DISCUSS: A dialogue move taxonomy layered over semantic representations

Lee Becker[1] Wayne H. Ward[1,2] Sarel van Vuuren[1] Martha Palmer[1]
{lee.becker, martha.palmer, sarel.vanvuuren}@colorado.edu,
wward@bltek.com

[1]University of Colorado at Boulder, [2]Boulder Language Technologies

Abstract

In this paper we describe DISCUSS, a dialogue move taxonomy layered over semantic representations. We designed this scheme to enable development of computational models of tutorial dialogues and to provide an intermediate representation suitable for question and tutorial act generation. As such, DISCUSS captures semantic and pragmatic elements across four dimensions: Dialogue Act, Rhetorical Form, Predicate Type, Semantic Roles. Together these dimensions provide a summary of an utterance's propositional content and how it may change the underlying information state of the conversation. This taxonomy builds on previous work in both general dialogue act taxonomies as well as work in tutorial act and tutorial question categorization. The types and values found within our taxonomy are based on preliminary observations and on-going annotation from our corpus of multimodal tutorial dialogues for elementary school science education.

1 Introduction

Past successes with conversational Intelligent Tutoring Systems (ITS) (Graesser et al., 2001), have helped to demonstrate the efficacy of computer-led, tutorial dialogue. However, ITS will not reach their full potential until they can overcome current limitations in spoken dialogue technologies. Producing systems capable of leading open-ended, Socratic-style tutorials will likely require more sophisticated models to automate analysis and generation of dialogue. A well defined tutorial dialogue annotation scheme can serve as a stepping stone towards these goals. Such a scheme should account for differences in tutoring style and question scaffolding techniques and should capture the subtle distinctions between different question types. To do this, requires a representation that connects a turn's communicative and rhetorical functions to its underlying semantic content.

While efforts such as DAMSL (Core and Allen, 1997) and DIT++ (Bunt, 2009) have helped to make dialogue act annotation more uniform and applicable to a wider audience, and while tutoring-specific initiatives (Tsovaltzi and Karagjosova, 2004; Buckley and Wolska, 2008) have helped to bring dialogue acts to tutorial dialogue, the move granularity in these schemas is too coarse to capture the differences in tutorial questioning styles exhibited in our corpus of Socratic-style tutorial dialogues. Conversely, question type categories (Graesser and Person, 1994; Nielsen et al., 2008) have been designed with education in mind, but they largely ignore how the student and tutor may work together to construct meaning. The DISCOUNT scheme's (Pilkington, 1999) combination of dialogue acts and rhetorical functions enabled it to better capture tutoring moves, but its omission of shallow semantics prevents it from capturing how content influences behavior.

Our long-term goals of automatic dialogue characterization, tutorial move prediction and question generation led us to design our own dialogue representation called DISCUSS (Dialogue Scheme for Unifying Speech and Semantics). Design of this dialogue move taxonomy was based on preliminary observations from our corpus of tutorial dialogues, and was influenced by the aforementioned research. We hope that undertaking this ambitious endeavor to capture not only a turn's pragmatic interpretation,

but also its rhetorical and semantic functions will enable us to better model the complexity of open-ended, tutorial dialogue.

The remainder of the this paper is organized as follows. In the next section we describe our tutorial dialogue setting and our data. Section 3 discusses the organization of the DISCUSS annotation scheme. Section 4 briefly explains the current status of our annotation. Lastly section 5 outlines our future plans and conclusions.

2 Tutorial Dialogue Setting and Data

My Science Tutor (MyST) (Ward et al., 2010) is a conversational virtual tutor designed to improve science learning and understanding for students in grades 3-5. Students using MyST investigate and discuss science through natural spoken dialogues and multimedia interactions with a virtual tutor named Marni. The MyST dialogue design and tutoring style is based on a pedagogy called Questioning the Author (QtA) (Beck et al., 1996), wherein the teacher facilitates discovery by challenging students with open-ended questions and by directly keying in on ideas expressed in the student's language.

To gather data for MyST system coverage and dialogue analysis, we ran Wizard-of-Oz (WoZ) experiments that allowed a human tutor to be inserted into the interaction loop. Project tutors trained in QtA served as Wizards and were responsible for accepting and overriding system actions. Over the past three years we have accumulated over five-hundred, 15-minute WoZ sessions across four modules Magnetism and Electricity, Measurement, Variables, and Water, each with 16 lessons. Student speech from these sessions was professionally transcribed at the word level.

3 The DISCUSS Annotation Scheme

The Dialogue Scheme for Unifying Speech and Semantics (DISCUSS) is a multifaceted dialogue move taxonomy intended to capture both the pragmatic and semantic interpretations of an utterance. A DISCUSS move is a tuple composed of values from four dimensions: *Dialogue Act*, *Rhetorical Form*, *Predicate Type*, and *Semantic Roles*. Together these dimensions convey the communicative action, surface form, and meaning of an utterance independent of the original utterance text.

We designed DISCUSS to serve as an intermediate representation that will enable future work in dialogue session characterization, dialogue strategy optimization, and automatic question generation. To facilitate these goals, we have endeavored to create a taxonomy that is both descriptive and curriculum-independent while allowing for expansion as necessary. A complete listing of all the DISCUSS moves and dimensions can be found in our forthcoming technical report.

In the following subsection we will describe the different DISCUSS move categories. Descriptions of the *Semantic Role* and *Predicate Type* are found in the subsection about semantic dimensions, while discussion about the *dialogue act* and *rhetorical form* has been placed in the pragmatic dimensions subsection. Throughout the rest of this paper we denote DISCUSS tuples using the following notation: Dialogue Act/Rhetorical Form/Predicate Type ⟨Semantic Role⟩.

3.1 Move Categories

DISCUSS moves are dictated by the dialogue act dimension and may belong to one of three broad categories: *Dialogue Control*, *Information Exchange*, and *Attention Management*. Dialogue Control moves are largely concerned with maintaining and enabling the flow of information. This includes dialogue acts such as *Acknowledge, Open, Close, Repeat, and RequestRepeat*. The Information Exchange moves relay content (often lesson-specific) between speakers using moves such as *Assert, Ask, Answer, Mark, Revoice*. For tutorial dialogue the bulk of student-tutor interactions reside in this category. Lastly, Attention Management moves indicate how a speaker exercises initiative over other speakers or topics. Dialogue acts found in the attention category are *Focus, Defer, Elicit, and Direct*.

3.2 Semantic Dimensions

The semantic dimensions define the objects, events, properties and relations contained within an utterance. The semantic roles at the lowest level of the DISCUSS hierarchy directly capture the propositional entities. Predicate Types summarize the interactions between all of the semantic roles found within an utterance.

Semantic Roles: The MyST system models a lesson's key concepts as propositions which are realized as semantic frames. For MyST natural language understanding, these frames serve as the top-level nodes for a manually written semantic grammar used by the Phoenix parser (Ward, 1994). Two example concepts/frames and Phoenix parses are shown below. Although these semantic frames form the basis of MyST dialogues, for DISCUSS annotation we sought a more domain-independent representation that would generalize across a wide range of subjects. We began with VerbNet (Schuler, 2005) for defining our set of semantic roles because of its intuitive balance between descriptiveness and portability. While we used a majority of the labels as is, we found that the definition of some roles needed to be modified or extended to properly cover our set of concepts. For example, many concepts that express proportionality relationships can not be easily represented using predicate argument structure, and are more easily decomposed into *cause* and *effect* roles. We also added the catch-all *keyword* label to reflect terms that may relate to the proposition, but are not part of the core representation.

For our annotation project, rather than manually tagging all of the utterances with VerbNet labels, we created a mapping layer between the Phoenix frame roles and the VerbNet roles. The table below shows two frames along with their role mappings. We envision that in future projects, the hand-tuned semantic grammars could be replaced with a statistically trained semantic role labeler.

Frame:	BatteryFunction	Frame:	MagnetsAttract
Description:	*The DCell is the source of electricity.*	Description:	*Magnets attract to certain objects.*
⟨Instrument⟩:	[Battery]	⟨Instrument⟩:	[Magnet]
⟨Predicate⟩:	[Source]	⟨Predicate⟩:	[Attract]
⟨Theme⟩:	[Electricity]	⟨Theme⟩:	[Object]

Predicate Type: Simply knowing an utterance's propositional content is insufficient for inferring what was stated. Consider the two exchanges shown in the table below. The mixture of semantic roles in both students' responses are identical. Additionally, we can not differentiate between the exchanges based solely on dialogue act or rhetorical form. We need additional information to know the first scenario seeks to elicit discussion about observations while the second scenario focuses on procedures. One can also imagine such information would be useful for identifying communication breakdowns. For example, responding with a description of a procedure to a request about a process may indicate that the student did not understand the question or that the student is unwilling or unable to address the question.

T12:	*Tell me about what's going on here in this picture.* Ask/Describe/**Observation**
S13:	*The wires connect the battery and the light bulb and then then light bulb lights up.* Answer/Describe/**Observation** ⟨Instrument⟩.*wires* ⟨Predicate⟩.*connect* ⟨Theme1⟩.*battery* ⟨Theme2⟩.*light bulb* ⟨Effect⟩.*bulb lights up*
T7:	*Tell me about how you got the bulb to light up.* Ask/Describe/**Procedure**
S8:	*To make the light go we connected the wires to the battery and the bulb.* Answer/Describe/**Procedure** ⟨Effect⟩.*light go* ⟨Predicate⟩.*connected* ⟨Instrument⟩.*wires* ⟨Theme1⟩.*battery* ⟨Theme2⟩.*bulb*

To address this need, we created the *Predicate Type* based partly on the rhetorical predicates used in the DISCOUNT (Pilkington, 1999) scheme. While DISCOUNT included discourse relations in the set of predicate types, we restrict predicate types to those that encapsulate or summarize the collection of semantic roles in an utterance. Example predicate types include *procedure*, *observation* and *purpose*. A complete list of predicate types can be found in our forthcoming technical report.

3.3 Pragmatic Dimensions

The pragmatic dimensions are composed of the dialogue act dimension and the rhetorical form dimension. The dialogue act expresses the communicative function of a move and is the most general dimension in DISCUSS. The rhetorical form expresses attributes of the utterance's surface realization and can be thought of as refining the intent of the coarser dialogue act.

Dialogue Act: The dialogue act dimension is the top-level dimension in DISCUSS with the values of all other dimensions depending on the value of this dimension. Like with the majority of dialogue act taxonomies, DISCUSS dialogue acts have a grounding in speech act theory with a focus on what action the utterance performs. While most of the dialogue acts in the Dialogue Control and Information Exchange move categories have direct corollaries to those found in other taxonomies like DIT++ or DAMSL, we needed to supplement them with two frequently used Questioning the Author discussion moves: *marking* and *revoicing*. In marking, the tutor highlights parts of the student's language to emphasize important points and to steer the conversation towards key concepts. Revoicing serves a similar purpose, but instead of highlighting, the tutor rephrases student speech to clarify ideas they may have been struggling with. Examples of these acts are shown below.

S5: *that when you stick a magnet to a rusty nail and then you stick it to a paper clip it sticks* Answer/Describe/Process T6: *I think I heard you say something about magnets sticking or attracting. Tell me more about that.* **Mark**/None/None, Ask/Elaborate/Process
S33: *well when you scrub the the paperclip to the magnet the paperclip is starting to be a magnet* Answer/Describe/Process T34: *very good, so if the magnet gets close to the paperclip it picks it up* Feedback/Positive/None, **Revoice**/None/None

Dialogue acts in the Attention Management move category also reflect many of the actions regularly seen in tutorial dialogue. *Focus* and *Defer* acts are often used to move to or away from lesson-specific topics. In our corpus *Direct* is typically used to give instructions related to the multimedia (e.g. "Click on the box" or "Look at this animation.").

Rhetorical Form: The DISCUSS *Rhetorical Form* dimension provides another mechanism for differentiating between utterances with identical semantic content. While the dialogue act dimension is useful for providing an utterance's pragmatic interpretation and for determining what sequences are licensed, by itself it provides no indication of how a speaker is advancing the topic under discussion. Additional information is needed to create an utterance's surface form. Consider the two transactions in the table below. The semantic parses in both scenarios would be identical, however the tutor's questions and the resulting student response serve very different functions. In the first, the tutor is asking for a description and in the second, identification. Selection of the DISCUSS rhetorical forms found in the Information Exchange move category were inspired by the sixteen top-level tags used in Rhetorical Structure Theory (RST) (Mann and Thompson, 1988). While RST uses a rhetorical relation to link clauses and to show the development of an argument, DISCUSS uses the rhetorical form to refine the dialogue act. A sequence of dialogue acts paired with rhetorical forms can show progressions in the dialogue and tutoring process such as a shift from open-ended to directed questioning.

T1: *Can you tell which one is the battery?* Ask/**Describe**/Visual S2: *The battery is putting out electricity.* Answer/**Describe**/Process	T1: *Can you describe what is going on with the battery?* Ask/**Identify**/None S2: *The battery is the one putting out the electricity.* Answer/**Identify**/None

4 Annotation Status

We are still in the early stages of this ambitious annotation project. We currently have approximately 60 transcripts singly-annotated with DISCUSS moves. Each of these transcripts represents roughly 15 minutes of conversation and 50 turns on average. The DISCUSS taxonomy is a work in progress. Though

we have created the tags for each dimension based on a wide body of prior research and on preliminary studies of our transcripts, we expect that future analysis of our annotation reliability and consistency will likely lead us to add, modify, and combine tags. We anticipate that DISCUSS's multidimensional nature will likely raise issues for inter-annotator reliability, and the ability to add multiple tags per turn will further complicate the process of evaluating agreement.

5 Future Work and Conclusions

We plan to use our corpus of DISCUSS annotated tutorial dialogues to build dialogue models for a variety of applications including assessment of tutorial quality and dialogue move prediction. This annotation will allow us to investigate what features of tutorial dialogue correlate with increased learning gains and what types of questions encourage greater student interaction. Data-driven dialogue characterization will also allow us to explore how tutorial tactics vary across domains and tutors. We envision this work as an important first step towards automatic question generation.

In this paper we introduced the DISCUSS dialogue move taxonomy. This scheme overlays dialogue act and rhetorical annotation over semantic representations. We believe this combination of pragmatic interpretations and semantic representations provide an intermediate representation rich enough to analyze the interactions in a complex task-oriented domain like tutorial dialogue. Furthermore, we think DISCUSS moves can succinctly summarize the actions of a speaker's turn, while still providing sufficient information for natural language generation of dialogue moves.

Acknowledgments This work was supported by grants from the NSF (DRL-0733322, DRL-0733323) and the IES (R3053070434). Any findings, recommendations, or conclusions are those of the author and do not necessarily represent the views of NSF or IES.

References

Beck, I. L., M. G. McKeown, J. Worthy, C. A. Sandora, and L. Kucan (1996). Questioning the author: A year-long classroom implementation to engage students with text. *The Elementary School Journal 96*(4), 387–416.

Buckley, M. and M. Wolska (2008). A classification of dialogue actions in tutorial dialogue. In *Proc. COLING*, pp. 73–80. ACL.

Bunt, H. (2009). The dit++ taxonomy for functional dialogue markup. In *Proc. EDAML 2009*.

Core, M. and J. Allen (1997). Coding dialogs with the damsl annotation scheme. In *AAAI Fall Symposium on Comm. Action in Humans and Machines*, pp. 28–35.

Graesser, A., X. Hu, S. Susarla, D. Harter, N. Person, M. Louwerse, B. Olde, and the Tutoring Research Group (2001). Autotutor: An intelligent tutor and conversational tutoring scaffold. In *Proc. AIED'01*, pp. 47–49.

Graesser, A. and N. Person (1994). Question asking during tutoring. *American Educational Research Journal 31*, 104–137.

Mann, W. C. and S. A. Thompson (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text 8*(3), 243–281.

Nielsen, R. D., J. Buckingham, G. Knoll, B. Marsh, and L. Palen (2008, September). A taxonomy of questions for question generation. In *Proc. WS on the Question Generation STEC*.

Pilkington, R. M. (1999). Analysing educational discourse: The discount scheme. Technical Report 99/2, Computer Based Learning Unit, University of Leeds.

Schuler, K. K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon.* Ph. D. thesis, University of Pennsylvania.

Tsovaltzi, D. and E. Karagjosova (2004). A view on dialogue move taxonomies for tutorial dialogues. In *Proc. SIGDial*, pp. 35–38. ACL.

Ward, W. (1994). Extracting information from spontaneous speech. In *Proc. ICSLP*.

Ward, W., R. Cole, D. Bolanos, C. Buchenroth-Martin, E. Svirsky, S. Van Vuuren, T. Weston, J. Zheng, and L. Becker (2010). My science tutor: A conversational multi-media virtual tutor for elementary school science. *ACM TSLP: Special Issue on Speech and Language Processing of Children's Speech for Child-machine Interaction Applications*.