# Discourse 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,siekmann,bao}@ags.uni-sb.de, {gabsdil,korbay,pinkal,dimitra,magda}@coli.uni-sb.de

## 1 Introduction

Dialogs about problem solving in mathematics are characterized by a mixture of telegraphic natural language text and embedded formal expressions. Behaving adequately in such an environment is extremely important for tutorial systems, as follows from Moore's empirical findings which show that flexible natural language dialog is needed to support active learning [3]. Motivated by the lack of empirical data for such kind of conversations, we have collected a corpus of dialogs with a simulated tutoring system for teaching proofs in naive set theory, to identify genre-specific variants of linguistic phenomena which impose specific requirements on natural language dialog management. This work is embedded in the DIALOG project [1] [4] whose goal is (i) to empirically investigate the use of flexible natural language dialog in tutoring mathematics, and (ii) to develop an experimental prototype system gradually embodying the empirical findings. The experimental system will engage in a dialog in written natural language (and later also in multimodal forms of communication based on diagrams, spoken language and animated mathematical function displays) to help a student understand and construct mathematical proofs.

## 2 Empirical Study

We conducted a *Wizard-of-Oz (WOz)* experiment, supported by a tool [1], in order to collect a corpus of tutorial dialogs in the naive set theory domain. 24 subjects with varying educational background and prior mathematical knowledge ranging from little to fair participated in the experiment. The experiment consisted of several phases: (1) *preparation and pre-test* on paper, (2) *tutoring session* proper, and (3) *post-test and evaluation questionnaire*, on paper again. The subjects had to prove theorems, by applying the De Morgan laws ($K((A \cup B) \cap (C \cup D)) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))$), and by making use of the concepts power set ($A \cap B \in P((A \cup C) \cap (B \cup C))$) and complement (When $A \subseteq K(B)$, then $B \subseteq K(A)$, where $K$ signifies *K*omplement in German).

The interface enabled the subject to type text or insert mathematical symbols by clicking on buttons. The subject was instructed to enter steps of a proof rather than the complete proof as a whole, in order to enable a dialog with the system. The tutor-wizard's task was to respond to the student's utterances following a given algorithm. The wizard first classified the completeness, accuracy, and relevance of the subject's utterance with respect to a valid proof of the theorem at hand. Then, the wizard decided what dialog moves to make next and verbalized them. The wizard was free to mix text with formulas [2].

## 3 Phenomena observed in student utterances

Examples of student utterances appear in Figure 1 (the original German versions of utterances together with their English translation). We have divided the phenomena observed into three categories: (1) references, (2) ambiguities, and (3) imprecision, all associated with multiple readings on general linguistic grounds. As opposed to a typical dialog managemant component of a task-oriented system, which would invoke several clarification dialogs, the presence of accurate domain knowledge and task control needs to be exploited to avoid this whenever possible in a tutorial environment.

[1] The DIALOG project is part of the Collaborative Research Center on *Resource-Adaptive Cognitive Processes* (SFB 378) at University of the Saarland. SFB 378 web-site: http://http://www.ling.gu.se/projekt/siridus/.

| | |
|---|---|
| References | (1) Potenzmenge enthält alle Teilmengen, also auch $(A \cap B)$<br>*A power set contains all subsets, hence also* $(A \cap B)$<br>(2) $K((A \cup B) \cap (C \cup D)) = K(A \cup B) \cup K(C \cup D)$ de Morgan Regel 2 auf beide Komplemente angewendet<br>*de Morgan rule 2 applied to both complements* |
| Ambiguities | (3) de Morgan Regel 1 gilt auch für $K(C \cup D)$ de Morgan Regel 2 besagt $K(A \cap B) = K(A) \cup K(B)$. In diesem Fall z.B. $K(A) =$ dem Begriff $K(A \cup B)$ und $K(B) =$ dem Begriff $K(C \cup D)$. Deshalb ist dann $K((A \cup B) \cap (C \cup D)) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))$<br>*de Morgan rule 1 also holds for* $K(C \cup D)$ *de Morgan rule 2 means* $K(A \cap B) = K(A) \cup K(B)$. *In this case e.g.* $K(A) =$ *the term* $K(A \cup B)$ *and* $K(B) =$ *the term* $K(C \cup D)$. *Therefore* $K((A \cup B) \cap (C \cup D) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))$ |
| Imprecision | (4) $(A \cup B)$ muß in $P((A \cup C) \cap (B \cup C))$ sein, da $(A \cap B) \in (A \cap B) \cup C$<br>$(A \cup B)$ *must be in* $P((A \cup C) \cap (B \cup C))$, *since* $(A \cap B) \in (A \cap B) \cup C$<br>(5) $(B \cup A)) \subseteq C$ $(B \cup A)) \subseteq D$. Wenn $A$ Teilmenge von $C$ und $B$ Teilmenge von $C$ dann müssen beide Mengen zusammen ebenfalls eine Teilmenge von $C$ sein. Gleiches gilt mit $D$ $K(C \cap D) \cup K(A \cap B)$ Anwendung der de Morgan Regeln. $((B \cup A)) \subseteq C$ $(B \cup A)) \subseteq D$.<br>*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* $D$. $K(C \cap D) \cup K(A \cap B)$ *applying the de Morgan rules.* $((B \cup A)) \subseteq C$ $(B \cup A)) \subseteq D$. |

**Fig. 1.** Examples of dialog utterances. The predicates $P$ and $K$ stand for power set and complement, respectively.

- *References* may be established by the use of semantic operators, such as "on *the left side*", and "for the *inner parenthesis*". which are incomplete specifications. "Left side" refers to an equation, and "inner parenthesis", which is metonymic, refers to the expression enclosed by it. Generic and specific references may appear within the same utterance (cf. (1), where "Potenzmenge" (powerset) is used as a generic reference, whereas $A \cap B$ is a specific reference to a subset of a specific instance of the power set). Likewise, the example (2) illustrates the relevance of domain-specific foci, where determining the referent of "both complements" requires interpreting sides of an equation as individual focus spaces.
- *Ambiguities* concern the use of propositional junctors, which may allow for varying scopings, and the use of punctuations, such as commas, which has been observed to be ambiguously used to mean enumeration, implication, and conjunction. Moreover, the relation "=", which is mostly used within equations or for indicating a value assignment, has also been found as an indicator for a term substitution in an axiom ((3), $K(A \cup B)$ to be substituted for $K(A)$, and $K(C \cup D)$ for $K(B)$, respectively).
- *Imprecision* of natural language expressions may result in an ambiguous reference to domain relations or to a formally incorrect but metaphorically interpretable reference. Within the domain of mathematical sets, examples for an ambiguous reference are the expressions "must be in" in (4) which can be interpreted as "element" or "subset"; and "both sets *together*" in (5), which may be interpreted as union or intersection). An example for a metaphoric reference is "The intersection of two sets is less or equal to the smaller one of *these* sets", where "less than" needs to be interpreted as "subset of".

## References


1. A. Fiedler and M. Gabsdil. Supporting Progressive Refinement of Wizard-of-Oz Experiments. In Proceedings of the Sixth International Conference on Intelligent Tutoring Systems—Workshop W6: Empirical Methods for Tutorial Dialogue Systems, 2002.
2. A. Fiedler and D. Tsovaltzi. Automating Hinting in Mathematical Tutorial Dialogue. Proceedings of the EACL-03 Workshop on Dialogue Systems: interaction, adaptation and styles of management, Budapest, 2003.
3. J. Moore. What makes human explanations effective? In *In Proc. of the Fifteenth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ. Earlbaum, 2000.
4. M. Pinkal, J. Siekmann, and C. Benzmüller. Projektantrag Teilprojekt MI3 — DIALOG: Tutorieller Dialog mit einem mathematischen Assistenzsystem. In *Fortsetzungsantrag SFB 378 — Ressourcenadaptive kognitve Prozesse*, Saarbrücken, Germany, 2001. Universität des Saarlandes.