Differences between human-human and human-computer dialogues, are there any?

Susanne Horst
Hauptseminar Dialogue Managment (Ivana Kruijff-Korbayová, SS 2002)
http://www.coli.uni-sb.de/~korbay/Lehre/DM-beschreibung.html

Revision: 30. September 2002

1 Introduction

There is relatively little material on differences between human-human and human-computer dialogues. The following gives an overview about what I have found out so far. There are some main areas of research in dialogue management which seem to be relevant: intentions of the dialogue partners, cues, which control the dialogue structure and initiative also a part of the dialogue structure.

2 Intentions

2.1 Definition

Kreutel and Matheson state that in the following example

```
A: Did Pete drive here?
B: I don't know/I don't want to talk about that.
```

intention-based dialogue models fail to explain why B actually responds to A's question.

They propose a set of inference rules which allows the construction of an intentional structure reflecting the relationships between the DP's intentions at a given state of a dialogue and see intentions as objects that are individualised with respect to the dialogue partners. The content of intentions is considered as interactive goals to be achieved in the course of a discourse and is associated with assertions and questions. Goals are assumed to be shared_belief(p) and resolved(q) where p and q are the contents of the assertion and the question, respectively.

2.2 Rules

They propose the following rules associating intentions with dialogue acts where I is a set of the DP's intentions and i is a member of I.

- assert: update I with i: shared_belief(p) where p is the content of the assertion
- ask: update I with i: resolved(q) where q is the content of the question
- answer: update I with i: resolved(q) where q is the content of the question the answer replies to
- info_request: update I with i: resolved(q) where q is the content of the question the argumentation act replies to

2.3 Inference rules

In order to determine the set of satisfied intentions (sat) they assume the following inference rules which add information to the DH. We define an intention as satisfied if its content occurs in the DH:

- accept: infer that shared_belief(p) where p is the content of the assertion accepted
- accept_answer: infer that resolved(q) where q is the content of the question that is resolved

3 Cues

3.1 Information about discourse structure

Previous research has identified by which speakers signal information about discourse structure to listeners: Cue words and phrases, intonations, pronominalisation.

3.2 Cue words

Whittaker and Stenton argue that in the cue words approach it was previously assumed that phrases like “because”, “so”, and “but” offer explicit information to listeners about how the speaker's current content relates to previous content. For example a speaker might use the expression “so” to signal that s/he is about to conclude what s/he has just said. An example would be that “and” or “but” signal to the listener that a new topic and set of referents is being introduced whereas "“anyway” and “in any case” indicate a return to a previous topic and reference set. A second way of signalling discourse structure is intonation when for example new topics are being signalled by changes in intonational contour. A finer more indirect cue to discourse structure is the speaker's choice of referring expressions and grammatical structure.

The problem with this approach is that the cue may only be an infrequent indicator of a particular type of shift. If a general theory of discourse is wanted then it is necessary to know about the whole range of cues serving this function. Their study therefore takes a different approach. They begin by identifying all shifts of control in the dialogue and then look at how each shift was signalled by the speakers.

3.3 Data

The data were recordings of telephone conversation between clients and an expert concerning problems with software. The tape recordings from four dialogues were then described and the analysis conducted on the typewritten transcripts rather than the raw recordings. There was a total of 450 turns in the dialogues.

3.4 Criteria for classifying utterance types

- Assertions: declarative utterances which were used to state facts or yes or no answers
- Commands: utterances which were intended to instigate action in their audience e.g.

  `''What I would do if I were you is to relink X''`

- Questions: utterances which were intended to elicit information from the audience
- Prompts: These were utterances which did not express propositional content. Examples of prompts where things like "Yes" and "Uhu".

They then went on to analyse how control was exchanged between participants at the boundaries of these phases. There was a total of 56 shifts of control over the 4 dialogues and they identified 3 main classes of cues used to signal control shifts. These were prompts, repetitions and summaries. They also looked at what happened when no signal was given (interruptions).

3.5 Example for prompts

```
1. E: ''And they are, in your gen you'll find that they've relocated into the
       labelled common area'' (E control)

2. C: ''That's right.'' (E control)
```

```
3. E: ‘‘Yeah’’

(E abdicates control with prompt)
```

3.6 Example for repetitions

```
1. Client: "These routines are filed as DS" (C control)

2. Expert: ‘‘That’s right, yes’’ (C control)

3. C: ‘‘DS’’ (C abdicates control with repetition)
```

Half the repetitions were accompanied by cue words. These were “and”, “well” and “so” which prefixed the assertion.

3.7 Summaries

In the dialogues the summaries seemed to be characterised by the concise reference to objects or entities which had earlier been described in detail, e.g.

```
(a) ‘‘Now, I’m wondering how the two are related’’’
```

in which “the two” referes to the two error messages which it had taken several utterances to describe previously. The other characteristic of summaries is that they contrast strongly with the extremely concrete descriptions elsewhere in the dialogues, e.g.

```
"err the system programm standard call file does not have a tail record"
```

followed by

```
‘‘And I’ve no clue at all how to get out of the situation".
```

Chu-Carroll and Brown describe explicit cues, which include explicit requests by the speaker to give up or take over the initiative and discourse cues which includes cues that can be recognized using linguistic information, such as the surface form of an utterance, and the recognized intentions of the utterances, such as how the current utterance relates to prior discourse and analytical cues which include cues that cannot be recognized without the hearer performing an evaluation of the speaker’s proposal using the hearer’s private knowledge. After the evaluation, the hearer may find the proposal invalid, suboptimal or ambiguous.

4 Initiative

Walker and Whittaker state that conversation between two people is usually of mixed-initiative with control over the conversation being transferred from one person to another. They apply a set of rules for the transfer of control to 4 sets of dialogues consisting of a total of 1862 turns. The application of the control rules lets them derive a domain-independent discourse sturcture. The derived structure indicate that initiative plays a role in the structuring of discourse.

Conversation between two people has a number of characteristics that have yet to be modeled adequately in human-computer dialogue. Conversation is bidirectional, information is exchanged by mixed-initiative, conversational partners not only respond to what others say, but feel free to volunteer information that is not requested and sometimes ask questions of their own.

From a practical perspective, there is ample evidence that limited mixed-initiative has contributed to lack of system usability.

The application of the control rules to these dialouges lets them derive domain-independent discourse segemnts with each segment being controlled by one or other discourse particiapant.

4.1 Utterance types

- assertions
- commands
- questions
- prompts

The paper did not make it quite clear if these are controlled by the speaker or the hearer.

4.2 Control shift types

- abdication:

```
``Okay, go on.''
```

- repetition/summary:

```
``That would be my recommendation
  and that will ensure that you get a logically integral set of files.''
```

- interruption:

```
``It is something new tough um.''
```

Abdications correspond to those cases where the controller produces a prompt as the last utterance of the segment. The class repetition/summary corresponds to the controller producing a redundant utterance and the controller supplies no new propositional content. Interruptions characterize shifts occuring when the noncontroller displays initiative by seizing control. This classification suggests that the transfer of control is often a collaborative phenomenon.

4.3 Application example

TRIPS is quite limited in the form of mixed-initiative interaction it supports. It supports taking discourselevel initiative, but does not allow shifting of task-level initiative during the interaction. The reason is that system behavior is driven by the dialogue manager, which focuses on interpreting user input. This means that the system's own independent goals are deemphasized. The behavior of a conversation agent should ideally be determind by three factors, not just one: the interpretation of the last user utterance (if any), the agent's own persistent goals and obligations, and exogenous events of which it becomes aware.

4.4 Proposals

Chu-Carroll and Brown state that in order for a dialogue system to successfully collaborate with its user on their interaction, it must be able to dynamically track initiative shifts during their interaction by recognizing the user's cues for initiative shifts, and by providing appropriate cues, when necessary, in its responses to user utterances.

Most existing models of initiative conflate two types of initiatives, which they call task initiative and dialogue initiative, and they show how distinguishing between these two types of initiatives accounts for phenomena in collabaorative dialogues that previous models were unable to explain. They find three cases: having neither task nor dialogue initiative, having dialogue but not task initiative and having both task and dialogue initiatives.

An agent is said to have the task initiative if she is directing how the agents task should be accomplished, i.e. if her utterances directly propose actions that she believes the agent(s) should perform, while an agent has the dialogue initiative when both agents expect the agent to communicate next.

Jordan and Siler define any utterance that provides new unsolicited information as a show of StudentInitiative. StudentInitiative was annotated based on whether the student did more than respond in the expected way relative to the category assigned to the tutor question.

- closed question from the tutor

  ```
  ''What is the direction of the force experienced by the car
  when it is hit from behind?''
  ```

- student response with no show of StudentInitiative

  ``‘‘the direction is towards the car that hit it.’’``

- open-ended question from the tutor

  ``‘‘Why should the neck get hurt?’’``

- closed question from the tutor

  ``‘‘What is caused by force?’’``

- student response with StudentInitiative

  ``‘‘acceleration, so I should mention force and acceleration instead of speed’’``

They wanted to achieve a more comprehensive definition of initiative and encourage students taking an active part in the learning process but they did not show any realworld data to confirm that they actually succeeded.

4.5 Conclusion

It appears that there are differences in dialogue act patterns between the human-humand and human-computer dialogues, but it seems to be too early to conclude anything about any relation between dialogue acts and user satisfaction.

Literatur

[1] John Aberdeen, Christina Doran, The MITRE Corporation, Human-computer and human-human dialogues

[2] Pamela W. Jordan, Stephanie Siler Student initiative and questioning strategies in computer-mediated human tutoring dialogues

[3] Jennifer Chu-Carroll, Michael K. Brown An evidential model for tracking initiative in collaborative dialogue interactions

[4] Steve Whittaker, Phil Stenton, Cues and control in expert-client dialogues

[5] Marilyn Walker, Steve Whittaker, Mixed initiative in dialouge: An investigation into discourse segmentation (Appeared in ACL90)

[6] Joern Kreutel, Colin Matheson, Obligations, intentions and the notion of conversational games

[7] David R. Traum, James F. Allen, Discourse obligations in dialogue processing

[8] The Trindi Book, The Trindi Consortium May 2001