A Grounding Approach to Modelling Tutorial Dialogue Structures

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

Abstract

Pedagogically motivated analyses of tutorial dialogue have identified recurring local sequences of exchanges which we propose to be analysed analogously to grounding structures. In this paper, we present a model describing such local structures in which a learner and a tutor collaboratively contribute to building a solution to a task. Such structures are modelled as "grounding" exchanges which operate at the task level, i.e. at the level of deep understanding of the domain. Grounding a learner's contributions depends on the tutor's beliefs as to the learner's level of understanding. We treat this explicitly by requiring sufficient domain-level evidence to be shown for a contribution to be grounded. This work attempts to link general theories of dialogue with observations from pedagogical science.

1 Motivation

Successful conversational communication depends strongly on the coordination of meanings and background assumptions as to the state of the world (Clark, 1992; Stalnaker, 2002; Thomason et al., 2006). Dialogue participants try to achieve a situation in which they mutually believe that their utterances are interpreted as intended and that their assumptions as to the shared knowledge, the *common ground*, agree. To this end, they engage in a process called *grounding* (Clark and Schaefer, 1989; Traum, 1999), whose purpose is to ensure explicit alignment of (mutual) beliefs. Grounding can serve to avoid or recover from communication failures arising from problems which may range from low level signal-related issues through the interpretation of the propositional content up to the level the communicative intentions of speech acts.

Grounding is a general pragmatic phenomenon in cooperative communication that is independent of the purpose of the verbal activity, be it socially-motivated spontaneous conversation or task oriented verbal communication such as information seeking, negotiation, problem solving or dialogue-based instruction. The latter scenario is additionally inherently prone to misalignment of beliefs beyond the level of the communicative intentions of speech acts: namely at the level of *deep understanding* of the tutored domain. First, tutoring is typically characterised by an asymmetry of knowledge possessed by the tutor and the learner (Munger, 1996; Lee and Sherin, 2004). Second, there is an uncertainty on the part of the tutor as to the learner's deep understanding and the overall knowledge state. In fact, empirical research shows that tutors tend to have difficulties in estimating the learner's deep understanding (Chi et al., 2004). Still, dialogue-based one-on-one instruction, even by non-experts, has been shown to produce higher learning gains than other forms of teaching (Bloom, 1984; Moore, 1993). One of the factors that makes a difference in the efficiency of instruction is adaptivity of tutorial feedback and explanation. Nückles et al. (2006) show that tutors who are better informed on the learners' prior knowledge can better adapt their feedback. Another important feature of efficient tutoring are locally targeted pedagogical actions. Graesser et al. (1995) show in an empirical study that tutors typically do not focus on cognitive alignment, i.e. do not strive to establish complete understanding of the students' state of beliefs. Instead they tend to perform specific targeted tutoring moves that locally address

Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue, June 2–4, 2008, London, U.K.

the student's (lack of) progress on the task at hand.

Motivated by these findings we have been investigating discourse and dialogue phenomena in the context of dialogue-based tutoring with the ultimate goal of building a tutoring system for mathematical theorem proving. Our approach to modelling tutorial dialogue draws on the empirical evidence from the above-mentioned studies and can be summarised by the following observations:

On the one hand, tutorial dialogue is in many respects different from other types of dialogue. The model of cooperative interpretation must address the learner's utterances not only in terms of their function as speech acts, but also as demonstrations of knowledge, that is, it is dependent on the adopted pedagogical strategy. Motivated by pedagogical goals and licenced by his authority, the tutor may be the "uncooperative" interlocutor in the sense that he/she may demand presentation of pieces of knowledge that the learner had left implicit, or may even refuse to provide information requested by the learner (overriding dialogue obligations valid in other dialogue genres) attempting to lead the learner to self-discovery of knowledge. The structure of tutorial dialogue is moreover characterised by systematically recurring sub-structures. The role of these is to address the learner's knowledge contributions and to monitor, at least to some extent, the learner's deep understanding, allowing feedback and cooperative behaviour to be adapted to what the student has previously shown to have understood.

On the other hand, tutorial dialogue is still a type of dialogue, that is, it is characterised by the general phenomena present in any dialogue genre and should lend itself to modelling in terms of general notions of dialogue. However, because it is a special type of dialogue, the model's parameters (e.g. the contents of the information state, models of dialogue state transitions, obligations, and cooperativity) must be adjusted to the genre's characteristics.

This work is an attempt to apply notions from general dialogue theory to tutorial dialogue. In particular, we will try to show parallels between the structure of grounding at the speech acts level and the local structures in tutorial dialogue which resemble grounding, but address the deep understanding of the domain. We will call these *communication level grounding* and *task level grounding* respectively. We start by exemplifying these local structures with dialogue excerpts from our corpora (Section 2.1). In Section 2.2 we briefly introduce grounding according to Traum (1999). In Section 2.3 we present our framework for task level grounding and point at the differences between it and Traum's model. Section 3 presents the model formally and steps through an example, before we summarise our conclusions in Section 4.

2 Tutorial Dialogue Structures as Grounding Exchanges

Tutorial dialogues have been shown to exhibit local patterns, referred to by Graesser et al. (1995) as *dialogue frames*, related to the pedagogical goals that tutors follow. We will argue that the structure of dialogue frames is similar in character to that of Traum's Discourse Units (Traum, 1999), the basic building blocks of which are utterances which contribute to achieving mutual understanding. Our goal is to attempt to unify these two views on (tutorial) dialogue structure in a grounding-based model of tutorial dialogue, which we present in the next section.

2.1 Dialogue Frames in Tutoring

In a corpus-based analysis of the collaborative nature of tutorial dialogue Graesser et al. identify local interaction patterns which make one-on-one tutoring, even by non-experts, effective in producing learning gains. They consist of the following steps, all of which but step 2 may be omitted:

Step 1 Tutor asks a question.

Step 2 The student offers an answer

Step 3 Tutor gives feedback on the answer

Step 4 Tutor and student collaboratively improve the quality of the answer, whereby the tutor can for instance elaborate on the answer, give a hint, pump the student for more information, or trace an explanation or justification.

Step 5 The tutor assesses the student's understanding of the answer, for instance by explicitly asking whether the student understood.

Similar structures were revealed by our analysis of the two corpora of tutorial dialogues on mathematical theorem proving (Wolska et al., 2004;

Benzmüller et al., 2006) which we have collected.[1] (1) and (2) are examples of such exchanges[2]:

(1)
S0 $(R \circ S)^{-1} = \{(x, y)|(y, x) \in (R \circ S)\}$
T0 correct
S1 $(R \circ S)^{-1} = \{(x, y)|(y, x) \in \{(x, y)|\exists z(z \in M \wedge (x, z) \in R \wedge (z, y) \in S\}\}$
T1-1 okay,
T1-2 but you could have done that more simply

(2)
S19-1: $(R \circ T) = (T^{-1} \circ R^{-1})^{-1}$ (by exercise W),
S19-2: then it must also hold that $(S \circ T) = (T^{-1} \circ S^{-1})^{-1}$
T25: Why does this follow from exercise W?
S20: $(R \circ S) = (S^{-1} \circ R^{-1})^{-1}$ (according to exercise W), then it must also hold that $(S \circ T) = (T^{-1} \circ S^{-1})^{-1}$ and $(R \circ T) = (T^{-1} \circ R^{-1})^{-1}$
T26-1: All other steps are appropriate,
T26-2: but the justification for $(R \circ T) = (T^{-1} \circ R^{-1})^{-1}$ is still missing.
S21: $(R \circ T)^{-1} = (T^{-1} \circ R^{-1})$ (by exercise W)
T27: Yes.

The building block of such exercises is the *proof step*, a contribution which consists of a formula which the step derives, a justification, premises, and possibly other components. Proof steps may be underspecified, for instance by only providing the derived formula. This leads to them possibly having to be augmented in order to be acceptable.

In (1) we see a simple case of a student's contribution being accepted by the tutor. In terms of Graesser's dialogue frames, S0 corresponds to step 2 and T0 to step 3. Because the tutor is immediately satisfied that the student has understood the answer, steps 4 and 5 are not performed. S1 and T1 form a new dialogue frame which is the same as the first except that step 4 is realised in T1-2 by the tutor, who elaborates on the answer.

(2) is a more complex example which begins with the student's contribution in S19 (Graesser's step 2). It consists of two contributions, however only the first one (S19-1) is discussed. Similarly to S0 in (1), the contribution is incomplete in that the student does not provide the premise that allowed him/her to conclude that the contribution in S19-1 holds, but rather leaves it implicit. Here however, the tutor is not satisfied with the incomplete step and responds with a request to elaborate the answer (step 4) in T25, skipping the feedback (step 3). Instead of addressing this request, the student offers a new contribution, leaving the request in T25 pending. In T26-1 the tutor gives feedback on both S19 and S20 (step 3 for both of these contributions) and continues by repeating the request for elaboration in T26-2. The student then addresses this request by supplying the missing premise in S21 (step 4) which the tutor accepts in T27, thereby closing the dialogue frame.

Our analysis of the two tutorial dialogue corpora revealed that structures such as the ones described above systematically recurred in the domain of proof tutoring in the context of the conducted experiment. Locally, the dialogue structures indeed typically reflect Graesser's steps 2 though 4, with individual proof steps being proposed (step 2) and subsequently optionally elaborated (step 3) and evaluated (step 4), in either order. Due to the student having the initiative in our experimental setup step 1 is seldom found in our data.

In the corpora elaboration requests were most commonly initiated because the inferences proposed by the students were only partially specified. Typically, the students provide a formula (or an equivalent worded statement) leaving out, for instance, the inference rule, the way it should be applied or the premises. This means that part(s) of the task-level steps are left implicit (or *tacit*), possibly resulting in them not being grounded. In the tutoring domain the question of whether an underspecified step (or more generally, an incomplete knowledge demonstration) can be accepted (i.e. grounded) depends, for instance, on pedagogical factors (in the case of mathematical proofs, for example on the tutor's notion of an "acceptable" proof (Raman, 2002)) and the tutor's beliefs as to the student's knowledge state.

2.2 The Grounding Acts Model

Traum (1999) defines a set of Grounding Acts which are identified with particular utterance units and perform specific functions towards the achievement of mutual understanding. The content of an utterance can become grounded as a result of an exchange containing Grounding Acts; such possibly multi-turn sequences are referred to as *Discourse Units* (DU). DUs can contain the following acts:

Initiate begins a new DU with a new utterance unit.

Continue adds content to an act.

[1] The corpora were collected in Wizard-of-Oz experiments. The 2004 corpus contains 22 dialogues (775 turns in total) in the domain of naive set theory. The 2006 corpus contains 37 dialogues (1917 turns) in the domain of binary relations.

[2] Sx and Tx label student and tutor utterances respectively.

Acknowledge is evidence of understanding of the current utterance unit. This evidence can be of differing strength, e.g. demonstration of the understood meaning or performance of a relevant continuation.

Repair changes the content of the current utterance unit.

ReqRepair requests repair of a previous act by signalling non-understanding.

ReqAck asks the dialogue participant to acknowledge understanding of a previous act.

Cancel abandons the current DU without grounding.

In the DU in example (3), taken from Traum (1999), the content of the initiating utterance I1-1 and the continuation I1-2 has been successfully grounded by the acknowledgement Grounding Act in R1.

(3)	**I1-1** Move the box car to Corning	init^I
	I1-2 and load it with oranges	cont^I
	R1 ok	ack^R

Our previous example (1) also exhibits this structure within the DU, where S0 and S1 are initiations and T0 and T1-1 are acknowledgements. We now show that in tutorial dialogue, in addition to the communicative level of Traum's model, grounding also operates at the task level.

2.3 Grounding in Tutorial Dialogue

We have exemplified the parallels between the structures found in tutorial dialogue and grounding exchanges and will now make these parallels more explicit. We will interpret dialogue frames as discourse units and the actions within dialogue frames as grounding acts.

What is grounded in the course of a discourse unit is a piece of domain content which contributes to the domain-level task. In tutoring this is a knowledge demonstration — we will use the term *solution step*. Proof steps become grounded by being first proposed and then accepted by the tutor, provided that the tutor has sufficient evidence to believe that the student has deeply understood how the step was derived. To reach this state the student may be obliged to supply evidence of having understood the step, and this evidence can be of varying strength. In this sense supplying evidence is similar to Traum's **Acknowledge**, and a request for evidence is similar to **ReqAck**. We list the set of actions as well as who can perform them in the course of grounding a solution step in Table 1.

Propose	S,T	propose a solution step
ReqEv	S,T	request evidence showing understanding of the current step
SuppEv	S	give evidence showing understanding of the current step
Accept	T	accept that the student has understood the current step
Reject	T	reject the step (due to incorrectness or non-understanding)

Table 1: Task level grounding actions and speakers

Augment	an elaboration of the current step
Reword	paraphrase of the current step
Claim	positive answer to "do you understand?"
Verbatim	repeat back the step verbatim

Table 2: Types of evidence of understanding

In the same way that Clark and Schaefer (1989) identify different types of evidence of understanding, the action **SuppEv** encompasses a number of different ways of showing understanding of a solution step. From our analysis of the data, we propose the four categories listed in Table 2 from strongest to weakest. Although verbatim repetition of the content being grounded is the strongest evidence type in Clark and Schaefer's communication level grounding model, at the task level it is the weakest form, since it does not show any understanding beyond recognition of the original signal. Claiming understanding is self-reflection on the student's own belief state, and for our purposes is a weak form of evidence. Rewording is a strong indication of understanding, but does not add anything to the current content which is being grounded. The strongest evidence type is augmenting the current solution step with further information. This shows that the student understands even those components which were not stated in the proposal phase of the discourse unit. In keeping with Clark and Schaefer's observation that evidence must be "sufficient for the current purpose", the tutor's decision of whether to consider this evidence *sufficient* to show understanding of the current content (and then to accept the step) depends on both a student model and the pedagogical strategy being followed. Indeed for different teaching domains this notion of *sufficient* will be defined differently according to the demands of the task and the domain dependent teaching goals.

According to this model the annotation of example (2) from the previous section, where subscripts index individual steps under discussion, is:

S19-1	**Propose**$_1$
S19-2	**Propose**$_2$
T25	**ReqEv**$_1$
S20	**Propose**$_3$
T26-1	**Accept**$_{2,3}$
T26-2	**ReqEv**$_1$
S21	**SuppEv**$_1$ (Augment: premise)
T27	**Accept**$_1$

The proposal made in S19-1 is eventually grounded in T27, but in between a new proposal is made (S20), showing that more than one solution step can be under discussion at once.

Contrasts with Traum's model We have borrowed many concepts from and shown parallels to Traum's Grounding Acts model, so here it is useful to highlight some key differences. The main difference is that task level grounding works at a higher level than Traum's communication level grounding model. Our model does not deal with meaning but rather with deep understanding, and the object being grounded is part of the task being explicitly talked about. Accordingly, actions contributing to task level grounding are motivated by task level goals, such as completing the current exercise, whereas Traum's Grounding Acts contribute to successful communication as a whole. Communication level grounding does however still operate as usual in parallel. We refer to example (2), in which the utterance T25 has two functions: at the communication level it grounds the propositional content initiated in S19 but at the task level it continues the discourse unit. A further difference is the roles of dialogue participants and their goals. In tutoring our model does not consider the roles of speaker and hearer, but rather student and tutor, necessary because of the asymmetry of roles in tutorial dialogue; students are obliged to demonstrate understanding but tutors are not.

In summary, we have found a correspondence between general grounding structures and the structures found in tutoring. In order to treat these subdialogues in terms of grounding we need a model of grounding with a higher level object: the task level step. In the next section we introduce the more formal machinery to model these sequences.

3 A Model of Task Level Grounding

Our discourse unit is a subdialogue which begins with the proposal of a task level step and which ends with this step being either accepted or rejected by the tutor. In the previous section we have motivated this choice by showing its equivalence to both Graesser's dialogue frames and Traum's Discourse Units. The objects which are under discussion and which are to be grounded in these subdialogues are *solution steps*, here proof steps, and the conditions which affect this are a student model, the tutor's pedagogical strategy, the correctness, relevance and granularity of the step, as well as some definition of what it means for evidence to be sufficient. The internal structure of solution steps should be defined for the task at hand — here we use a solution step for mathematical proofs consisting of a formula which is derived, a justification for deriving the formula, and the premises used by the justification. In this section we present the machinery necessary to model these phenomena and step through example (2).

We assume that the dialogue system has access to two expert systems: a pedagogical manager and a mathematical domain reasoner. The pedagogical manager (Fiedler and Tsovaltzi, 2003) is responsible for the teaching strategy that the system follows, as well as for maintaining the student model. The domain reasoner (Dietrich and Buckley, 2007; Schiller et al., 2007) evaluates solution steps with respect to correctness, granularity and relevance, and can resolve missing components of underspecified steps.

The model uses the categorisations of utterance types in terms of their function in the DU (Table 1) and evidence types (Table 2) that play a role in the grounding exchanges we are considering. We will now additionally define a dialogue state which represents intermediate stages of the discourse unit, followed by a finite state machine which encodes the transitions between dialogue states and their effects.

3.1 Dialogue State

The dialogue state used in our model is an extension of our previous work on common ground (Buckley and Wolska, 2007), reduced to those aspects relevant to this presentation. It consists of four parts and is shown in Figure 1. The common ground (CG)

$$
\begin{bmatrix}
\textsc{cg} & \begin{bmatrix}\textsc{s-steps} & \langle\rangle\end{bmatrix} \\
\textsc{pending} & \left\langle \begin{bmatrix} \textsc{id} & \texttt{step1} \\ \textsc{formula} & \text{“}(R \circ T) = \ldots\text{”} \\ \textsc{justifs} & \text{“}\texttt{exercise\_W}\text{”} \\ \textsc{premises} & \end{bmatrix} \right\rangle \\
\textsc{lu} & \begin{bmatrix} \textsc{speaker} & \texttt{student} \\ \textsc{moves} & \texttt{propose, ...} \\ \textsc{s-step} & \texttt{step1} \end{bmatrix} \\
\textsc{output} &
\end{bmatrix}
$$

Figure 1: The dialogue state

contains an ordered list[3] of the solution steps which have been grounded in the process of solving the task (S-STEPS). The solution steps which are currently under discussion but are not yet grounded are stored in PENDING. The latest utterance (LU) in the dialogue is represented by a structure containing information about the speaker who performed the utterance, the dialogue moves it realised, and the solution step, if any, that it contained. Finally the dialogue moves that the system should verbalise next are collected in OUTPUT. Both LU/MOVES and OUTPUT store complete dialogue moves, however here we only list task-level grounding actions. When task-level grounding has been successful, the solution step moves from PENDING to CG/S-STEPS.

3.2 Transitions between Dialogue States

Figure 2 presents a finite state machine encoding the transitions between dialogue states in a discourse unit. A **Propose** moves the dialogue into a state in which there is an ungrounded solution step. From here the tutor can either accept the step directly, thus grounding the step, or ask for further evidence of understanding, after which it is necessary for the student to supply evidence before the discourse unit can be in the state in which the solution step is grounded.

The transitions (Table 3) are given as sets of preconditions and effects on the dialogue state. We omit additional processing such as information exchange with system modules. The conditions we use are stated informally — “evidence (in)sufficient” is decided by the pedagogical module, drawing on information from the dialogue state as well as its own

t1	pre eff	**Propose** $\in$ LU/MOVES PENDING := LU/S-STEP
t2	pre eff	evidence insufficient, ne(PENDING) OUTPUT := **ReqEv**
t3	pre eff	evidence sufficient, ne(PENDING) OUTPUT := **Accept**,(feedback) push(CG/S-STEPS,pop(PENDING))
t4	pre eff	**SuppEv** $\in$ LU/MOVES possibly update solution step
t5	pre eff	evidence insufficient OUTPUT := **ReqEv**
t6	pre eff	evidence sufficient OUTPUT := **Accept**,(feedback) push(CG/S-STEPS,pop(PENDING))

Table 3: Preconditions and effects of transitions (ne denotes “non-empty”)

student model. Transition t3 moves from a state in which a solution step has been proposed to a state in which that solution step has been grounded. If the evidence for understanding the step is sufficient, and there is content under discussion (ne(PENDING)), then an **Accept** and possibly some feedback is generated, and the solution step is moved from PENDING to CG/S-STEPS. This transition equates to Graesser’s step 3 in the dialogue frame. Transitions t2 and t5 both cover the situation where the evidence presented is not sufficient to show understanding, and both result in **ReqEv** being generated, and the solution step(s) that were in PENDING remain there (Graesser’s step 4). When evidence is supplied, we follow transition t4, which updates the solution step in the event that evidence of the type **Augment** was supplied. Although it is not included in the FSA, at any stage a discourse unit can be abandoned, possibly with a **Reject** action. This decision can be taken for instance in the state “evidence supplied” when the tutor believes that the student will not be able to show understanding of the step.

Because there can be more than one solution step under discussion at one time, as in example (2), we assume that a separate instance of the FSA is run for each one. An acceptance can thus address more than one solution step. Like downdating questions under discussion, we allow acceptances to ground as many solution steps as necessary. We also note that transitions in the model are only made in reaction to task-level grounding actions, so that as long as other actions are being performed, the FSA stays

[3] This is a strong simplification — a complete treatment would require a more detailed structure for solution steps.

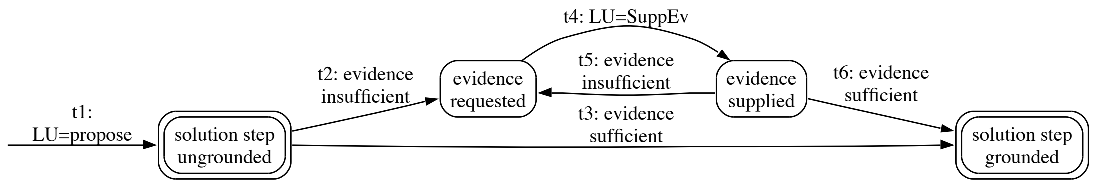

Figure 2: The FSA describing task-level discourse units

in the same state. This allows other levels in the dialogue to be modelled, for instance communication level grounding, off-topic talk or meta talk. Indeed this model can be integrated with a computational model of communication level grounding such as presented by Matheson et al. (2000) if we assume that their grounding acts are dealt with before generating any task level grounding actions. This way problems at the communication level are handled before understanding problems at the task level.

Example Figure 1 shows the dialogue state after utterance S19 in example (2), where the **Propose** in utterance S19-1 has put a solution step in PENDING. The tutor considers that with the current context and student model, there is not sufficient evidence of understanding of the solution step. Transition t2 is therefore executed, generating a **ReqEv** action, realised in utterance T25. Skipping forward to S21 (S20 to T26-2 deal similarly with a different solution step), we recognise a **SuppEv** action, which takes us through transition t4. Since the evidence supplied in S21 is of type **Augment**, we update the solution step by adding the premise the student stated as shown:

$$\left[\begin{array}{ll} \textsc{pending} & \left\langle \left[\begin{array}{ll} \textsc{id} & \texttt{st1} \\ \textsc{formula} & \text{“}(R \circ T) = \ldots\text{”} \\ \textsc{justifs} & \text{“}\texttt{exercise\_W}\text{”} \\ \textsc{premises} & \text{“}(R \circ T)^{-1} = \ldots\text{”} \end{array}\right] \right\rangle \end{array}\right]$$

Now the tutor can reassess whether this more complete solution step is evidence that the student has understood fully, and finds that it is. The transition t6 then generates the **Accept** in T27 and additionally moves the solution step to the common ground. The final dialogue state is shown in Figure 3.

$$\left[\begin{array}{ll} \textsc{cg} \mid \textsc{s-steps} & \left\langle \left[\begin{array}{ll} \textsc{id} & \texttt{st1} \\ \textsc{formula} & \text{“}(R \circ T) = \ldots\text{”} \\ \textsc{justifs} & \text{“}\texttt{exercise\_W}\text{”} \\ \textsc{premises} & \text{“}(R \circ T)^{-1} = \ldots\text{”} \end{array}\right] \right\rangle \\ \textsc{pending} & \langle\rangle \\ \textsc{lu} & \left[\begin{array}{ll} \textsc{speaker} & \texttt{tutor} \\ \textsc{moves} & \texttt{accept, ...} \\ \textsc{s-step} & \end{array}\right] \\ \textsc{output} & \end{array}\right]$$

Figure 3: Final dialogue state

4 Conclusions and Related Work

We take advantage of observations about recurring local structures in tutorial dialogue highlighted by Graesser's analysis and recognise that there exist systematic parallels between these and Traum's grounding exchanges. This motivates our computational model, which is analogous to Traum's but operates on a level which directly addresses understanding of the domain. Our model sees these local structures as discourse units whose objects are solution steps, and thus operates at the task level. It captures learners' deep understanding of the domain, and so acts higher than the communication level.

Grounding serves to build up a model of interlocutors' belief states. In tutoring this is particularly important because the tutor's model of the student's belief state is a parameter which affects the adopted pedagogical strategy. The local dialogue structure that our model describes allows the pedagogical model to elicit evidence of understanding and thus reach conclusions about the student's belief state. While we do not make any claims about how such a student model should be constructed, our model does provide input for the construction of a representation of the student's knowledge.

Rickel et al. (2002) also use a general dialogue model in a tutoring system which combines pedagogical expertise with collaborative discourse theory and plan recognition. Their approach models the knowledge state based on steps that the student has been exposed to, however without considering whether these were fully understood. Zinn et al. (2005) present a tutorial dialogue system which maintains common ground in the dialogue model, however they do not make use of grounding status to structure the dialogue locally. Baker et al. (1999) highlight the necessity for communication level grounding in collaborative learning, but admit that this does not guarantee "deeper" understanding. In general task-oriented dialogues Litman and Allen (1987) derive the structure of clarification subdialogues based on task plans and the discourse structure. Our approach is conceptually similar, however our task model is maintained externally to the dialogue model. Finally, our work relates to that of Thomason et al. (2006) and Benotti (2007) in the sense that the task level grounding model attempts to ground objects that can be viewed as tacit actions.

Our future work will include extending the model to allow more student initiative, for example in the case of domain level clarification requests by the student, as well as looking into more fine-grained structures within the common ground, for instance to support a model of the salience of task level objects.

References

M. Baker, T. Hansen, R. Joiner, and D. Traum. 1999. The role of grounding in collaborative learning tasks. In *Collaborative Learning. Cognitive and computational approaches*, pages 31–63. Pergamon, Amsterdam.

L. Benotti. 2007. Incomplete knowledge and tacit action: Enlightened update in a dialogue game. In *Proc. of DECALOG-07*, Rovereto, Italy.

C. Benzmüller, H. Horacek, H. Lesourd, I. Kruijff-Korbayová, M. Schiller, and M. Wolska. 2006. A corpus of tutorial dialogs on theorem proving; the influence of the presentation of the study-material. In *Proc. of LREC-06*, pages 1766–1769, Genoa, Italy.

B. Bloom. 1984. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6):4–16.

M. Buckley and M. Wolska. 2007. Towards Modelling and Using Common Ground in Tutorial Dialogue. In *Proc. of DECALOG-07*, pages 41–48, Rovereto, Italy.

M. T. H. Chi, S. A. Siler, and H. Jeong. 2004. Can Tutors Monitor Students' Understanding Accurately? *Cognition and Instruction*, 22(3):363–387.

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

H. H. Clark, editor. 1992. *Arenas of Language Use*. University of Chicago Press and CSLI.

D. Dietrich and M. Buckley. 2007. Verification of Proof Steps for Tutoring Mathematical Proofs. In *Proc. of AIED-07*, pages 560–562, Los Angeles, USA.

A. Fiedler and D. Tsovaltzi. 2003. Automating Hinting in Mathematical Tutorial Dialogue. In *Proc. of the EACL-03 Workshop on Dialogue Systems: Interaction, Adaptation and Styles of Management*.

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

V. R. Lee and B. L. Sherin. 2004. What makes teaching special? In *Proc. of ICLS-04*, pages 302–309.

Diane J. Litman and James F. Allen. 1987. A Plan Recognition Model for Subdialogues in Conversation. *Cognitive Science*, 11(2):163–200.

C. Matheson, M. Poesio, and D. Traum. 2000. Modelling grounding and discourse obligations using update rules. In *Proc. of NAACL-00*, pages 1–8.

J. Moore. 1993. What makes human explanations effective? In *Proc. of the 15^{th} Meeting of the Cognitive Science Society*, pages 131–136, Hillsdale, NJ.

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

M. Nückles, J. Wittwer, and A. Renkl. 2006. How to make instructional explanations in human tutoring more effective. In *Proc. of the 28th Annual Conference of the Cognitive Science Society*, pages 633–638.

M. J. Raman. 2002. *Proof and Justification in Collegiate Calculus*. Ph.D. thesis, UC Berkeley.

J. Rickel, N. Lesh, C. Rich, C. Sidner, and A. Gertner. 2002. Collaborative discourse theory as a foundation for tutorial dialogue. In *Proc. of ITS-02*, pages 542–551.

M. Schiller, D. Dietrich, and C. Benzmüller. 2007. Towards computer-assisted proof tutoring. In *Proc. of 1st SCOOP Workshop*, Bremen, Germany.

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

R. Thomason, M. Stone, and D. DeVault. 2006. Enlightened update: A computational architecture for presupposition and other pragmatic phenomena. In *Presupposition Accommodation*. (draft).

D. Traum. 1999. Computational models of grounding in collaborative systems. In *Working Papers of the AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*, pages 124–131.

M. Wolska, B. Q. Vo, D. Tsovaltzi, I. Kruijff-Korbayova, E. Karagjosova, H. Horacek, M. Gabsdil, A. Fiedler, and C. Benzmüller. 2004. An annotated corpus of tutorial dialogs on mathematical theorem proving. In *Proc. of LREC-04*, pages 1007–1010, Lisbon.

C. Zinn, J. D. Moore, and M. G. Core. 2005. Intelligent information presentation for tutoring systems. In *Intelligent Information Presentation*. Kluwer.