models: Practical Considerations
  - Set up, Evaluation
  - One objection: isn't this AI complete?
- More Objections
- Conclusions

Objection:
AI-completeness of dialogue

- free, unrestricted conversation requires full intelligence (Turing 1950)

- free, unrestricted conversation requires "background knowledge that is not formalizable" (Larsson 2005)

- but if "dissociation hypothesis" is correct, you can identify sub-tasks of dialogue that do not require full intelligence, and can study those.

Summary

- aiming for "human-likeness" as primary goal is viable
- will tell us a lot about constraints on human language processor
- increasing level of detail: from functional to module level, to algorithmic / representational...

  from modeling behaviours, to modeling principles behind behaviours
- goal of doing something useful with SDS (= transferability) may suffer

thanks.

funded by: