Arguing with the Devil: Teaching in controversial domains

Symeon Retalis[1], Helen Pain[2] and Mandy Haggith[2]

[1] Department of Electrical Engineering, National Technical University of Athens, Athens, Greece
email: retal@softlab.ntua.gr
[2] Department of Artificial Intelligence, University of Edinburgh, 80 South Bridge, Edinburgh, EH1 1HN, UK
email: helen@aisb.ed.ac.uk, hag@aisb.ed.ac.uk

Abstract. There are certain problems inherent in teaching in controversial domains. In this paper, we present a prototype of an intelligent tutoring system called OLIA which addresses some of these problems. We focus here on OLIA's **Devil's advocate** strategy which is a method for teaching through a debate between the user and the system.
This paper takes advantage of existing work using a meta-level argumentation framework for knowledge representation in controversial domains. In this framework statements are linked by primitive relations such as disagreement, equivalence and justification, and higher order meta-level ones such as counter-arguments and corroborations. We investigate the use of this framework as the basis for the OLIA system.

1 Introduction

Teaching in controversial domains can be problematic. There is not one viewpoint to be taught as the correct one; teaching has to encourage students to explore different and often conflicting opinions, consider their justifications and weigh up the various arguments for and against a position. Tutoring systems, however, tend to focus on teaching a particular expert view. Cases of inconsistency or various beliefs are handled but usually considered to be evidence of misconceptions on the part of the student. However, in the real world absolute consensus or consistency is rare.

In this paper we present a prototype of an intelligent tutoring system called OLIA, which teaches in controversial domains. The OLIA tutoring system is based on a meta-level argumentation framework (Haggith 1995). It is the application of this framework to teaching in controversial domains that is addressed in this paper. This framework, unlike traditional knowledge based systems, supports controversy and is used for handling and exploring inconsistency. The OLIA tutor supports multiple teaching strategies (see Retalis, 1995), playing primarily the role of a Coach and also the role of a **Devil's Advocate**.

In the Devil's Advocate approach the user debates with the system on a controversial topic and both parties try to defend their opinions using supporting arguments. This teaching method is an innovative one and aims to provide the student with a global knowledge of the domain in order to avoid him jumping to a conclusion without being made aware of the various opinions in the domain.

2 Teaching in Controversial Domains

Conflicting opinions are not a problem but an opportunity to teach; an opportunity for interesting reasoning. For example people were interviewed in order to get their opinions about the greenhouse effect (Haggith 1993). Their viewpoints are conflicting.

Edinburgh will be submerged in a foot of water, claims an informant, as *It is near the sea* and *The sea level will rise* because:

1. The polar ice-caps are mostly ice.
2. The polar ice will melt.
3. When the ice becomes water then the level of the sea will rise.

On the other hand, other informants state: Mm... *Edinburgh won't be submerged in a foot of water* since *There wont be anywhere near the amount of sea level change people make out*, ... *The ice melting doesn't matter*. Why? Because:

1. The ice will melt.
2. The majority of ice-melt will be at the Arctic.
3. Arctic ice sits on the sea and displaces the water.
4. When the ice melts, all that volume is replaced by water.
5. Nine-tenths of the ice sits below the surface of the water, and its density is lower.

Although some of the above opinions of the scientists are equivalent (*The ice will melt* and *The polar ice will melt*) they are used to justify conflicting viewpoints which disagree. In a case where different opinions exist in the domain we need a tool for exploring the disagreements or the equivalences or for navigating around the domain by following the justifications of the propositions.

Different opinions on a topic often appear in different knowledge bases. Traditional knowledge based systems tend to present the knowledge of only one expert since they have difficulty in coping with the inconsistencies: their goal is usually to get rid of them. There is little work on the benefit that might be gained from exploring rather than removing such inconsistencies. However, Haggith (1995) proposes a meta-level argumentation framework for the exploration of multiple knowledge bases in which conflicting opinions are expressed. It is used as a basis for building the OLIA intelligent tutoring system for teaching in controversial domains. We focus here only on the Devil's Advocate strategy.

Systems such as DOMINIE (Spensley & Elsom-Cook 1988), or Meno-tutor (Woolf 1987) are not committed by design to a single tutoring approach but, as OLIA, support multiple teaching strategies. The difference between these systems and OLIA is that they deal with the viewpoints of only one expert and do not deal with controversies. Systems designed to support learning where several opinions are handled are rare, though there are examples such as Nicol's Intelligent Student System (Nicols 1993) and Belvedere (Suthers & Weiner 1995), and those where disagreement is permitted in relation to the student model (Bull et al, 1995). A number of others, which focus primarily on argumentation in relation to dialogue strategies, are discussed in (Treasure-Jones 1995). The success of systems such as Belvedere suggests that research in this area is of interest and benefit.

This paper introduces the Devil's Advocate strategy. In the previous example we had two scientists presenting their arguments. In OLIA one informant is the user and the other is the system who plays the role of the Devil's Advocate. A Devil's Advocate is a person who champions the weaker case for the sake of argument (Columbia 1950). The two parties debate by each presenting their arguments, and supportor justification for their respective arguments. The debate goes on until either the user has learned enough or the system gives up arguing (because of lack of arguments) saying "I rest my case!"

The **Devil's Advocate** term is commonly used in journalism, literature and law. In the Roman Catholic Church the *advocatus diaboli* was an official appointed to argue against a candidate (*advocatus Dei*) for a beatification or canonisation (Columbia 1950). In this paper we introduce the Devil's Advocate as a teaching approach in order to facilitate teaching in controversial domains into the ITS literature.

3 A knowledge representation framework for conflicting points of view

In Haggith (1995) a framework is described which enables multiple conflicting bodies of knowledge to be represented. This is achieved by restricting the knowledge representation to a high level of abstraction, in which the contents of statements are not formalised in any particular object level language, such as predicate logic; instead, what is formalised are the relationships between statements. The framework is thus meta-level - it enables reasoning *about* the statements in the knowledge base, and their relationships, rather than reasoning *with* them to draw particular conclusions. A basic assumption is that where controversies exist, it is useful to reason about the nature of the controversy, ie to argue. No attempt is made to resolve the conflict.

In this paper we explore how this framework makes possible the teaching of topics involving controversy. We assume that students are not trying to decide which of two conflicting points of view is 'correct', but rather learning about the nature of the controversy thus enabling them to weigh up the various sides of the argument. Understanding arguments, and how they relate, is necessary before a student is able to decide which argument is actually persuasive.

This meta-level framework treats statements or propositions as atoms, which do not contain internal structure which can be reasoned with. The primitive notions in the framework are four binary relations between propositions or sets of propositions.

1. *equivalent* - two propositions mean the same thing, or are in agreement; e.g. equivalent('The ice will melt.', 'The polar ice-caps will melt.').
2. *disagree* - two propositions state points of view which are in conflict; e.g. disagree('The level of the sea will rise', 'I don't think there'll be anywhere near the amount of sea-level change people make out.').
3. *justification* - a proposition is justified or supported by a set of other propositions; e.g. justification('I don't think there'll be anywhere near the amount of sea-level change people make out.', ['The ice melting does not affect sea-level', 'The ice will melt.']).
4. *elaboration* - a proposition is expanded upon by a set of other propositions; e.g. elaboration('The ice melting does not affect sea-level', ['The majority of ice-melt will be at the Arctic.', 'Arctic ice sits on the sea and displaces the water.', 'When the ice melts, all that volume is replaced by water.']).

The most important structures in the framework are arguments. An argument is a recursive structure with a conclusion and a collection of propositions which lend support to that conclusion. For example, if P is a conclusion supported by propositions Q, R and S, this would be written :

P <= [Q, R, S].

Each support can in turn be supported, so if R is itself supported, we get the argument in which R's argument is nested within the argument for P :

P <= [Q, R <= [U,V], S].

The argument connective <= represents that what is on its left hand side is supported by what comes on its right hand side, and the framework allows for various interpretations of the idea of support. For example, if we allow support to include equivalences, justifications and elaborations, we can construct a 'hybrid argument' for the statement above that the informant does not think that there will be much sea-level rise. This would be as follows :

> 'I don't think there'll be anywhere near the amount of sea-level change people make out.' <= ['The ice melting does not affect sea-level' <= ['The majority of ice-melt will be at the Arctic.', 'Arctic ice sits on the sea and displaces the water.', 'When the ice melts, all that volume is replaced by water.'] 'The ice will melt.']

Substituting single letters for the propositions we can see this structure more clearly :

P <= [Q <= [R, S,T], U]

P is supported by both Q and U. Q in turn is supported by all three of R, S, and T. Using this nesting structure, arbitrarily complex arguments can be constructed. The loss of detailed representational expressiveness of the contents of the propositions is compensated for by being able to represent very rich structures of connections between propositions.

The framework also provides definitions of more complex structures, 'higher order objects' which express relationships between arguments. These structures commonly occur in debates and in the literature on argumentation, and some of them exploit the existence of disagreements within the knowledge base. They include definitions of rebuttal (when an argument has a conclusion which disagrees with the conclusion of another argument), corroboration (two arguments whose conclusions are equivalent or support each other), enlargement (two arguments one of which argues for an elaboration of the conclusion of the other) and counterargument (when an argument's conclusion disagrees with any statement in another argument). Details of these structures and the rest of the framework can be found in Haggith 1995.

4 Devil's advocate strategy

We introduce the Devil's Advocate strategy as a method for tutoring through debate. The system will play the role of the Devil's Advocate and it will try to oppose any proposition that the user presents as support to his viewpoint. The user has selects one of the propositions from the *top level disagreements* in order to defend it. The Devil's Advocate will oppose this opinion and will support the opposite one. The possible options that are offered to the user are: *Support*, *Users higher order objects* and *Devils higher order objects*.

With the **Support** choice the user presents the *supports* for the propositions that he defends. However, the Devil's Advocate presents its supporting propositions in order to defend its part. Each interaction is a new level of debate. The new propositions that appear belong to this *New Level*. All the levels of debate are shown to the user (i.e. propositions that both parties – user and Devil's Advocate – have presented) in order that he be informed of how the discussion is "going". If the user cannot present any other support propositions to defend his opinion because all the propositions of the last level are self evident or assumptions then he is asked to stop interacting or to use only the options for higher order arguments. On the other hand if the Devil's Advocate cannot oppose anything he "rests his case" but the user can continue supporting his opinion since this is a way to learn more about the domain.

Two choices *Users higher order objects* and *Devils higher order objects* present information about the way the propositions that one party defends relate via higher order meta-level relations (rebuttals, undercuttings, counterarguments, etc.) to the ones of the "opponent" party. The following example is an instance of a debate between the user and OLIA as the Devil's Advocate. OLIA allocates one window to the user when they are debating with the system and another similar window presents the opinions of the Devil.

User: Edinburgh won't be submerged in a foot of water
DEVIL: Edinburgh will be submerged in a foot of water

User chooses to support his opinion

User: Edinburgh won't be submerged in a foot of water because:
- It is built on hills
- There won't be anywhere near the amount of sea level change people make out

DEVIL: No, no, no! Edinburgh will be submerged in a foot of water because:
- It is near the sea
- The sea level will rise

The user is now called to try again to defend his opinions or try to learn how his opinions and those of his opponent are related to each other via higher order relations
User selects support

User: Edinburgh is built on hills is self evident.
There won't be anywhere near the amount of sea level change people make out because:
- The ice will melt
- The ice melting does not affect sea level

DEVIL: Edinburgh is near the sea is self evident. The sea level will rise because:
- The polar ice-caps will melt.
- The polar ice-caps are mostly ice.
- When the ice becomes water then the level of the sea will rise.

The debate ends when the Devil's Advocate *rests his case* because of lack of arguments.

The Devil's Advocate allows the user to learn about the domain while debating with the system on a subject. However, there are situations in a discussion about a theme where different parties use the same arguments in a different way in order to defend their opinion. The Devil's Advocate strategy that is implemented in the OLIA tutoring system gives the user the option to learn more about the agreements between the "opponents" while they debate. It checks the possibility of equivalent points of view that the different parties have used and it shows some pieces of information such as which proposition an argument supports, with what other arguments it supports that proposition and so on. For example, the equivalent opinions *The ice will melt* and *The polar ice-caps will melt* are being used with different viewpoints to support conflicting opinions.

5 Evaluation and Discussion

Three initial studies were undertaken in order to carry out preliminary evaluation of the OLIA tutoring system. The focus here is primarily on the feasibility of the approach, and does not assess the effect on students' learning. In the long term, we would hope to use such an approach for improving both student's understanding of the arguments in specific domains, and their general understanding of arguments per se.

The first study: involved six postgraduate students, taking an MSc Knowledge Based Systems course. They used a small knowledge base concerning the greenhouse effect and arguments for and against the opinion that Edinburgh will be submerged in a foot of water. This involved a short briefing on the domain representation and the philosophy of the OLIA tutoring system, and a period of unguided exploration of the domain, exploring OLIA's Coaching strategy (see Retalis, 1995, for details) in addition to the Devil's Advocate strategy. Their interaction with the system was observed. This was followed by an informal interview, including questions on the friendliness, usability and limitations of the system.

In general, the results of the evaluation were encouraging. All six users agreed that OLIA was potentially useful for learning about controversial domains. The way in which they interacted with the system varied. They chose different routes (options) for learning the domain. All of them selected the Devil's Advocate strategy as the most interesting strategy. OLIA appeared to be of use to users who already have some idea of the domain or who expect more sophisticated information about the domain. In general, they found that OLIA was user friendly and almost all information was clearly presented.

The users commented that the Devil's Advocate strategy was interesting and potentially useful for tutoring about domains where conflicting opinions exist. They agreed that this strategy was necessary to prevent them from jumping to a conclusion without having general knowledge about the subject that they were exploring. The users also indicated that the Devil's Advocate was useful for teaching how the different opinions are related to each other when it was used with the option of *showing the agreements.*

The second study: involved 10 undergraduate students taking a first year course in Artificial Intelligence. In this study, an alternative domain was implemented: two sets of propositions represented alternative positions in relation to the mind-body problem, the Materialist position (Minsky) and the Dualist position (Descartes). Students were asked to interact with OLIA using both the Devil's Advocate and Coaching strategies, in order to try to answer questions about the mind-body debate. This study took place during a routine practical session, and provided a follow up to material that had been discussed in a previous tutorial.

Students were asked to comment, in writing, after the session. Example questions were, "Did you learn anything about the mind-body problem in this session? Did you learn anything about argumentation structure?" They were also asked to compare working with this material in tutorials to working with it in the OLIA system, and to suggest improvements to the system. We do not go into a detailed analysis of their responses here, but focus on general formative comments on the system that may influence design modifications and a provide a preliminary assessement of its potential use.

The general findings were that all students who expressed a preference preferred using the Devil's Advocate to the coaching strategy. The majority felt that they improved their understanding of the mind-body problem in the course of using OLIA. 5 students felt that using OLIA also improved their knowledge of argument structure. Some students said they preferred using OLIA to learning in a tutorial, because it was 'more systematic', and the arguments 'were presented better', Most students who expressed a preference preferred working in tutorials because the arguments could be 'deeper', they could ask their tutor for clarification, and they could express their own 'personal arguments and opinions'. A number of improvements were suggested, both to the interface and to the way in which the system might be used. Most felt that OLIA's Devils' Advocate strategy provided a useful means of exploring arguments, particularly when presented in relation to other means of presenting these arguments, such as reading texts.

The third study: used the same knowledge base (the mind-body problem), but here OLIA was used as an evaluation exercise for a group of seven postgraduate and honours level students taking a specialist course in Artificial Intelligence and Education. The structure and questions were similar to those of the second study, though there was no comparison with coverage of the material by other means (such as tutorials).

The results indicated that several students learned something about the mind-body problem in the session, though others felt not (though several commented that they were already familiar with the arguments). Most were able to answer a fairly complex question, after working with the Devil's Advocate, that they were not able to answer before using it: "Did you find any proposition that supports two other propositions which disagree with each other?" Some commented that they learned something about argument structure, though felt that the distinctions between some of the relations were not made clear by the system. Most preferred the Devil's Advocate strategy to the basic Coaching strategy. All made useful suggestions for improving the interface.

6 Limitations and Conclusions

Evaluation of the use of the Devil's Advocate strategy in the OLIA system is only in preliminary stages. There are a number of limitations to the approach, some of which will be considered briefly here. The Devil currently always uses the same type of argument move as the user. More flexiblility should be provided to enable a wider variety of interaction. The performance of this strategy has not been tested when more than two opinions appear as *top level disagreements*. If there were more than two conflicting opinions on a topic then the user could choose one and the Devil's Advocate could defend each one of the others. The current algorithm for this strategy should support this option. Nonetheless, we need to investigate ways which the user might debate with two **Devil's advocates** who in their turn might debate with each other. The graphical interface (built in Prolog) as it is at the moment does not enable the implementation of more than one Devil's Advocate. Additionally, there are no facilities at present for the student to add their own propositions: this would be a vital extension for more widespread use.

We have demonstrated the use of the argumentation framework for teaching in controversial domains. We believe that the meta-level argumentation framework could facilitate the use of other strategies such as the Socratic one. This would be further evidence of how powerful this framework is for tutoring. Both the meta-level framework and the OLIA intelligent system are designed to be domain independent (Haggith 1995): it is straightforward to use OLIA with other knowledge bases representing knowledge about other controversial domains, as we have demonstrated through our second and third studies. Further analysis of the data collected is currently being carried out, and improvements to the design of the tutor being considered. Evaluation of the effectiveness of this approach in respect of improving students' learning is also needed. Further study of student's use and understanding of the argument structures and relations is also required.

References

S. Bull, H. Pain and P. Brna. (1995). Mr. Collins: A A Collaboratively Constructed, Inspectable Student Model for Intelligent Computer Assisted Language Learning. *Instructional Science*, 23:65–87.

The Columbia Encyclopedia. Columbia University Press, 1950.

M. Haggith. Towards a formal analysis of dissagreement. Discussion Paper 128, 1993.

M. Haggith. A meta-level framework for exploring conflicting knowledge bases. In J.Hallam, editor, *Hybrid Solution, Hybrid problem.* IOS Press, 1995.

M. Minsky. *The Society of mind.* Simon and Schuster, New York, 1988.

D. Nicols. *Intelligent Student Systems: An Application of Viewpoints to Intelligent Learning Environments.* Unpublished PhD thesis, University of Lancaster, 1993.

S. Retalis. Investigating the use of a meta-level argumentation framework as a basis for tutoring in controversial domains. Unpublished M.Sc. thesis, Edinburgh University, 1995.

F. Spensley and M.T. Elsom-Cook. Dominie: Teaching and assessment strategies. Technical Report 37, Milton Keynes, UK: Centre for Information Technology in Education, The Open University, 1988.

D. Suthers and A. Weiner. Belvedere: Engaging students in critical discussion of science and public policy issues. In *Proc. 7th World Conference on Artificial Intelligence in Education (AI–ED 95).* Washington, D.C, 1995.

T. Treasure-Jones. Support and exploration of disagreement within intelligent tutoring systems. First Year PhD Report. Computing Department Lancaster University, 1995.

B. Woolf. Theoretical frontiers in building a machine tutor. In Kearsley G., editor, *Artificial Intelligence and Instruction. Applications and Methods*, pages 229–265. Addison-Wesley, 1987.

Acknowledgements
We would like to thank Amy Isard for her comments on this paper.

This article was processed using the LaTeX macro package with LLNCS style