Language Phenomena in Tutorial Dialogs on Mathematical Proofs

Christoph Benzmüller[1], **Armin Fiedler**[1], **Malte Gabsdil**[2], **Helmut Horacek**[1], **Ivana Kruijff-Korbayová**[2], **Dimitra Tsovaltzi**[2], **Bao Quoc Vo**[1], **Magdalena Wolska**[2]
[1]Fachrichtung Informatik [2]Fachrichtung Computerlinguistik,
Universität des Saarlandes, Postfach 15 11 50, D-66041 Saarbrücken, Germany
{chris,afiedler,horacek,bao}@ags.uni-sb.de, {gabsdil,korbay,dimitra,magda}@coli.uni-sb.de

1 Introduction

Empirical findings (Moore, 1993) show that flexible natural language interaction promotes active learning. In the DIALOG project [1] (Pinkal et al., 2001), we are investigating and modelling semantic and pragmatic phenomena in tutorial dialogs focused on problem solving skills in mathematics. The language used in this domain is particularly challenging because of the mixture of natural language and formal symbolic expressions. We have collected a corpus of dialogs with a simulated tutoring system for proofs in naive sets theory. Corpus investigation into the use of natural language in mathematics enables us to identify the linguistic phenomena that will impose specific requirements on dialog management. The goal of the project is to develop a prototype system gradually embodying the results of our analyses.

2 Empirical Study

To collect a corpus of tutorial dialogs, we have conducted a *Wizard-of-Oz (WOz)* experiment (Benzmüller et al., 2003a). 24 subjects with varying educational backgrounds, and little to fair prior mathematical knowledge participated in the experiment which consisted of 3 phases: (1) *preparation and pre-test* (on paper), (2) *tutoring session* (mediated by the *WOz* tool (Fiedler and Gabsdil, 2002)), (3) *post-test and evaluation questionnaire* (on paper). Subjects were asked to prove 3 theorems[2]: (a) $\overline{(A \cup B) \cap (C \cup D)} = (\overline{A} \cap \overline{B}) \cup (\overline{C} \cap \overline{D})$; (b) $A \cap B \in \wp((A \cup C) \cap (B \cup C))$; and (c) When $A \subseteq \overline{B}$, then $B \subseteq \overline{A}$. The subjects were instructed to enter proof steps, rather than complete proofs at once, to encourage dialog with the system. The wizard's task was to respond to the student's utterances following a given algorithm (Fiedler and Tsovaltzi, 2003): first evaluate their completeness, accuracy, and relevance with respect to a valid proof, then select the next dialog move and verbalize it. The subjects and the wizard were free to mix natural and formal language. The dialogs were carried out in German, and had between 1 and 15 student turns; 5 on average.

3 Language phenomena in utterances

Figure 1 shows example responses; only English translations are included for brevity. We have observed that mathematical and natural language expressions are tightly interleaved, and the language is telegraphic: sentences are fragmented, symbols are used to name relations. We have identified recurring structural and semantic phenomena, examples of which we present below.

Structural phenomena, firstly, concern parsing the interleaved symbolic and natural language ("A also $\subseteq$ B") and identifying the status of mathematical expressions (considering scoping issues, as in "B contains no $x \in A$"). Secondly, they concern recognizing typical patterns expressing derivation steps ("If..., then...", "So..."), and possible structural ambiguities in cases such as coordination ("[$x \in B$] and so [$x \subseteq \overline{B}$ and $x \subseteq \overline{A}$] given

[1] DIALOG is part of the Saarland University Collaborative Research Center on *Resource-Adaptive Cognitive Processes* (SFB 378) (http://www.coli.uni-sb.de/sfb378/)

[2] $\wp$ stands for power set

(1)	A power set contains all subsets, hence also $(A \cap B)$
(2)	**T:** ...$\overline{(A \cup B) \cap (C \cup D)} = \overline{A \cup B} \cup \overline{C \cup D}$... **S:** de Morgan rule 2 applied to both complements
(3)	de Morgan rule 1 also holds for $\overline{C \cup D}$ de Morgan rule 2 means $\overline{A \cap B} = \overline{A} \cup \overline{B}$. In this case e.g. $\overline{A} =$ the term $\overline{A \cup B}$ $\overline{B} =$ the term $\overline{C \cup D}$. So $\overline{(A \cup B) \cap (C \cup D)} = (\overline{A} \cap \overline{B}) \cup (\overline{C} \cap \overline{D})$
(4)	$(A \cup B)$ must be in $\wp((A \cup C) \cap (B \cup C))$, since $(A \cap B) \in (A \cap B) \cup C$
(5)	$B \cup A \subseteq C B \cup A \subseteq D$. If A is a subset of C and B a subset of C, then both sets together must also be a subset of C. The same holds for D. $\overline{C \cap D} \cup \overline{A \cap B}$ applying the de Morgan rules.

Figure 1: Examples of students' utterances.

[the assumption]" vs. "[[$x \in B$] and so [$x \subseteq \overline{B}$]] and [[$x \subseteq \overline{A}$] given [the assumption]])" [3].

Semantic phenomena mainly involve imprecise, ambiguous, and/or informal references to domain relations, or formally incorrect, but metaphorically interpretable references. For example, "must be in" in (4) and "contains" in (1), can express the relation *element_of* or *subset_of*; "have common elements" refers to the non-empty intersection; "both sets together" in (5) may be interpreted as union or intersection; the relation "=", used mainly in equations or to indicate a value assignment, has been also used for term substitution in an axiom (in (3), $\overline{A \cup B}$ is to be substituted for $\overline{A}$, and $\overline{C \cup D}$ for $\overline{B}$). Another class of phenomena are references to structural parts of terms and formulae such as "the left side" or "the inner parenthesis" which are incomplete specifications: the former refers to a part of an equation, the latter, metonymic, to an expression enclosed in parenthesis. Moreover, structural parts of formulae constrain the search space for anaphor resolution (e.g. the left- and right-hand side of the equation in (2) for resolving "both complements"). Finally, generic and specific references can appear within one utterance (cf.(1), "a power set" is a generic reference, whereas "$A \cap B$" is a reference to a subset of a specific power set).

4 Ongoing work

Presently, we are implementing a formulae parser capable of (a) identifying formulae, (b) analysing their structural parts. Formulae analysis is a preprocessing step toward a semantic interpretation of utterances. Here, we are pursuing two approaches: (1) shallow recognition of derivation steps based on recurring keywords used to introduce them, and (2) syntactic and semantic analysis using a lexically-based grammar[4]. The input analysis is embedded within the context of implementing an Information State-based dialog manager that is to interact with a mathematical assistant in conducting a tutoring session (Benzmüller et al., 2003b).

[3] Square brackets illustrate alternative structural grouping of coordinated constituents

[4] We are using an open source CCG parser and realizer (http://openccg.sourceforge.net/)

References

C. Benzmüller, A. Fiedler, M. Gabsdil, H. Horacek, I. Kruijff-Korbayová, M. Pinkal, J. Siekmann, D. Tsovaltzi, B. Vo, and M. Wolska. A Wizard-of-Oz Experiment for Tutorial Dialogues in Mathematics. *Proc. of the AIED Workshop on Advanced Technologies for Mathematics Education*, Sydney, Australia, 2003a.

C. Benzmüller, A. Fiedler, M. Gabsdil, H. Horacek, I. Kruijff-Korbayová, M. Pinkal, J. Siekmann, D. Tsovaltzi, B. Vo, and M. Wolska. Tutorial Dialogs on Mathematical Proofs. *Proc. of the IJCAI Workshop on Knowledge Representation and Automated Reasoning for E-Learning Systems*, Acapulco, Mexico, 2003b.

A. Fiedler and M. Gabsdil. Supporting Progressive Refinement of Wizard-of-Oz Experiments. *Proc. of the 6th Inter. Conference on Intelligent Tutoring Systems Workshop on Empirical Methods for Tutorial Dialogue Systems*, San Sebastian, Spain, 2002.

A. Fiedler and D. Tsovaltzi. Automating Hinting in Mathematical Tutorial Dialogue. *Proc. of the EACL-03 Workshop on Dialogue Systems*, Budapest, Hungary, 2003.

J. Moore. What makes human explanations effective? *Proc. of the 15th Annual Meeting of the Cognitive Science Society*, Hillsdale, NJ, 1993.

M. Pinkal, J. Siekmann, and C. Benzmüller. Projektantrag Teilprojekt MI3–DIALOG:Tutorieller Dialog mit einem mathematischen Assistenzsystem. *Fortsetzungsantrag SFB 378–Ressourcenadaptive kognitve Prozesse*, Saarbrücken, Germany, 2001.