

Modeling anaphora in informal mathematical dialogue

Magdalena Wolska Ivana Kruijff-Korbayová

Fachrichtung Computerlinguistik
Universität des Saarlandes, Postfach 15 11 50
66041 Saarbrücken, Germany
{magda,korbay}@coli.uni-sb.de

Abstract

We analyze anaphoric phenomena in the context of building an input understanding component for a conversational system for tutoring mathematics. In this paper, we report the results of data analysis of two sets of corpora of dialogs on mathematical theorem proving. We exemplify anaphoric phenomena, identify factors relevant to anaphora resolution in our domain and extensions to the input interpretation component to support it.

1 Introduction

Our goal is to develop a discourse understanding module for a dialog-based system for tutoring mathematics. A number of computational anaphora resolution approaches have been proposed (Mitkov, 2002), including solutions specific to modeling reference to entities other than nominals (Byron, 2004), as well as approaches specific to dialogue (Eckert and Strube, 1999; Jain et al., 2004). We can partly draw on those solutions, however, our domain differs from the domains these approaches address in that it involves formalized mathematical notation. While parsing and interpretation techniques for mixed natural and symbolic language do exist (Baur, 1999; Zinn, 2003; Wolska and Kruijff-Korbayová, 2004), referentiality phenomena have not, to our knowledge, been thoroughly studied. An additional challenge is posed by formal errors and sloppiness in students' proofs that may introduce referential ambiguity.

(Wolska et al., 2004; Wolska and Korbayová, 2006) presented two corpora of tu-

torial dialogs on mathematical theorem proving collected in a Wizard-of-Oz setup. We conducted an analysis of this data in order to guide the development of an anaphora resolution algorithm suitable for dialogs in the domain of mathematics. Our goal is to (i) systematically investigate reference phenomena specific to mathematical dialog, (ii) based on empirical findings, propose a co-reference resolution method for our domain.

In this paper, we report the first results of data investigation. We concentrate on the peculiarities of the genre at hand: notably, references to mathematical concepts and expressions. With this focus in mind, we present and exemplify anaphoric phenomena observed in the two corpora. Second, we discuss our observations on implications for anaphora resolution and the functionality of the input interpretation component necessary to support it.

The paper is organized as follows: In Section 2, we briefly present the corpora we study. In Section 3, we show corpus examples of reference phenomena. In Section 4, we present our observations related to modeling anaphora in our domain, and extensions to an input interpretation module needed to support anaphora resolution.

2 Corpus

Our analysis is based on two tutorial dialog corpora¹ collected in Wizard-of-Oz experiments: Corpus-I (Benzmüller et al., 2003) and Corpus-II (Wolska and Korbayová, 2006). In both experiments, the subjects were told that they were interacting with a conversational tutoring system. They were using natural lan-

¹The corpora are available online.

Let R be a relation on a set M . Prove: $R = (R^{-1})^{-1}$

A relation is defined as a set of pairs. The above equation expresses an equality between sets. Set equality can be proven by The Principle of Extensionality, where one shows that every element of one set is also an element of the other set. Let (a,b) be a pair on $M \times M$. We have to show that $(a,b) \in R$ if and only if $(a,b) \in (R^{-1})^{-1}$. $(a,b) \in (R^{-1})^{-1}$ holds by definition of the inverse relation if and only if $(b,a) \in R^{-1}$ and this again holds by the definition of the inverse relation if and only if $(a,b) \in R$, which was to be proven.

Figure 1: Example proof from Corpus-II.

guage (German) typed on the keyboard as well as mathematical symbols. Both the subjects and the tutors were unconstrained in the way they formulated their turns. Corpus-I contains 66 dialogs (775 turns) on proofs in the domain of naive set theory, and Corpus-II 37 dialogs (1615 turns) on binary relations.

Analysis of the corpora reveals various phenomena that present challenges for modeling anaphora and anaphora resolution. The prominent phenomenon is reference to (parts of) the formal mathematical notation. This raises questions about introducing discourse entities for mathematical expression parts as well as requires extensions to the standard functionality of input processing sub-components. We discuss the extensions in Section 4.3, but first, illustrate the phenomena with examples from the corpora.

3 Phenomena

To indicate the overall complexity of the anaphora resolution task in our setting, we present an overview of common reference phenomena. First, we give a brief characterization of the language of informal mathematical discourse, and then present anaphoric phenomena specific to the domain: reference to (parts of) mathematical expressions and mathematical propositions.

3.1 Language of informal mathematical discourse

Informal mathematical discourse can be characterized as a mixture of natural language interleaved with conventionalized formal expressions. Formal mathematical language consists of a vocabulary of symbols and operators, and technical terminology specific to

a sub-field. Mathematical expressions include terms (denoting abstract mathematical objects) and statements (formulas) built from the vocabulary, both of arbitrary structural complexity. An informal proof consists of a sequence of assertions derived by application of inference rules. Figure 1 shows an example proof from Corpus-II presented to a subject at the end of a tutoring session.² In the course of the proof exposition, symbols that denote domain-objects (here: e.g. relations, pairs, sets) are mentioned and anaphoric devices are used to refer to abstract entities they denote or their specific (symbolic) instantiations in the discourse.

Below, we illustrate examples of references in informal mathematical dialogue from the point of view of the type of entity referred to. The phenomena themselves are not new, but the formal domain adds complexity to them, in particular from the point of view of referential ambiguity and functionality needed for anaphor resolution in general. The dialog excerpts to which we refer here are included in the Appendix.

3.2 Referring to (parts of) symbolic notation

Using pronouns and pronominal adverbs

In (1), a pronoun, “it”, is used to refer to a term in a formula, a set variable “B”, whose syntactic/semantic function in the formula can be viewed as that of a subject/agent, parallel to the semantic function of the anaphor. In (2), a pronoun is referring to a variable naming a member of a set. In (3) the same name, “x”, was introduced with the intention of denoting two different

²We present only an English translation here for space reasons.

entities. The entities are moreover of different types (in one case, a pair is a variable in a set abstract³ and “x” refers to an element of this pair, in the other case, “x” refers to a set-member variable of a simple form). Ambiguous designation is invalid in a mathematical proof and the tutor issues a clarification sub-dialog, in which, in turn, the pronominal reference in **S19** has an ambiguous denotation.

In (4), a pronominal adverb “davon” (en. *of it*) is used to refer to a complex term, “ $R \circ S$ ”, on the left-hand side of the definition. In principle, the reference is ambiguous: a competing antecedent for “davon” is the definiens part of the definition.

Using noun phrases In our analysis, we include bridging references. We have found certain types of bridging references to systematically recur in our corpora. For example, noun phrases such as “the inner bracket” and “the left side” refer to a formula’s structural part: a term in a formula. Both need a metonymic re-interpretation: “the left side” refers to the term to the left of the top-node operator in a formula (rather than a topological area), while “the inner bracket” refers to a bracketed sub-term of a bracketed term in a formula (rather than to a bracket itself).

There are two ways of interpreting the definite noun phrase “the powerset” in the student utterance “I have problems with the powerset, I don’t know how to compute it. . .” On the one hand, it may be referring to a term headed by the powerset operator (rather than the powerset operator itself) in the previous student turn that contains the following expression: “ $P((A \cup C) \cap (B \cup C)) = P(C \cup (A \cap B))$ ”. In this case, it needs a metonymic extension. Under this interpretation, the reference is ambiguous as there are two powersets in the expression. On the other hand, it is more plausible to interpret the reference generically; the student has a general problem in understanding the concept of a powerset.

Analogously, the quantified noun phrase,

³A set abstract is a set-denoting expression of the form $\{v : \phi\}$, where v is a variable and ϕ a formula

“beide Komplemente” in (5) needs a metonymic re-interpretation. Moreover, the reference is truly ambiguous in that there are five complement-headed terms in the preceding formula. A resolution algorithm must, therefore, not only decide on distributive vs. collective reading of the plural, but also identify plausible scopes for antecedent search.

In (7), the definite noun phrase “diese Menge” (en. *this set*) in **S35** is again a bridging reference to the set denoted by a term in **S34** (where the type of the result of the top-node operator is set).

Using demonstratives and discourse deixis

In (6), the deictic reference “der obere Ausdruck” (en. *the above expression*) refers to the entire formula in the preceding turn, while the demonstrative pronoun “dies” (en. *this*) in (7) refers to a term in the previous formula.

3.3 Referring to propositions

Pronouns, demonstratives and adverbial pronouns may be used to refer to propositions as well as partial proofs constructed in the course of a dialog. In (8) the adverbial pronoun “damit” (en. *with this*) in **S7**, refers to the proposition stated in the first clause of the utterance. The pronominal adverb “somit” (en. *with that*) in **S8** in the same excerpt may refer to the conjunction or implication of the assertions in **S7** or only the last assertion (marked with j in the example). In (9), the pronoun “es” (en. *it*) is referring to the proposition in the tutor’s turn **T19**.

3.4 Referring to domain-concepts

Both **definite and bare noun phrases** are used generically to refer to concepts in the domain, e.g. “the union” in: “The union of sets R and S contains all elements from R and all elements from S .”. In “Powerset contains all subsets therefore also $(A \cap B)$ ”, “powerset” is a generic reference, whereas “ $(A \cap B)$ ” is a specific reference to a subset of a specific instance of a power set introduced earlier. Moreover, named theorems and lemmata may be referred to by their **proper names**, for example, “deMorgan rule 2”.

	Corpus-I	Corpus-II
math. expr. part	26	13
proof step	35	81
formula	19	46
mixed	16	35
Total	61	94

Table 1: References to domain objects

To summarize, the first and most obvious observation based on the corpus is that anaphoric references are used to refer to the formal notation of mathematical expressions. References may address entire formal expressions or their parts, and antecedents may lie in either own or the other party’s turns. In spite of a seemingly high potential for ambiguity, only in one case was an explicit clarification dialog initiated by the tutor to clarify an ambiguous reference. Below, we present details of our corpus analysis and observations relevant for modeling anaphora.

4 Modeling anaphora in tutorial dialogues on proofs

We looked at all occurrences of references to *domain objects* in both corpora. For the purpose of this paper, by domain objects we mean (i) symbolic mathematical expressions and their parts, (ii) domain relevant propositions (mathematical assertions); e.g. proof steps proposed by the student expressed either formally or in words.⁴ Below, we present a quantitative result of our analysis, summarize the observations concerning referentiality phenomena with in our context, and present extensions to the input understanding module we have implemented to support anaphor resolution in our domain.

4.1 Quantitative corpus analysis

Overall, of the 1269 student turns in both corpora, 140 turns contained references to

⁴We do not include in this analysis proper name named entity references referring to domain concepts, theorems, lemmata, etc., such as “The Second De Morgan Law”, “The Distributivity Law”.

	PRP	prp. or loc. adv.	defn.art./dem.	def.NP
simple term	2	0	0	2
complex term	3	2	2	28
Total	5	2	2	30
ante. in S-turn	5	2	2	18
0 (same turn)	3	1	0	4
1	0	1	2	5
2 or earlier	2	0	0	9
ante. in T-turn	0	0	0	12
1 (prev. turn)	0	0	0	4
2 or earlier	0	0	0	8
of that in task def.	0	0	0	10

Table 2: References to parts of mathematical expressions

some domain object: 46 out of 332 (14%) in Corpus-I and 94 out of 709 (13%) in Corpus-II. The details of the analysis are presented in three tables which we discuss below.

Table 1 presents an overview of references to domain objects: parts of mathematical formulas and propositions (proof steps). There were overall 155 anaphoric references. The relatively large number of references to proof steps in the second corpus, we think, is related to the style in which proofs were conducted. Most students built their proofs by re-writing preceding terms, and referring to the previous step either with discourse markers, such as “hence” or “therefore” or with pronominal adverbs (e.g. “somit”, en. *with that*).

Table 2 shows an overview of references to (parts of) mathematical expressions. Here we include references to simple terms (i.e. symbolic identifiers such as variables A , B , x , etc.) and complex terms (terms containing at least one operator symbol). Of the 27 references, the largest proportion are nominal bridging references to formula parts (such as “left side” or “inner brackets” exemplified in Section 3.2). The antecedent tends to be found either in the student’s own turn or in the task definition (the goal formula to be proven).

Table 3 presents a summary of references to propositions. There are 116 instances of such references, the majority of which are realized with German pronominal or locative

	PRP	prp. adv. or loc. adv.	demonstr. or def.art.	def. NP
formula	0	28	27	10
mixed nl+formula	2	31	13	5
Total	2	59	40	15
ante. in S-turn	1	59	38	12
0 (same turn)	0	21	22	4
1	1	38	16	6
2 or earlier	0	0	0	2
ante. in T-turn	1	0	2	3
1 (previous turn)	1	0	2	0
2 or earlier	0	0	0	3
of that in task def.	0	0	0	3

Table 3: References to propositions

adverbs (59) and demonstrative pronouns or definite articles (40). A large proportion of these were found in Corpus-II. References to propositions tend to be local: most of the time, the antecedent is found in the student’s own turn, in the same turn as the anaphor or preceding turn with respect to the anaphor.

4.2 Factors in anaphor resolution

Our corpus analysis of anaphoric reference to domain objects, yields the following observations relevant to anaphora resolution:

Sources of information There appear to be three major information sources to which an anaphor resolution module in our domain needs access:

- (i) The semantic interpretation of the utterance and the utterance’s function;
In order to provide information on the semantic content, in particular, with respect to proof contribution, the utterances in the student turn must be parsed and interpreted in the context of the given domain. In the further discussion, we assume the approach to interpretation as the one presented in (Wolska and Kruijff-Korbayová, 2004) and discuss required extensions. Of particular importance for anaphor resolution is whether according to the assigned interpretation, the given utterance is intended to convey a proof step (domain contribution) or not.

- (ii) The correctness status of the last student’s proof step;

For example, in re-writing style of proofs, students tend to make references to the last correct proof step (or partial proof) to indicate that it justifies the current step. We will return to this when we discuss salience of propositions below.

- (iii) The semantic content of the last tutor move;

The tutor dialog moves include, among others, proof step evaluations (e.g. “That is not correct.”) and hints (e.g. “How about starting the proof like this: ...”). If the last tutor’s turn contains a hint that gives away the correct step expected at the time, the student is likely to refer to that step. Moreover, the first tutor’s dialog contribution defining the exercise (the goal proposition) is also often referred to.

Antecedent candidates in references to (parts of) formulas

Anaphoric references to mathematical expression parts appear to have local scope. In most cases, the referent occurred in the same or immediately preceding turn with respect to the anaphor, as exemplified in (1). In all cases of “it”-references (neuter personal pronouns) the anaphor was the entity on the left side of the candidate mathematical expression of type formula. This can be explained by the fact that in the verbalized form of such expressions, the entity on the left side plays the role of the subject or agent of the predicate.

Moreover, the structure of mathematical expressions is a strong indicator in identifying the search space for antecedents. This holds both in case of noun phrase references to topographical structure (e.g. “inner bracket” or “left side”) as well as in case of quantified phrases referring to sub-structure. In the latter case, the topographical structure may help in guiding the search (e.g. in (5)).

In order to support resolution of references to (parts of) mathematical expressions, an input interpretation module must include a

mathematical expression tagger and a deep parser for mathematical expressions, in particular, the parser must be capable of identifying all the relevant sub-structures of mathematical expressions. On the domain modeling side, it needs procedures for dealing with metonymic references to formula sub-structures.

Salience of propositions As the student develops the proof, the cognitive salience of the propositions that are part of the proof (proof steps) changes. At the beginning of the dialog, the most salient proposition is the goal formula (the exercise definition). According to our observations, as the proof progresses, the most salient proposition at a given time is the last correct proof step. If the student made several incorrect steps, no correct steps, and the tutor has not given away any steps, the goal formula in the exercise definition remains the most salient proposition even after several turns.

4.3 Extensions to input understanding module

To resolve references to (parts of) mathematical expressions, two issues must be taken into account: first, as mentioned above, we need a comprehensive analysis of mathematical expressions, and second, we need to include the entities specific to mathematical expression analysis in the domain-specific knowledge representation. Below, we summarize our implementation of domain modeling extensions required for reference resolution in the corpora we analyzed.

Mathematical expression parsing The mathematical expression parser uses simple indicators to identify mathematical expressions within sentence- and word-tokenized text. They include single character tokens, designated strings for mathematical symbols, and new-line characters.

The parser converts the infix notation used in the input into an expression tree whose nodes are marked as to whether they denote operators or variables; the expression type is marked on the root-node operator (e.g. FOR-

MULA, TERM, etc.). Moreover, the parser has access to domain-knowledge on the type of result of mathematical operations (e.g. the subset relation takes two sets and the type of the result is a truth-value). The expression tree is an input structure to subroutines relevant for reference resolution.

Considering the complexity of the mathematical expressions, we take a pragmatic approach in modeling reference to mathematical expression sub-parts, in that at the time of parsing we only create a discourse referent for the entire expression⁵, but not for every sub-structure entity relevant for anaphor resolution. Instead, the mathematical expression parser includes subroutines that on-demand recover (i) specific parts of mathematical expressions in specific PART-OF relations to the original expression, (ii) their types.

The choice of identified sub-structures is motivated by systematic reference in natural language to those parts (see Section 3.2) and includes: (i) topological features (such as “sides” of terms and formula); (ii) linear orders (e.g. “first”, “second” argument); (iii) structural groupings (bracketed sub-expressions) with information on their embedding. Execution of those subroutines is triggered by lexical semantic interpretation of the utterances (e.g. the meaning of “side” together with its modifier “left” in the representation of the noun phrase “the left side”).

Domain modeling Objects associated with types of mathematical expressions (e.g. FORMULA, TERM) as well as substructure delimiters (e.g. bracket, vertical bar of a set abstract) are represented in an ontological representation of domain objects.

Motivated by the systematicity in metonymic references to mathematical expression sub-parts, as part of the domain-model we encode “metonymy rules” that allow to re-interpret utterances with certain sortal restriction conflicts. Currently, the choice of rules is guided by phenomena found in our two cor-

⁵See (Wolska and Kruijff-Korbayová, 2004) for parsing the mixed language.

pora and includes the following:⁶

1. SIDE : TERM;
2. BRACKET : TERM;
3. OBJECT : RESULT;
4. RESULT : OPERATOR;
5. OPERATOR : SUB-TREE.

For example, the noun phrase “this set” referring to the expression $(S \cup R) \circ S^{-1}$ in (7), can be then interpreted by applying rule 3 first and then rules 4 and 5.

Discourse modeling Our preliminary implementation of the discourse model, includes a data structure storing a dialog history. Aside from the interpretation of student input utterances, a dialog history stores information on the semantic content of tutor moves, in particular, information about the correctness of the proof steps proposed by the student, as well as symbolic representation of proof steps that were disclosed to the student during tutoring.

5 Conclusions

Based on experimentally collected data, we presented examples of anaphoric phenomena in tutorial dialogs on mathematical proofs and a quantitative analysis of two corpora from the point of view of reference to entities specific to the genre: mathematical expressions and propositions expressing proof steps. We discussed corpus observations relevant to building an anaphor resolution algorithm for the domain. These observations yield constraints and preference criteria for forming sets of candidate antecedents and for antecedent search, that we can directly incorporate into a first implementation of an anaphor resolution algorithm which we are presently developing. We also presented our extensions to an input interpretation component necessary to support the anaphor resolution algorithm.

⁶The presentation included here is only schematic; for example, SIDE : TERM means that the concept “side” (left or right) may be alternatively interpreted as “term” in a formula to the left or right accordingly.

References

- Judith Baur. 1999. *Syntax und Semantik mathematischer Texte*. Diplomarbeit, Fachrichtung Computerlinguistik, Universität des Saarlandes, Saarbrücken, Germany.
- Christoph Benz Müller, Armin Fiedler, Malte Gabsdil, Helmut Horacek, Ivana Kruijff-Korbayová, Manfred Pinkal, Jörg Siekmann, Dimitra Tsovaltzi, Bao Quoc Vo, and Magdalena Wolska. 2003. A Wizard-of-Oz experiment for tutorial dialogues in mathematics. In *AIED2003 Supplementary Proceedings*, volume VIII: Advanced Technologies for Mathematics Education, pages 471–481, Sydney, Australia.
- Donna K. Byron. 2004. Resolving pronominal reference to abstract entities. Technical Report 815, University of Rochester Computer Science Dept.
- M. Eckert and M. Strube. 1999. Resolving discourse deictic anaphora in dialogues. In *Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics*, pages 37–44, Bergen, Norway.
- Prateek Jain, Manav Ratan Mital, Sumit Kumar, Amitabha Mukerjee, and Achla M. Raina. 2004. Anaphora resolution in multi-person dialogues. In *Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue at HLT-NAACL-04*, pages 47–50, Ann Arbor, MI.
- Ruslan Mitkov. 2002. *Anaphora Resolution*. Longman.
- Magdalena Wolska and Ivana Kruijff Korbayová. 2006. Factors influencing input styles in tutoring systems: the case of the study-material presentation format. In *Proceedings of the ECAI-06 Workshop on Language-enabled Educational Technology*, Riva del Garda, Italy. To Appear.
- Magdalena Wolska and Ivana Kruijff-Korbayová. 2004. Analysis of Mixed Natural and Symbolic Language Input in Mathematical Dialogs. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL-04)*, pages 25–32, Barcelona, Spain.
- Magdalena Wolska, Bao Quoc Vo, Dimitra Tsovaltzi, Ivana Kruijff-Korbayová, Elena Karagjosova, Helmut Horacek, Malte Gabsdil, Armin Fiedler, and Chris Benz Müller. 2004. An annotated corpus of tutorial dialogues on mathematical theorem proving. In *Proceedings of International Conference on Language Resources and Evaluation (LREC-04)*, pages 1007–1010, Lisbon, Portugal. ELDA.
- Claus Zinn. 2003. A Computational Framework For Understanding Mathematical Discourse. *Logic Journal of the IGPL*, 11(4):457–484.

Appendix – Dialog excerpts

- (1) **S6:** ... Da, wenn $A_j \subseteq K(B_i)$ sein soll, A_j Element von $K(B_i)$ sein muss. Und wenn $B_i \subseteq K(A_j)$ sein soll, muss es_i auch Element von $K(A_j)$ sein.
Because if it should be that $A_j \subseteq K(B_i)$, A_j must be an element of $K(B_i)$. And if it should be that $B_i \subseteq K(A_j)$, it must be an element of $K(A_j)$ as well.
- (2) **S1:** Wie ist $R \circ S$ definiert?
How is $R \circ S$ defined?
T1: $R \circ S := \{(x,y) \mid \exists z(z \in M \wedge (x,z) \in R \wedge (z,y) \in S)\}$
 ...
S4: ist z_i nur fuer die Definition eingefuehrt oder hat es_i einen anderen Sinn?
is z introduced only for the definition or does it have a different meaning?
- (3) **S18:** ... Daraus folgt $(R \cup S) \circ T = \{(x,y) \mid \exists z(z \in M \wedge (x,z) \in \{x? \mid x? \in R \vee x? \in S\} \wedge (z,y) \in T)\}$
... From that follows $(R \cup S) \circ T = \{(x,y) \mid \exists z(z \in M \wedge (x,z) \in \{x? \mid x? \in R \vee x? \in S\} \wedge (z,y) \in T)\}$
T19: Was bedeutet die Variable x_i bei Ihnen?
What is the meaning of your variable x ?
S19: x_i hat zwei Bedeutungen es kommt in zwei verschiedenen Mengen vor
 x has two meanings it is contained/comes in two different sets
T20: Benutzen Sie bitte fuer die zwei verschiedenen Bedeutungen von x zwei verschiedene Bezeichnungen.
Please use two different designations for the two different meanings of x .
- (4) **S:** $R \circ S_i := \{(x,y) \mid \exists z(z \in M \wedge (x,z) \in R \wedge (z,y) \in S)\}$
T: Das ist richtig!
This is correct!
S: Nun will ich das Inverse davon;
Now I want the inverse of it
- (5) **T1:** Bitte zeigen Sie: $K((A \cup B) \cap (C \cup D)) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))!$
Please show: ...
S2: de morgan regel 2 auf beide komplemente angewendet
de morgan rule 2 applied to both complements
- (6) **T1:** Bitte zeigen Sie: $[K((A \cup B) \cap (C \cup D)) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))]_i!$
S1: laut De-Morgan-Regel-2 gilt: $(K(A) \cap K(B)) = K(A) \cup K(B)$, damit kann ich den oberen Ausdruck _{i} wie folgt schreiben: ...
By De-Morgan-Law-2 holds: $(K(A) \cap K(B)) = K(A) \cup K(B)$, given that I can write the above expressions as follows: ...
- (7) **S33:** Nach Aufgabe W ist $(S \circ (SUR)^{-1})^{-1} = [(SUR)^{-1}]^{-1} \circ S^{-1}]_i$
By Exercise W holds: ...
T34: Das ist richtig!
S34: Dies _{i} ist nach Theorem 1 gleich $[(SUR) \circ S^{-1}]_j$
This is by Theorem 1 equal to ...
T35: Das ist auch richtig!
S35: Ein Element (a,b) ist genau dann in dieser Menge _{j} , wenn es ein $z \in M$ gibt mit $(a,z) \in SUR$ und $(z,b) \in S^{-1}$
An element (a,b) is in this set if and only if there is a $x \in M$ such that $(a,x) \in SUR$ and $(x,b) \in S^{-1}$
- (8) **S7:** Also [ist $(z,x) \in S$ und $(y,z) \in R$] _{i} und damit _{i} auch $[(y,x) \in R \circ S]_j$
Therefore holds $[(z,x) \in S \text{ and } (y,z) \in R]_i$ and by that i also $[(y,x) \in R \circ S]_j$
T7: Sie haben recht.
You are correct.
S8: Somit _{j} ist $(x,y) \in (R \circ S)^{-1}$
Given that it holds that $(x,y) \in (R \circ S)^{-1}$
- (9) **S12:** $\exists z \in M$, so dass $(x,z) \in S^{-1}$ und $(z,y) \in R^{-1}$
 $\exists z \in M$ such that $(x,z) \in S^{-1}$ and $(z,y) \in R^{-1}$
T18: Richtig. Wissen Sie, ob ein solches z existiert?
Correct. Do you know whether such z exists?
S13: Nein
No
T19: Erinnern Sie sich daran, [dass es ein z gibt mit $(x,z) \in S^{-1}$ und $(z,y) \in R^{-1}$.] _{i}
Do you remember that there is a z such that $(x,z) \in S^{-1}$ and $(z,y) \in R^{-1}$.
S14: Ja, ich habe es_i vorausgesetzt
Yes, I made such assumption