Towards Modelling and Using Common Ground in Tutorial Dialogue[*]

Mark Buckley and **Magdalena Wolska**
Department of Computational Linguistics
Saarland University
66041 Saarbrücken, Germany
{buckley|magda}@coli.uni-sb.de

Abstract

In order to avoid miscommunication participants in dialogue continuously attempt to align their mutual knowledge (the "common ground"). A setting that is perhaps most prone to misalignment is tutoring. We propose a model of common ground in tutoring dialogues which explicitly models the truth and falsity of domain level contributions and show how it can be used to detect and repair students' false conjectures and facilitate student modelling.

1 Motivation

In order to communicate efficiently, participants in a dialogue take into account the information believed to be mutually known to them: the "common ground" (Clark and Schaefer, 1989). This concerns not only knowledge accumulated in the course of dialogue, but also common ground (context) that is presupposed prior to the interaction. In order for a piece of information to become common ground (henceforth CG), it must be explicitly or implicitly acknowledged by the interlocutor in the process called *grounding* (Clark and Schaefer, 1989; Traum, 1994). Lack of grounding may lead to incorrect beliefs about CG or, using Stalnaker's term, "defective context" (Stalnaker, 2002) i.e. a situation in which discourse participants presuppose different things. This, in turn, may lead to miscommunication.

A setting that is perhaps most prone to misalignment in discourse participants' beliefs is tutoring. Student-tutor teaching interactions are characterised by an inherent asymmetry of knowledge possessed by the tutor and the learner (Munger,

1996; Lee and Sherin, 2004). In order to effectively guide the student in a problem-solving task, the tutor must make inferences as to the student's state of knowledge based on his/her interpretation of student's utterances. Empirical research shows, however, that untrained tutors tend to perform specific targeted moves (e.g. use curriculum scripts, example problems, give immediate evaluative feedback) that locally address the student's progress (or lack thereof) on the task at hand, instead of focusing mainly on coordinating the CG, i.e. establishing complete understanding of the students' state of beliefs, and so cognitive alignment (Graesser et al., 1995). Considering this, it is difficult for tutors to identify, let alone repair, deep misconceptions that underlie students' errors. However, when a misalignment in student's beliefs as to CG becomes apparent based on the linguistic content of the student's utterance, the tutor may choose to explicitly address it by challenging the student's statement. Moreover, an explicit model of CG can feed into a module that monitors the student's performance (McArthur et al., 1990; Merrill et al., 1995).

In this paper, we propose a preliminary model of CG in tutoring dialogues on problem solving, in the context of building a conversational intelligent tutoring system for mathematical proofs. As the dialogue progresses the CG in our model develops as a store of the truth and falsity of the contributions that the student has made, based on the evaluations that the tutor has given. In addition, discourse referents for formulas are introduced to support modeling their salience. The CG forms the context for judging whether an utterance constitutes evidence that misalignment has occurred.

We begin by discussing examples of student-tutor interactions that exemplify the use of CG in this domain and motivate the need for modelling CG in an automated system (Section 2). We present the structure of the CG maintained by the

dialogue model, the mechanism of updating the CG, and discuss the uses of CG in the tutorial dialogue: detecting and repairing student's false conjectures and facilitating student modelling (Section 3). We give a walk-though of our examples in Section 4 and related work is discussed in Section 5.

2 Linguistic Data

To collect data about naturalistic human-computer tutorial dialogue interactions, we conducted two experiments in a Wizard-of-Oz paradigm (Kelley, 1984). The subjects and the wizards were using natural language (German; typed on the keyboard) and mathematical symbols available on the GUI to interact with the simulated system, and they were unconstrained in their language production.

In the first experiment (Wolska et al., 2004) the subjects were tutored using one of the following strategies: *minimal feedback* (students were given feedback only on correctness and completeness of proposed proof-steps), *didactic* (when the student made no progress, the tutor disclosed the expected reasoning step), or *socratic* (a pre-defined hinting algorithm was used to guide the student toward the solution). In the second experiment (Benzmüller et al., 2006) the tutors did not follow any tutoring algorithm, but were given general instructions on *socratic* tutoring. The first experiment concerned naive set theory and the second binary relations. In both, students were given study material before the tutoring session containing the basic mathematical knowledge required to solve the exercises. Overall, the collected corpora consist of 22 and 37 dialogue logfiles for the first and second experiment respectively.

2.1 Dialogue Phenomena

The CG that student and tutor believe currently exists in the dialogue can become misaligned as the dialogue progresses, for instance due to misunderstanding, misconception, or the boundedness of attentional resources. Evidence of misalignment of CG can be observed, for example, in certain situations in which informationally redundant utterances (IRUs) (Karagjosova, 2003; Walker, 1993) are performed. An utterance is informationally redundant if the proposition it expresses is entailed, presupposed or implicated by a previous utterance in the discourse.

If an unmarked IRU is performed then the in-

formation it contains, which has already been grounded, is being repeated without the speaker indicating that this repetition is being done on purpose. This indicates to the hearer that this information was not in what the speaker believes to be the CG of the dialogue. The hearer must then conclude that the CG has become misaligned.

Sometimes it is necessary to repeat information that has already been grounded, for example to make a known fact salient again to support an argument (Walker and Rambow, 1994). Such utterances are informationally redundant. To prevent the hearer concluding from such utterances that misalignment has occurred, the speaker explicitly indicates that he is aware that he is repeating shared information (i.e. that the fact should be CG) by marking with phrases such as "of course". In tutorial dialogue hints which remind students of previously proved facts and concepts from the tutoring domain are IRUs. Students also use IRUs to check if what they believe the CG to be is actually that which the tutor believes it to be. In the following examples from the corpus we give English translations and include the original German utterances where they are illustrative.

In (1) the domain content of utterance S10, that the assertion that the formula embedded in the utterance holds, is repeated in utterance S18.

(1) **S10:** It holds that $(R \cup S) \circ T = \{(x,y)|\exists z(z \in M \land (x,z) \in (R \cup S) \land (z,y) \in T\}$

   T10: That's right!

     . . .

   S18: By definition it holds that $(R \cup S) \circ T = \{(x,y)|\exists z(z \in M \land (x,z) \in (R \cup S) \land (z,y) \in T\}$

   T18: That's right! You've already performed this step.
(*German: Diesen Schritt haben Sie vorhin schon vollzogen.*)

The confirmation (T10) of this fact puts the truth of the formula in S10 in CG, and therefore when utterance S18 is performed it is an IRU. Because S18 is unmarked for informational redundancy, the tutor concludes that misalignment of context has occurred, i.e. the fact concluded in S10 is no longer CG for the student. He augments his confirmation (T18) with the indication ("already") that the step in S18 had already been performed, telling the student that misalignment occurred and realigning their CG.

In example (2) the student explicitly introduces by assumption a fact that he has already proved.

(2) **S3:** Let $(a, b) \in (R \circ S)^{-1}$. Then it holds that $(b, a) \in (R \circ S)$

　　T3: That's right.

　　　　…

　　S6: Let $(b, a) \in (R \circ S)$. Then …

　　T6: Since you already know that $(b, a) \in (R \circ S)$, you don't need to postulate it again.

In S3 the student has successfully proved $(b, a) \in (R \circ S)$, and the tutor's confirmation of this (T3) makes it CG. The student later wants to use this fact as the antecedent of a derivation (S6), but wrongly introduces it as a new assumption. Thios shows that the truth of this formula is no longer CG for the student, i.e. misalignment has taken place. In order to realign the CG the tutor reminds the student that he has previously proved the formula, and this utterance is marked with "already" (T6).

An example (3) an IRU performed by the tutor is marked so that the student does not mistakenly conclude that misalignment has taken place.

(3) **S2:** $A \cap B = \emptyset$

　　　　…

　　T4: Right. Now what?

　　　　…

　　T8: … The justification could for instance be: Let x be an arbitrary element of B, then it can't be in A (since of course $A \cap B = \emptyset$) … (*German: … (da ja $A \cap B = \emptyset$) …*)

The student has proved a formula in S2 which was confirmed in T4, making it CG. In T8 the tutor recaps the solution proof. The formula $A \cap B = \emptyset$ is part of the proof, and is thus in the CG, so the tutor marks the reminder of this fact with the particle "of course".

An example of a student's marked IRU is shown in utterance S4 of (4), in which the IRU is used by the student to check suspected misalignment. "Doch" is a modal particle which, when deaccented, marks old or shared information.

(4) **S3:** and for the powerset it holds that: $P(C \cup (A \cap B)) = P(C) \cup P(A \cap B)$

　　T4: Do you really mean: $P(C \cup (A \cap B)) = P(C) \cup P(A \cap B)$?

　　S4: But I think: $P(A) \cup P(B) = P(A \cup B)$ (*German: ich denke doch: $P(A) \cup \ldots$*)

　　T5: That's not right! Maybe you should have another look in your study material.

　　S5: sorry, it holds of course that: $P(C \cup (A \cap B)) \subseteq P(C) \cup P(A \cap B)$

　　T6: Really?

　　S6: oh, no.. . the other way around

　　T7: That's right at last!

In S3 the student claims a fact which is then questioned by the tutor. This causes the student to suspect a misalignment, because a rule he used in deriving the fact and which he believed to be true is in fact false. In S4 the student then checks whether this rule is in fact in the CG by stating it explicitly. He considers S4 to be uninformative, and therefore marks it explicitly with "doch" (meaning "but I thought..."). However S4 actually is informative, in the sense that it is not subsumed by the set of facts in the CG when it is uttered. This leads the tutor to conclude that misalignment has taken place. In addition to rejecting the rule, he also directs the student to the study material. The next student proof step (S5) is again rejected (T6). In S6 the student gets the rule right, which is confirmed in T7. The student adds the corrected rule to his CG, completing the realignment that began in S4.

The data shows that misalignment occurs between student and tutor, and that it can be observed in the case of informationally redundant utterances. Unmarked IRUs (such as S18 in example (1) and S6 in example (2)) are evidence that CG has become misaligned, and should trigger strategies for realignment. Conversely, when IRUs are to be generated as part of pedagogical strategies (T8 in example (3)), these should be marked as such in order to avoid the student falsely concluding that misalignment has occurred. Finally, misalignment can be evidenced by utterances which are marked for informational redundancy but are in fact not IRUs (S4 in example (4)). To account for such phenomena a model of CG is necessary that allows the detection of which utterances are informationally redundant and which are not, at the level of truth in the domain. The CG must therefore model the utterances that were performed and whether their content was accepted by the tutor, and thus grounded.

3 Modelling Common Ground

Our model is developed within the wider scope of a tutorial environment for mathematics. The student's task is to build a proof of a mathematical theorem. The student does this by conducting a dialogue with the tutor in which he/she performs utterances which may contain proof steps. The environment includes study material for the theory at hand. Turn-taking during tutoring sessions is strictly controlled. Each correct proof step extends

the current partial proof. The task is completed when the student has constructed a complete proof of the theorem.

3.1 Elements of the Architecture

We now briefly describe the roles played by those system modules in the environment which are directly relevant to analysing CG.[1]

Discourse Interpreter A discourse interpretation module analyses students' natural language utterances.[2] The result of the analysis is the linguistic meaning of the utterance, the dialogue move that represents its function and, in case of domain contributions, a formal representation p for the proof manager (realised as, for example, in (Wolska and Kruijff-Korbayová, 2004)). In particular, the linguistic meaning of modal particles such as "doch" is reflected in the representation in that a feature MARKED is present.

Proof Manager A proof manager maintains the solution the student is building and evaluates proof steps in this context (Dietrich and Buckley, 2007). It can check the correctness and relevance of proof steps by accessing a domain reasoner such as ΩMEGA (Siekmann et al., 2006) or Scunak (Brown, 2006).

Tutorial Manager A tutorial manager stores pedagogical expertise on when and how hints should be given and maintains a student model (Tsovaltzi et al., 2004). The tutorial manager computes what dialogue moves the system should perform. Two possible dialogue moves which are relevant for this model are `accept` and `reject`. It performs the content selection step for output generation, which includes deciding whether utterances which are informationally redundant should be marked as such. It also decides whether to realise moves in the declarative or interrogative mood, as in utterance T4 in example (4).

3.2 Our Model

We model CG as being similar to that of DeVault and Stone (2006). Their *objective normative context* is a product of the actions taken by the dialogue participants. In our case, actions in the dialogue result in the dialogue participants having beliefs about the truth (or falsity) of the propositions that are contributed by the student and evaluated by the tutor. This is combined with the knowledge in the study material that the students are given before the tutorial session. We assume that it is part of the CG at the start of the dialogue. In our model the CG contains the facts that propositions were uttered, the evaluations of those utterances by the tutor, and the facts that the student knows about the domain as a result of preparatory study.

3.2.1 Types of Entities in the Model

There are two types of entities in the model: discourse referents (for entities introduced in the discourse) and propositions.

Domain contributions contain or refer to formulas that the student uses or concludes. For each domain contribution the discourse interpreter delivers the discourse referent for the proposition that the utterance expresses. Our model includes these discourse referents in the common ground. When references are made to substructures of formulas, for instance "the left-hand side of ..." we add new referents as needed.[3]

The fact that a proposition was uttered is modeled as $\texttt{uttered}(speaker, p)$, where speaker is the dialogue participant who performed the utterance. Having $\texttt{uttered}(speaker, p)$ in the CG tells us only that the event took place, and does not tell us anything about the truth of the content P of the utterance. Evaluations of the propositions that were uttered have the form either $\texttt{holds}(p)$ or $\neg\texttt{holds}(p)$, depending on whether they were accepted or rejected by the tutor. For previous knowledge that the student is assumed to have we use $\texttt{prev}(p)$.

Finally we model the utterances which are performed in the dialogue as objects u, and access the proposition p expressed by an utterance u with $p = \texttt{expresses}(u)$. In this way we can access the proof step that an utterance contained. The propositions expressed by utterances are treated as verbatim formulas.

The entities described above are represented in the dialogue model as shown in Figure 1, where

---

[1]We omit a discussion of other modules which are part of the architecture.

[2]For the purposes of this exposition we only consider assertion-type dialogue moves which contain domain contributions (here labelled with `domcon`), that is, possibly underspecified proof steps. For example we do not treat questions or meta-level communication management etc.

[3]This account of discourse referents is intentionally simple — a full account would require for instance referents for actual utterances in order to resolve references like "what I said above".
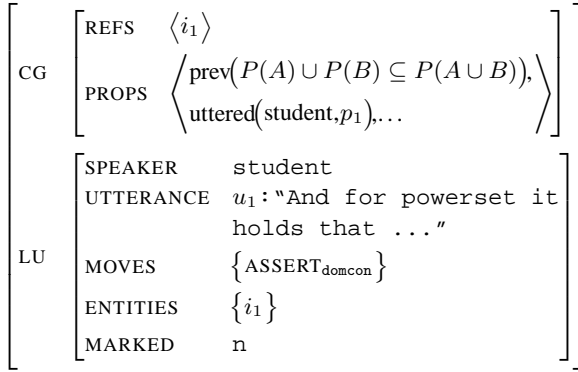
$$\begin{bmatrix} \text{CG} & \begin{bmatrix} \text{REFS} & \langle i_1 \rangle \\ \text{PROPS} & \left\langle \begin{array}{l} \text{prev}\big(P(A) \cup P(B) \subseteq P(A \cup B)\big), \\ \text{uttered}\big(\text{student}, p_1\big), \ldots \end{array} \right\rangle \end{bmatrix} \\ \text{LU} & \begin{bmatrix} \text{SPEAKER} & \text{student} \\ \text{UTTERANCE} & u_1 : \text{"And for powerset it} \\ & \text{holds that } \ldots \text{"} \\ \text{MOVES} & \big\{\text{ASSERT}_{\text{domcon}}\big\} \\ \text{ENTITIES} & \big\{i_1\big\} \\ \text{MARKED} & \text{n} \end{bmatrix} \end{bmatrix}$$

Figure 1: Dialogue model representation

CG is the common ground and LU is the last utterance in the dialogue. CG/REFS contains the discourse referents which have been introduced so far. CG/PROPS contains the three types of propositions that we model, namely `uttered`, $(\neg)$`holds` and `previous`. Both of these slots have the type ordered set which provides functions for membership, push and pop. We use this representation as simple account of salience. In CG/REFS, the most salient discourse entity is the one whose referent is at the top of the stack.

The LU part of the dialogue model stores a representation of the last utterance. It includes which dialogue participant performed the utterance, the actual utterance itself, the set of dialogue moves that the utterance realised as well as the discourse referents that were addressed. The flag `marked` indicates whether the utterance was marked for informational redundancy.

3.2.2 Updating the Common Ground

Information is added to the CG as a result of utterances performed by the student and the tutor. This corresponds to implicit grounding at the understanding level (Clark, 1996).[4]

We model three dialogue moves which lead to an update of the CG: `domcon`, `accept` and `reject`. Domain contributions claim the truth of formulas derived by proof steps, and in our model they correspond to Clark's proposal or presentation phase. In the case of a `domcon` we make the updates push(`uttered`(s,p),CG/PROPS) and push(i,CG/REFS), where p is the content of the

---

Predicate	Definition
exists(i)	$i \in$ CG/REFS
exists(p)	holds(p) $\in$ CG/PROPS
	$\vee \neg$holds(p) $\in$ CG/PROPS
salient(i)	$i = $ top(CG/REFS)
iru(u)	exists(expresses(u))

Table 1: Predicates on common ground.

proposition and i is the discourse referent introduced by p.

The tutor's evaluations of domain contributions performed by the student are represented by `accept` and `reject` moves, which lead to the updates push(holds(p),CG/PROPS) and push($\neg$holds(p),CG/PROPS) respectively. Here we rely on the discourse interpreter being able to determine which domain contribution the tutor is responding to and, in turn, what its propositional content p was. Because they have the effect of putting things in CG, `accept` and `reject` moves correspond to Clark's acceptance phase. We make the assumption that the tutoring scenario gives the tutor the authority to ground content simply by evaluating it. In effect, the student is expected to accept what the tutor says as being true. A further update is made to the CG when a substructure of a formula is accessed. New discourse referents j for subformulas are generated on demand and added to the CG by push(j,CG/REFS).

3.2.3 Testing and Using the Common Ground

Now that we can update the CG to reflect the current state of mutual belief, we define a set of predicates (see Table 1) that test properties of given propositions and referents. The predicate `exists`(x) holds when the discourse referent or proposition x has already been introduced, and `salient`(i) holds of the most salient discourse referent i. Utterances are IRUs if they satisfy the predicate `iru`(u), that is, if the proposition they express is already in the CG. Since `expresses` treats formulas verbatim, `iru`(u) can only hold when the formula is a case-insensitive match of u. We also define an operation `makesalient`(i) which promotes an existing discourse referent i to the top of CG/REFS, making it the most salient referent. The `makesalient` operation is performed when `iru`(u) is detected because a formula is being mentioned for a second time and should become salient again.

The `exists` predicate allows us to determine whether utterances are informationally redundant

in the context of the dialogue, and using `exists` we can now define a test on the dialogue model which tells us whether the last utterance is evidence that the common ground has become misaligned. Informally, we can conclude that misalignment has occurred if an IRU is not linguistically marked $(-)$ for informational redundancy or if a non-IRU is marked $(+)$ for informational redundancy. In terms of the predicates introduced above, we express this condition with the predicate `misaligned`:

misaligned *iff* (LU/MARKED$-$ $\wedge$ iru(LU/UTTERANCE)) $\vee$
LU/MARKED$+$ $\wedge$ $\neg$iru(LU/UTTERANCE)

In determining when to mark tutor utterances as informationally redundant, the tutoring manager uses the CG as input when it is asked to generate tutorial content. This way it can check if `iru`(u) holds of a planned utterance u and if so add marking.

4 Examples

We now illustrate how our model accounts for the examples above. For the purpose of example (1) we let p_1 stand for the proposition embedded in utterance S10, so that $p_1 = $ `expresses(S10)`. The domain contribution realised in S10 triggers the updates `push(uttered`(s, p_1)`,CG/PROPS)` and `push(`i_{p_1}`,CG/REFS)`. The `accept` performed by the totor then triggers the update `push(holds`(p_1)`,CG/PROPS)`, and the resulting CG is

$$
\begin{bmatrix} \text{CG} & \begin{bmatrix} \text{REFS} & \langle i_{p_1}, \dots \rangle \\ \text{PROPS} & \langle \texttt{uttered}(s, p_1), \texttt{holds}(p_1), \dots \rangle \end{bmatrix} \end{bmatrix}
$$

When S18 is performed the predicate `misaligned` becomes true. This is because S18 is an IRU (the proposition it expresses is a match of p_1) but is unmarked for informational redundancy. The system concludes that misalignment has taken place and the tutoring module generates the reminder that the student should already believe that `holds`(p_1), helping him to realign.

Example (2) shows a direct reference to a fact that the student should know. As in example (1), the contribution S3 followed by the acceptance T3 results in CG/PROPS containing `holds`(p_2), where $p_2 = (b, a) \in (R \circ S)$. In S6 the student assumes this fact again and the system can determine that `misaligned` holds because the formula in S6

matches p_2. It performs `makesalient(`i_{p_2}`)`, and the tutoring module reminds the student that he already knows the fact p_2 (T6).

As a result of the utterances S2 and T4 in example (3), the CG includes `uttered`(s, p_3) and `holds`(p_3), where p_3 is the formula $A \cap B = \emptyset$. When the system recapitulates the solution in T8, one of the utterances expresses a proposition which matches p_3. That means that `exists`(p_3) holds and that this utterance is an IRU. So that the student does not mistakenly conclude that misalignment took place, the system generates the utterances augmented with the marking "of course" to indicate informational redundancy.

We treat example (4) in more detail because it shows how misalignment can be detected and repaired. In Figure 1 we saw the state of the dialogue model after utterance S3. T4 is a `reject`, so the model is updated to

$$
\begin{bmatrix} \text{CG} & \begin{bmatrix} \text{REFS} & \langle i_{p_4}, \dots \rangle \\ \text{PROPS} & \langle \neg\texttt{holds}(p_4), \texttt{prev}(P(A) \cup \dots) \rangle \end{bmatrix} \end{bmatrix}
$$

where $p_4 = $ `expresses(S3)`. The student has a misconception that $P(A) \cup P(B) = P(A \cup B)$ holds. Since this rule is not correct, there is no proposition `prev`$(P(A) \cup P(B) = P(A \cup B))$ in CG/PROPS. That means that when the utterance S4, in which the student checks whether the misconceived rule is correct or not, is performed, we have $\neg$`iru(S4)`. However, the marking of S4 with the particle "doch" signals that the student assumes it to be shared knowledge. From these two facts the system detects that misalignment occurred. This type of misalignment informs the tutoring module to execute a strategy to resolve a misconception, namely, the student is referred to the study material. The resulting state is

$$
\begin{bmatrix} \text{CG} & \begin{bmatrix} \text{REFS} & \langle i_{p_5}, i_{p_4}, \dots \rangle \\ \text{PROPS} & \langle \neg\texttt{holds}(p_5), \neg\texttt{holds}(p_4), \dots \rangle \end{bmatrix} \end{bmatrix}
$$

In S5, with $p_6 = $ `expresses(S5)`, the student tries to correct the proof step that was rejected in utterance S3 by using a different rule, but the rule he applies (that $P(A) \cup P(B) \supseteq P(A \cup B)$) is not the correction of his original misconception, and the step is rejected (T5). The update is analogous, and we now have

$$
\begin{bmatrix} \text{CG} & \begin{bmatrix} \text{REFS} & \langle i_{p_6}, i_{p_5}, i_{p_4}, \dots \rangle \\ \text{PROPS} & \langle \neg\texttt{holds}(p_6), \neg\texttt{holds}(p_5), \dots \rangle \end{bmatrix} \end{bmatrix}
$$

The domain contribution S6 is an accepted (T7), correct application of the previously misconceived rule. By applying the correct rule the student shows that he has resolved the misconception that became apparent in utterance S4. The CG has now been realigned by adding $\texttt{holds}(P(A) \cup P(B) \subseteq P(A \cup B))$, and this information can be passed to the tutoring module.[5]

5 Related Work

Jordan and Walker (1996) compare models of how agents in a collaborative setting decide to remind each other about salient knowledge, and argue for an approximate rather than detailed model of the attentional state. In tutoring, the decision to remind is further influenced by pedagogical strategies. Our model provides input to this decision making process, however, the decision itself is made by the tutoring module. For example, the contents of CG could be used to realise a *Point-to-information* hint which is part of the hint taxonomy proposed by (Tsovaltzi et al., 2004).

Baker et al. (1999) argue that learning from grounding is the basis of collaborative learning and Pata et al. (2005) show how student grounding acts serve to inform tutoring scaffolds. Intelligent tutoring systems, such as AutoTutor (Person et al., 2000) and Ms Lindquist (Heffernan and Koedinger, 2002), with simple dialogue models have no model of CG, but capture misconceptions using explicit buggy rules. In those systems, there is no clear separation between modeling the dialogue itself and modeling the tutoring task. The dialogue advances according to the local tutoring agenda. (Zinn, 2004) presents a dialogue-based tutoring system in which discourse obligations are generated from a store of task solution descriptions and the CG is maintained in the dialogue model. However, the choice of tutoring actions is not informed by the state of the CG, but rather is explicitly encoded.

6 Conclusion and Further Work

We presented a preliminary model of common ground for a domain where grounding propositional content is crucial. However in principle this model is general enough to be be applied to other domains. The model takes into account Clark's

---

[5]The tutoring manager should record and make use of the fact that the student had and repaired a defective context. However we do not treat this topic here.

distinction between proposal and acceptance of dialogue contributions.

Indeed the current model is somewhat simplistic. There are a number of aspects of grounding which we observe in our corpus which this model does not account for but could be extended for, for instance domain content which is in a "pending" state when argumentation is taking place. Our further work will include extending the model to a larger set of dialogue moves including grounding acts. To obtain a more fine-grained model of context we need to further investigate what additional information about the problem-solving steps the domain reasoner can provide to the dialogue model, and thus to the tutoring manager. Furthermore we need a model of salience of propositions and steps in the problem-solving task, which may require a more flexible data structure. In a broader context it may be necessary to consider deletion of propositions however the conditions under which deletion rather than decay should occur need to be investigated. Current work includes implementation in TrindiKit.

References

Michael Baker, Tia Hansen, Richard Joiner, and David Traum. 1999. The role of grounding in collaborative learning tasks. In Pierre Dillenbourg, editor, *Collaborative Learning. Cognitive and computational approaches*, Advances in Learning and Instruction Series, pages 31–63. Pergamon, Amsterdam, The Netherlands.

Christoph Benzmüller, Helmut Horacek, Henri Lesourd, Ivana Kruijff-Korbayová, Marvin Schiller, and Magdalena Wolska. 2006. A corpus of tutorial dialogs on theorem proving; the influence of the presentation of the study-material. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2006)*, pages 1766–1769, Genoa, Italy.

Chad Edward Brown. 2006. Dependently Typed Set Theory. SEKI-Working-Paper SWP–2006–03, Saarland University.

Herbert H. Clark and Edward F. Schaefer. 1989. Contributing to discourse. *Cognitive Science*, 13(2):259–294.

Herbert H. Clark. 1996. *Using Language*. Cambridge University Press.

David DeVault and Matthew Stone. 2006. Scorekeeping in an uncertain language game. In *Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue (brandial)*, pages 139–146, Potsdam, Germany.

Dominik Dietrich and Mark Buckley. 2007. Verification of Proof Steps for Tutoring Mathematical Proofs. In *Proceedings of the 13th International Conference on Artificial Intelligence in Education*, Los Angeles. To appear.

A. C. Graesser, N. K. Person, and J. P. Magliano. 1995. Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology*, 9:495–522.

Neil T. Heffernan and Kenneth R. Koedinger. 2002. An Intelligent Tutoring System Incorporating a Model of an Experienced Human Tutor. In *Proceedings of the 6th International Conference on Intelligent Tutoring Systems*, pages 596–608, London, UK.

Pamela W. Jordan and Marilyn A. Walker. 1996. Deciding to Remind During Collaborative Problem Solving: Empirical Evidence for Agent Strategies. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence AAAI*, volume 1, pages 16–23, Portland, OR.

Elena Karagjosova. 2003. Marked informationally redundant utterances in tutorial dialogue. In Ivana Kruijff-Korbayová and Claudia Kosny, editors, *Proceedings of the 7th Workshop on the Semantics and Pragmatics of Language (DiaBruck)*, pages 189–190, Saarbrücken, Germany.

J. F. Kelley. 1984. An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems*, 2(1):26–41.

Victor R. Lee and Bruce L. Sherin. 2004. What makes teaching special? In *ICLS '04: Proceedings of the 6th international conference on Learning sciences*, pages 302–309, Santa Monica, CA.

D. McArthur, C. Stasz, and M. Zmuidzinas. 1990. Tutoring techniques in algebra. *Cognition and Instruction*, 7(3):197–244.

D. C. Merrill, B. J. Reiser, S. K. Merrill, and S. Landes. 1995. Tutoring: Guided learning by doing. *Cognition and Instruction*, 13:315–372.

Roger H. Munger. 1996. Asymmetries of knowledge: What tutor-student interactions tell us about expertise. Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Milwaukee, WI.

Kai Pata, Tago Sarapuu, and Raymond Archee. 2005. Collaborative scaffolding in synchronous environment: congruity and antagonism of tutor/student facilitation acts. In D. D. Suthers T. Koschman and T.-W. Chan, editors, *Computer Supported Collaborative Learning 2005: The next 10 years*, pages 484–493. Kluwer.

Natalie K. Person, Laura Bautista, Roger J. Kreuz, Arthur C. Graesser, and the Tutoring Research Group. 2000. The Dialog Advancer Network: A Conversation Manager for AutoTutor. In *Proceedings of the Workshop on modeling human teaching tactics and strategies at the 5th International Conference on Intelligent Tutoring Systems (ITS-00)*, LNCS, pages 86–92, Montreal, Canada.

Jörg Siekmann, Christoph Benzmüller, and Serge Autexier. 2006. Computer Supported Mathematics with ΩMEGA. *Journal of Applied Logic*, 4(4):533–559.

Robert Stalnaker. 2002. Common ground. *Linguistics and Philosophy*, 25(5):701–721.

David R. Traum. 1994. A computational theory of grounding in natural language conversation. Technical Report TR545, University of Rochester, Rochester, NY, USA.

Dimitra Tsovaltzi, Armin Fiedler, and Helmut Horacek. 2004. A multi-dimensional taxonomy for automating hinting. In James C. Lester, Rosa Maria Vicari, and Fábio Paraguaçu, editors, *Intelligent Tutoring Systems — 7th International Conference (ITS-04)*, number 3220 in LNCS, pages 772–781. Springer.

Marilyn Walker and Owen Rambow. 1994. The Role of Cognitive Modeling in Achieving Communicative Intentions. In *Proceedings of the 7th International Workshop on Natural Language Generation.*, pages 171–180, Kennebunkport, ME.

Marilyn Walker. 1993. *Informational Redundancy and Resource Bounds in Dialogue*. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

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-Korbayova, Elena Karagjosova, Helmut Horacek, Malte Gabsdil, Armin Fiedler, and Christoph Benzmüller. 2004. An annotated corpus of tutorial dialogs on mathematical theorem proving. In *Proceedings of International Conference on Language Resources and Evaluation (LREC-04)*, pages 1007–1010, Lisbon, Portugal.

Claus Zinn. 2004. Flexible dialogue management in natural-language enhanced tutoring. In *Konvens 2004 Workshop on Advanced Topics in Modeling Natural Language Dialog*, pages 28–35, Vienna, Austria.