

WOULD I LIE TO YOU? MODELLING MISREPRESENTATION AND CONTEXT IN DIALOGUE

Carl Gutwin
Alberta Research Council¹
6815 8th Street N. E.
Calgary, Alberta T2E 7H7, Canada
Internet: gutwin@skyler.arc.ab.ca

Gordon McCalla
ARIES Laboratory, University of Saskatchewan²
Saskatoon, Saskatchewan S7N 0W0, Canada

ABSTRACT

In this paper we discuss a mechanism for modifying context in a tutorial dialogue. The context mechanism imposes a pedagogically motivated misrepresentation (PMM) on a dialogue to achieve instructional goals. In the paper, we outline several types of PMMs and detail a particular PMM in a sample dialogue situation. While the notion of PMMs are specifically oriented towards tutorial dialogue, misrepresentation has interesting implications for context in dialogue situations generally, and also suggests that Grice's maxim of quality needs to be modified.

1. INTRODUCTION

Most of the time, truth is a wonderful thing. However, this research studies situations where not saying what you believe to be the truth can be the best course of action. Intentional misrepresentation of a speaker's knowledge appears to be a common and highly pragmatic process used in many different kinds of dialogue, especially tutorial dialogue.

We use imperfect or incomplete representations in response to constraints and demands imposed by the situation: for example, many models of the real world are extremely complex, and misrepresentations are often used as useful, comprehensible approximations of complicated systems. People use idealized Newtonian mechanics, the wave (or particle) theory of light, and rules of default reasoning stating that birds fly, penguins are birds, and penguins don't fly. Some systems which cannot be simplified are purposefully ignored: for example, higher order

¹ This research was completed while C. Gutwin was a graduate student at the University of Saskatchewan. All correspondence should be sent to the first author.

² Visiting scientist, Learning Research & Development Centre, University of Pittsburgh, 1991-92

differential equations are left out of engineering classes because of their complexity. Simplified and imperfect representations are often found in tutoring discourse.

Misrepresentation as a pedagogic strategy holds promise for extending the capabilities of intelligent tutoring systems (ITSs), but the concept also affects computational dialogue research: it builds on the idea of discourse focus and context, extends work on adapting to the user with multiple representations of knowledge, and challenges Grice's maxims of conversation.

2. MOTIVATION AND BACKGROUND

Misrepresentations are alterations to a perceived reality. When they have sincere pedagogic purposes, we name them *Pedagogically Motivated Misrepresentations*, or PMMs. PMMs can reduce the complexity of the dialogue and of the concepts to be learned, provide focus in a busy environment, or facilitate the communication of essential knowledge.

PMMs share themes with research into computational dialogue and ITS. PMMs are intimately connected to ideas of instructional and dialogue focus, the latter of which was explored by Grosz [1977], who stated that task-oriented dialogue could be organized into focus spaces, each containing a subset of the dialogue's purposes and entities. The collection of focus spaces created by the changing dynamics of a dialogue could be gathered together into a focusing structure which assisted in interpreting new utterances.

Adaptation to the hearer is also a concern in dialogue research: beliefs about the hearer or about the situation can be used to vary the structure, complexity, and language of discourse to optimally suit the hearer. Several projects (e.g. [McKeown et al 1985], [Moore & Swartout 1989], [Paris 1989]) have

looked at adapting the level or tenor of explanations to a user's needs. Paris's [1989] TAILOR system varies its output (descriptions of complex devices) depending upon the hearer's expertise.

Another concern in both dialogue research and ITS research is multiple representations of domain knowledge. TAILOR, for example, uses two different models of each device to construct its explanations. Tutoring systems like SMITHTOWN [Shute and Bonar 1986] and MHO [Lesgold et al. 1987] organize different representations around distinct pedagogic goals; in the domain of electrical circuits, QUEST [Frederiksen & White 1988] provides progressively more sophisticated representations, from a simple qualitative model to quantitative circuit theory.

Lastly, any discussion of misrepresentation in dialogue is bound to reflect on Grice's first maxim of quality, "do not say that which you believe to be false." The conversational maxims of H. Paul Grice [1977] are a well-known set of observations about human discourse frequently used in computational dialogue research (for example [Joshi et al 1984], [Moore and Paris 1989], [Reichman 1985]). However, people sometimes accept the truth of Grice's maxims too easily. A close examination reveals difficulties with a literal interpretation of the first maxim of quality. While this maxim seems a reasonable rule to use in dialogue, examination of human discourse shows many instances where uttering falsehoods is legitimate behaviour. For example, in some first year computer science courses, students are told that a semicolon is the terminator of a Pascal statement. This utterance misrepresents reality (a semicolon actually separates statements), but the underlying purpose is sincere: the misrepresentation allows students to begin programming without forcing them to learn about syntax charts, parsing algorithms, or recursive definitions.

Grice's maxims have avoided major criticism by the computational dialogue community, and the maxims have been successfully used in limited domains to help dialogue systems interact with their users. Realizing that misrepresentations often occur in tutorial discourse, however, provides us with a context for investigating limits to the Gricean approach.

3. OVERVIEW OF PEDAGOGICALLY MOTIVATED MISREPRESENTATIONS

We have identified and characterized several types of PMM that can occur in tutorial discourse. We define each type as a computational structure that, when invoked, alters the dialogue system's own

reality and hence the student's perception of reality, for sincere pedagogic purposes. There are five essential computational characteristics governing the use of PMMs: preconditions, applicability conditions, removal conditions, revelation conditions, and effects.

These conditions are predicates matched against information in the dialogue system's essential data structures: a domain knowledge representation (in this system, a granularity hierarchy after [Greer and McCalla 1989], as shown in Figure 1); a model of the student; and an instructional plan (in this system, a simplified version of Brecht's (1990) content planner, from which a sample partial plan is shown in Figure 2). Each step in the instructional plan provides a teaching operator (such as prepare-to-teach) and a concept from the knowledge base which becomes the focus of the instructional interaction.

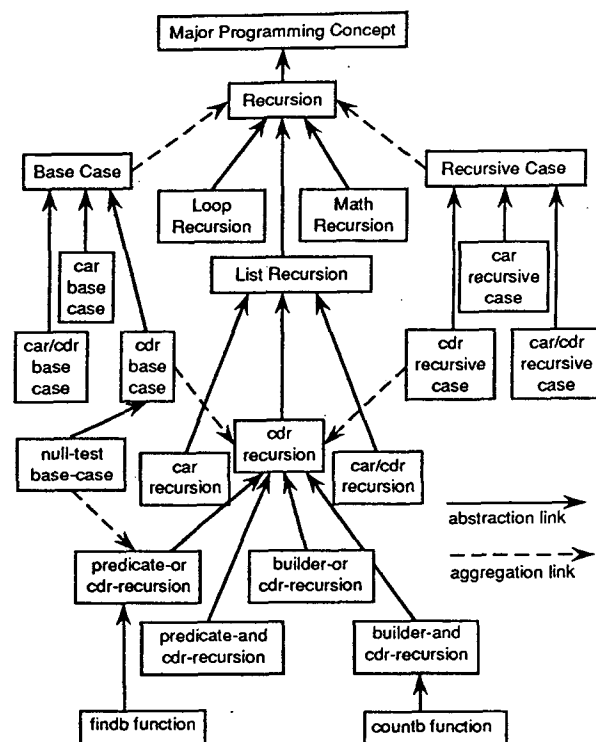


Figure 1. A fragment of the domain representation

In this implementation, PMMs act by manipulating the dialogue system's blackboard-based internal communication. An active PMM intercepts relevant messages before the knowledge base can receive them, then returns misrepresented information instead of the "true" information to the blackboard.

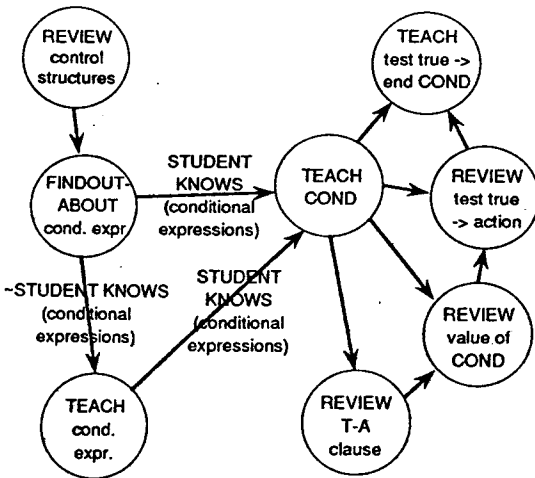


Figure 2. A partial content plan from Brecht's [1990] planner.

The first step in using a misrepresentation involves the PMM's preconditions and applicability conditions. Preconditions are definitional constraints characterizing situations in which a particular PMM is conceivable. Applicability conditions actually determine the suitability of a PMM to a situation. Each applicability condition examines one element of the current instructional context, from the student model, the domain representation, or the instructional plan. The individual conditions are combined to determine a final "score" for the PMM, using a calculus akin to MYCIN's certainty factors ([Shortliffe 1976]). For example, one applicability condition states that less student knowledge about a domain concept can provide evidence for the PMM's greater applicability, and more knowledge implies less applicability.

A PMM's removal conditions provide a facility for determining when the misrepresentation is no longer useful and may be removed. However, a dialogue system also needs to know when a PMM is not working well; after all, there are certain dangers associated with the use of misrepresentations. For example, a student may realize the discrepancy between the altered environment and reality. These situations are monitored by a PMM's revelation conditions, guiding the system in cases where it must be ready to abandon the misrepresentation and reveal the misrepresentation.

If preconditions and applicability conditions are satisfied, a PMM's procedural effects can be applied to the domain representation, implementing the

'alternative reality' presented to the student through the dialogue.

The way in which the student's perceived environment is altered and restored plays a crucial part in a misrepresentation's success. The dialogue actions which accomplish these changes compose two unique subdialogues. An alteration subdialogue must make a smooth transition to the altered environment; a restoration subdialogue has the opposite effect: it must restore the "real" environment, knot all the loose ends created by the misrepresentation, and help the student transfer knowledge from the misrepresented environment to the real environment. Restoration subdialogues must guard against another potential danger of misrepresentation: that students may retain incorrect information even after the misrepresentation has been retracted at the close of the learning episode.

4. DETAILS OF THE PMM MODEL

We have identified several types of pedagogic misrepresentations, and have implemented and evaluated them in a partial tutorial dialogue system. The implemented system concentrates on the function of the misrepresentation expert, and therefore the dialogue system is not fully functional: for example, it does not process or generate surface natural language. We have implemented the misrepresentation expert and the PMM structures, the blackboard communication architecture, the student model, and the domain knowledge (see Figure 1). The content planner and other system components are implemented as shells able to provide necessary information when needed.

Input to the system is a teaching situation including information from the content planner, the student model, and the domain. The system's output is a log of system actions detailing the simulation of the teaching situation.

Figure 3 shows the organization of the implemented PMMs, some of which inherit shared conditions and effects. The implemented PMMs have a variety of uses: Ignore-Specializations PMM simplifies concepts by reducing the number of kinds that a concept has; Compress-Redirect PMM collapses a part of the granularity hierarchy to allow specific instantiations of general concepts. There are also extended versions of these two PMMs which have more wide-reaching effects. The remaining PMMs are Entrapment PMM, which uses a misconception to corner a student and add weight to

the illustration of a better conception, and Simplify-Explanation PMM, which reduces the complexity of a concept's functional explanation. The remaining restriction PMM, Restrict-Peripheral PMM, is detailed in the following section to illustrate the concept of misrepresentation and the elements of the PMM model, and to show the PMM's use in an actual dialogue.

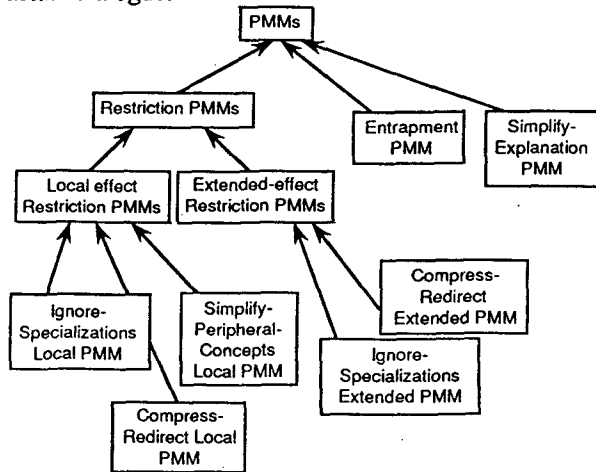


Figure 3. The PMM hierarchy.

The purpose of the "Restrict Peripheral Concepts" PMM is to simplify concepts related to the current teaching concept. For example, during an initial discussion of base cases (while learning programming in Lisp), a student might benefit from a misrepresentation which restricts recursive cases to a single type, the variety of recursive case used with cdr recursion. The restriction allows both participants in the dialogue to discuss and refer to a single common object, and allows the student to concentrate on base cases without needing to know the complexities of recursive cases.

This PMM's preconditions check that there are peripheral concepts in the current instructional context. Applicability conditions determine whether those concepts should be simplified, by considering the domain's pedagogic complexity and the student's capabilities. For example, the PMM considers the difficulty ratings of the current concept and the peripheral concept, the student's knowledge of these concepts and any existing difficulties with them as shown in the student model. In addition, the PMM considers other factors such as the student's anxiety level and their ability with structural relationships.

Removal conditions for this PMM consider factors such as whether or not instruction about the current

concept has been completed, or whether the instructional context has changed so markedly that the PMM can no longer be useful. Revelation conditions cover two other cases for a PMM's removal: when the student challenges the misrepresentation, and when the student or another part of the dialogue system requires a hidden part of the domain.

If applied, the effect of this PMM is to restrict peripheral concepts related to the current concept such that all but one of their specializations are hidden. The PMM carries out the restriction, but does not choose the specializations that will remain visible: that decision is left to the pedagogic expert, using the instructional plan and the student model.

5. EXAMPLE DIALOGUE

PMM "Restrict Peripheral Concepts" is illustrated below in an example dialogue. The dialogue is based on an actual trial of the implemented system, which determined when to invoke the PMM, when to revoke it, and all the interactions between the knowledge base and the dialogue system. However, the surface utterances are fabricated to illustrate how the misrepresentation system would function in a completed tutorial discourse system.

The teaching domain in the dialogue is recursion in Lisp (as shown in Figure 1), and the system believes the student to be a novice Lisp programmer.

T: ... the next thing I'd like to show you is the part of recursion that stops the reduction.

The system's current instructional context contains a teaching operator, "prepare to teach x," and a current concept, "base case." The current situation satisfies the preconditions of PMM "Restrict Peripheral Concepts," and its applicability score ranks it as most applicable to the situation. The PMM thus determines that the peripheral concept "recursive case" will be restricted to one specialization, and the pedagogic expert chooses 'cdr recursive case' as the most appropriate specialization for novice students.

The system asks the instructional planner to replan given the altered view of the domain, and enters into an alteration subdialogue with the student. Although these subdialogues are only represented as stubs in the system's internal notation, the discourse could proceed as follows:

T: Do you remember the last example you saw?
S: Yes.

T: OK. Remember that I pointed out the parts of the recursive function, the base case and the recursive case?

S: Yup.

T: Great. Now, I'll just put that example back on for a second. You'll notice that the recursive case looks like "(t (allnums (cdr liszt)))" Got that?

S: Yup.

T: Ok. For when we look at the base case, I want you to assume that this recursive case is the only kind of recursive case that there is. Then when we write some programs, you won't have to worry about the recursive case part. Does that sound ok?

[At this point the system has already imposed its alteration on the knowledge base, and when the system asks for the specializations of 'recursive case,' it will receive only 'cdr recursive case' as an answer.]

S: Sure.

T: Great. So the thing to remember is, whenever you need a recursive case, use a recursive case like you have in the example.

So. Let's move on to looking at the way the base case works; let's start with that example we had up. First, you identify the base case...

Later in the dialogue, the student is constructing a solution to another problem:

S: I'm not sure about the base case for this one ... I think I'll do the recursive case first. What does the recursive case do again?

T: A recursive case reduces the problem by calling the function again with reduced input. The recursive case is the default case of the "cond" statement, and it calls the function again with the cdr of the list input.

[Here the PMM again alters perceived reality, restricting 'recursive case' to 'cdr recursive case']

S: Right. So the recursive case is (t (findb (cdr liszt)))?

T: Yep.

[The PMM is again used to verify the student's query.]

S: OK. Now the base case ...

This exchange shows that the misrepresentation is useful in focusing the dialogue on the current concept of base case, by making the recursive case easy to synthesize.

The system continues investigating and teaching base case until the student can analyse and synthesize simple base cases. The instructional plan then raises its next step, "complete base case." Arrival at this plan step satisfies one of the removal conditions for the PMM, so the system engages in a restoration subdialogue with the student, which might go as follows, preparing the student for the next context:

T: Ok. The next thing we'll do is look a little closer at recursive case. Although I told you that there was only one kind of recursive case, there are actually more. The reason we only used one kind of recursive case is because I wanted to make sure you learned the way a base case works without needing all the details of recursive cases. Recursive cases still do the same thing (that is, reducing the input) but the specific parts might do different things than the recursive case we used. Does that sound ok to you?

S: ok.

T: So let's look at recursive cases. We'll only deal with the kinds used with cdr recursion. ...

6. RESULTS AND DISCUSSION

Evaluative trials for the PMM system have been aimed specifically at both the individual PMMs and the PMM model. Twenty-six different types of situations have been designed to test the PMMs' relevance, consistence, and coherence. Through these trials the individual PMMs demonstrated their integrity, and the PMM model itself was shown to be capable of working within a dialogue system architecture. Full details of evaluation methodology and results can be found in [Gutwin 1991].

This research project has shown that PMMs can be represented for use in a tutorial dialogue system, and supports their value as a pedagogic tool. However, the foremost contribution of the PMM system to computational dialogue may be how it extends the notion of focus currently used in dialogue research. Grosz and Sidner [1986] see dialogue as a collection of focus spaces which shift in reaction to changes in the discourse's purposes and salient entities. This research suggests that within any of these focus spaces, there can exist a further structure: a context that provides a specific interpretation of the knowledge represented in the system. The same knowledge is "in focus" throughout the focus space, but different contexts can color or interpret that knowledge in different ways. A pedagogically motivated misrepresentation is thus a context mechanism that alters the domain knowledge for an educational purpose. It is possible that we always use

some kind of alternate interpretation or misrepresentation to mediate between our knowledge and other dialogue participants.

Focusing structure has traditionally been used in interpretation: in several projects ([Grosz 1977], [Sidner 1983]), context structures are shown to be useful in tasks like pronoun resolution or anaphora resolution. Pragmatic contexts, such as those created by a PMM, can direct generation of discourse as well. They are active reflections of the larger situation, rather than local representations of dialogue structure, and they are able to alter the discourse in order to further some goal. Responding to patterns in the world outside the dialogue allows pragmatic context mechanisms such as PMMs to consider fitness and suitability of a dialogue situation in addition to a focus space's subset of goals and salient entities.

Another issue of importance to this research is that of tailoring. While some existing dialogue systems tailor an explanation to the user's level of expertise (e.g. [Paris 1989], [McKeown et al 1985]), the PMM system instead tailors the domain to the learner. The PMM system does not make basic decisions about either content or delivery in a dialogue, but attempts to shape the content's representation into a form which will be best suited to the learning situation.

The PMM model also touches on research into multiple representation, in that it provides a mechanism for encapsulating several different interpretations of a knowledge base. The mechanism might be able to model and administer alternate representations of other kinds as well, such as analogy.

The usefulness and ubiquity of PMMs also suggests that a literal interpretation of Grice's maxims, particularly the maxim of quality, is inappropriate. Clearly, we often say things we know to be false! However, the maxim of quality can be rescued by indicating the relationship between truth and dialogue purposes: from the original, "do not say that which you believe to be false," we create a new maxim, "do not say that which you believe to be false to your purposes." The new maxim shifts emphasis from an absolute standard of truth in dialogue to the more pragmatic idea of truth relative to a dialogue's goals, and better reflects the way humans actually use discourse.

Much remains to be accomplished in this research. There are undoubtedly other as yet undiscovered

PMMs. The notion of intentional misrepresentation itself may just be an instance of a more general context mechanism that underlies all dialogue, an idea that should be explored by considering other kinds of dialogue from the perspective of PMMs, and by a closer examination of existing theories of discourse context. Finally, all of the oracles used in the PMM system should be replaced by functioning components so that a dialogue system with complete capabilities can stand alone as proof of the PMM concept. Nevertheless, this research points the way towards the possibility of a new and widely applicable mechanism for modelling dialogue.

ACKNOWLEDGMENTS

The authors wish to thank the Natural Science and Engineering Research Council of Canada for financial assistance during this research.

REFERENCES

- [Brecht 1990] Brecht (Wasson), B. *Determining the Focus of Instruction: Content Planning for Intelligent Tutoring Systems*, Ph.D. thesis, University of Saskatchewan, 1990.
- [Frederiksen & White 1988] Frederiksen, J.R., and White, B. Intelligent Learning Environments for Science Education. in *Proceedings of the International Conference on Intelligent Tutoring Systems*, Montreal 1988, pp. 250-257.
- [Greer and McCalla 1989] Greer, J., and McCalla, G. "A computational framework for granularity and its application to educational diagnosis" in *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Detroit MI, 1989, pp. 477-482.
- [Grice 1977] Grice, H.P. "Logic and Conversation" in *Syntax and Semantics*, Vol. 3, New York: Academic Press, 1975, pp.41-58.
- [Grosz 1977] Grosz, B. "The Representation and Use of Focus in a System for Understanding Dialogs" in *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Cambridge, Massachusetts, 1977, pp. 67-76.
- [Grosz & Sidner 1986] Grosz, B.J., and Sidner, C. "Attention, Intentions, and the Structure of Discourse" in *Computational Linguistics* 12, 1986, pp.175-204.
- [Gutwin 1991] Gutwin, C. *How to Get Ahead by Lying: Using Pedagogically Motivated Misrepresentation in Tutorial Dialogue*. M.Sc. Thesis, University of Saskatchewan, 1991.
- [Joshi et al 1984] Joshi, A., Webber, B., and Weischedel, R. "Preventing False Inferences" in

- Proceedings of the 10th International Conference on Computational Linguistics*, 1984, pp.134-138.
- [Lesgold et al 1987] Lesgold, A., Bonar, J., Ivil, J., and Bowen, A. An intelligent tutoring system for electronics troubleshooting: DC-circuit understanding. in *Knowing and Learning: Issues for the Cognitive Psychology of Instruction*, L. Resnick ed., Hillsdale NJ: Lawrence Erlbaum Associates.
- [McKeown et al 1985] McKeown, K., Wish, M., Matthews, K. "Tailoring Explanations for the User" in *Proceedings on the 5th International Joint Conference on Artificial Intelligence*, Los Angeles, August 1985, pp.794-798.
- [Moore and Paris 1989] Moore, J., and Paris, C. "Planning Text for Advisory Dialogues" in *Proceeding of the 27th Conference of the Association for Computational Linguistics*, 1989, pp. 203-211.
- [Moore and Swartout 1989] Moore, J., and Swartout, W. R. "A reactive approach to explanation," in *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Detroit, 1989 pp.
- [Paris 1989] Paris, Cecile. "The use of explicit user models in a generation system for tailoring answers to the user's level of expertise" in *User Models in Dialog Systems*, A. Kobsa and W. Wahlster, eds. Berlin: Springer-Verlag, 1989, pp. 200-232.
- [Reichman 1985] Reichman, R. *Getting Computers to Talk Like You and Me*. Cambridge, MA: the MIT Press, 1985.
- [Shortliffe 1976] Shortliffe, E.H. *Computer-Based Medical Consultation: MYCIN*. New York: Elsevier.
- [Shute & Bonar 1986] Shute, V., and Bonar, J.G. "An intelligent tutoring system for scientific inquiry skills." in *Proceedings of the Eighth Cognitive Science Society Conference*, Amherst MA, pp.353-370.
- [Sidner 1983] Sidner, C. "Focusing in the Comprehension of Definite Anaphora" in *Computational Models of Discourse*, M. Brady and R. Berwick, eds. Cambridge, Mass: MIT Press, 1983, pp. 267-330.