A Human-Computer Dialogue System for Educational Debate: A Computational Dialectics Approach

Tangming Yuan, *Computer Science Department, University of Akureyri, Sólborg, 600 Akureyri, Iceland*
yuan@unak.is

David Moore, **Alec Grierson,** *Innovation North, Leeds Metropolitan University, Beckett Park, Leeds LS6 3QS, UK*
d.moore@leedsmet.ac.uk, a.j.grierson@leedsmet.ac.uk

Abstract. This paper reports research concerning issues involved in adopting a computational dialectics approach to develop a human-computer dialogue system for educational debate. In particular, we propose a dialogue model and a set of computationally usable strategic heuristics to enable the computer to engage its users in debate on a controversial issue. The system operationalising the proposed dialogue model and strategy has been fully implemented. It has also been subject to initial evaluation exercises. The results of the evaluation are essentially favourable. It is anticipated that the work reported in this paper will contribute toward the development of human-computer dialogue in general and of computer-based educational debate in particular, and help to illuminate research issues in the field of dialectics itself.

Keywords. Computational dialectics, educational human-computer debate, dialogue model, computational strategic heuristics

INTRODUCTION

The recent development of computer based learning (CBL) systems and the emergence of the World Wide Web and the Internet has changed the study life of many people (e.g. Beck & Stern, 1999; Cumming & McDougall, 2000; Cunningham-Atkins, Powell, Moore, Hobbs, & Sharpe, 2004; Grasso & Leng, 2003). However, an assumption underlying many of these computer based educational systems is that the computer does all the informing, the student being merely a passive receiver of the information. The teaching interaction, that is, may become unduly didactic (Moore, 2000; Baker, 2000). There is therefore a need for more dialogue within computer based learning systems.

There are many types of dialogue interactions in which people reason together, such as debate, persuasion, inquiry and information-seeking (Walton & Krabbe, 1995). The debating style of dialogue interaction is argued by Maudet and Moore (2001) to be important in critical thinking and developing debating and reasoning skills, and is also suggested by Pilkington and Mallen's (1996) educational discourse analysis to be effective and to have rich educational benefit. A particular concern with our research therefore is to investigate issues surrounding a computer based system for educational debate.

The educational benefit of such a debating system seems clear, in that the importance of discussion and debate in education is frequently stressed at primary (e.g. National Curriculum Council, 1990a), secondary (e.g. National Curriculum Council, 1990b), and tertiary (e.g. Garrison, 1991) levels.

The argument is given extra force by Laurillard's (Laurillard, 1995) mention of an "emphasis on whole-class teaching.... at the expense of opportunities for discussion and interaction", and her claim that "reflection is too often neglected in the teaching-learning process". The scope for application is wide, we argue, for as Self points out: "... it is rarely possible to define a unique 'correct' viewpoint to be communicated to a student" (Self, 1992, cf. Bouwer, 1999).

To enable a debate with a computerised system on a controversial issue such as capital punishment, the student will need to appreciate the need to defend his point of view against attacks by giving convincing reasons. The symmetrical nature of such a debate also enables the student to build his own position and to attack the opponent's view. This may encourage the student to think critically during the debate. In addition, by using a formal dialectical system (discussed below) to regulate such a debate, the students can learn the rules of debate, and improve their debating skills. A further educational advantage of using a formal dialectical system is that the student can learn the logical structure of the argument, and improve his reasoning skills. In addition, by revealing the issues under discussion, the student is given an opportunity to become familiar with the issues in the domain, and to perceive issues from different angles. Similar arguments for the advantage of such an educational debate are also made in (Pilkington & Mallen, 1996; Moore, 1993, 2000; Moore & Hobbs, 1996; Maudet & Moore, 2001; Ravenscroft & Pilkington, 2000; Pilkington, 1998, 1999; Hatcher, 1999; Quinn, 1997; Garrison, 1991).

This paper outlines our work in building such a debating system. The remainder of this paper is structured as follows. Firstly, we provide a brief introduction to the computational dialectics approach we have adopted in dealing with dialogue. Secondly, we discuss the dialogue model "DE", which we developed as the dialogue model for our human-computer debate system. We then discuss dialogue strategies for the computer to adopt to enable it to function as a dialogue participant. Finally, we discuss our evaluation of the debating system and outline our plans for further work concerning its development.

COMPUTATIONAL DIALECTICS

Dialectics involves a logical account of interaction in terms of rules for particular kinds of interaction. Computational dialectics is the use of such accounts as the dialogue model for human-computer or agent-agent dialogue systems. A common approach to dialectics is to utilise "Dialogue Game Theory" models developed within the field of Informal Logic to prescribe how dialogue should be regulated. A dialogue game (dialectical system) is seen as a rule governed structure for organised conversation where there is an exchange of arguments between two parties reasoning together on a turn-taking basis aimed at a collective goal (Walton, 1998). An important component of a dialogue game is the commitment store, which records what a dialogue participant has said or accepted during the evolving dialogue (Hamblin, 1971). A dialogue game also contains a set of available move types, a set of dialogue rules and a set of commitment rules. Dialogue rules prevent illegal events during the evolving dialogue, and commitment rules specify the effect of different move types on the participants' commitment stores. Some dialogue games also contain winning and losing rules (Walton & Krabbe, 1995).

Dialogue games, as so characterised, have a number of attractions from the point of view of their utilisation within a computer dialogue system. Walton (1998) argues that the use of dialogue games can show people how to prevent fallacious arguments and common errors. Moore and Hobbs (1996) argue that using dialogue games, if the game system is valid, can yield "fair and reasonable" dialogue. Further, Maudet and Moore (2001) argue that by constraining dialogue to a specific set of moves, each defined in terms of its effect on the commitment stores, a computer dialogue system based on dialogue games is able to bypass the need for complex pragmatic parsing, while still providing opportunity for educationally useful interaction.

Further, dialogue games use commitments rather than beliefs. Commitments are what a dialogue participant has said or accepted during the evolving dialogue. Commitments are public and thus it is easy to know what someone's commitments are in a dialogue (Hamblin, 1971; Mackenzie & Staines, 1999; Walton & Krabbe, 1995; Maudet & Chaib-Draa, 2002; Moore & Hobbs, 1996). Commitments are not necessarily the same as the dialogue participant's deeply-seated beliefs. An honest dialogue participant speaks his beliefs, but a good debater may make claims he does not believe, for the sake of argumentation. It is therefore difficult to know one's interlocutor's beliefs during argumentation. Dialogue games, by using commitment rather than beliefs, avoid this difficulty and allow debate participants to learn without exposure of their beliefs in a controversial domain, but with their domain knowledge and debating skills potentially being enhanced. In addition, the separation of commitment from belief has the advantage that one can allow for machines to argue without having to concede that they have beliefs (Moore & Hobbs, 1996).

Further support for the adoption of dialogue games as the basis for computerised dialogue systems can be found in the context of computer based learning. For example, Maudet and Moore (2001) use Mackenzie's (1979) DC to enable potentially educationally valuable "devil's advocate" discussions. Bench-Capon, Geldard and Leng (2000) investigate the use of dialectical systems in a legal educational context, and develop a dialogue game, "TDG", for this purpose. Ravenscroft and Pilkington (2000) develop a dialogue game framework to enable the computer to act as a "facilitating tutor" and the student the "explainer". The evaluation of their computerised system demonstrates its effectiveness in stimulating improvements in students' understanding of the physics of motion (Ravenscroft & Matheson, 2001, 2002).

Further, there has been an increasing use of a dialogue game approach in the areas of human computer interaction (e.g. Moore, 1993; Maudet & Moore, 2001; Ravenscroft & Pilkington, 2000), agent communication (e.g. Reed, 1998; Maudet & Chaib-Draa, 2002; McBurney & Parsons, 2001), mediation of legal reasoning (e.g. Bench-Capon, Geldard, & Leng, 2000; Freeman & Farley, 1996) and Artificial Intelligence in general (Walton, 2000). Indeed, in some literature, computational dialectics is seen as a new sub-field of Artificial Intelligence (Gordon, 1996).

THE DIALOGUE MODEL "DE"

A fundamental element of any computer dialogue system is the dialogue model. Following Moore (1993), which in turn derives from Walton (1998), it can be argued that the following set of criteria should be met by an underlying model for human-computer debate:

- the model needs to be persuasion style and symmetric in nature
- the set of move types provided by the model should be adequate for expression
- the model should leave enough room for strategy formation
- the model should be able to prevent fallacious argument
- the model should impose only a light cognitive load on the user
- the model should be computationally tractable

Using these criteria, we have conducted a comparative study of the most recent developments of dialogue models in the area of informal logic and computational dialectics (Yuan, 2004). The study suggests that Bench-Capon's (1998) system is explanation based rather than debate oriented. Prakken's (2000) and Lodder and Herczog's (1995) systems lack a question move type, and this may prevent students from asking the tutor questions and tutors from questioning the student's understandings; this is undesirable from an educational point of view (Veerman, Andriessen, & Kanselaar, 2002). Ravenscroft and Pilkington's (2000) system is asymmetrical and the dark side commitment of Walton and Krabbe's (1995) PPD would raise the cognitive load of the user. This is not of course to deny the general worth of these systems, but rather to suggest that they may not be a perfect match for the specific requirements of educational human computer debate. However, Moore's (1993) utilization of Mackenzie's system "DC" (Mackenzie, 1979) arguably meets most of the requirements. Further advantages of adopting DC as the underlying dialogue model are discussed by Moore and Hobbs (1996): the design of rules increases the computational tractability, and its symmetric nature enables either the tutor or the students to build their own positions. Further, Walton (1984) also suggests that the set of DC rules is practically useful. For these reasons, the dialogue game DC was chosen as the base system for further study.

We developed an amended dialogue model "DE" based on DC (Mackenzie, 1979) as the underlying model for our debating system. The motivation behind this development is that the underlying dialogue model of the debating system is required to have the ability to pick out fallacious argument and common errors when they occur during the course of debate (Yuan, Moore, & Grierson, 2003a). We had earlier used computational agents to debate with each other, using DC as their dialogue model, and thus to generate dialogue transcripts. Analysis of the transcripts facilitated the evaluation of the dialectical system DC (Yuan, 2004). The analysis revealed weaknesses in DC in preventing the fallacy of question begging, inappropriate challenges and the straw man fallacy, and in appropriate handling of the issue of repetition. In the light of this we constructed a modified system, "DE". DE was further evaluated using similar agent-based conversational simulations, and the result shows improvement over DC in preventing fallacious arguments and common errors. In particular, DE appears advantageous over DC in preventing the fallacy of question begging, inappropriate challenges and the straw man fallacy, and appropriate handling of the issue of repetition. Details of the specific differences between DC and DE are discussed in (Yuan, Moore, & Grierson, 2003b). The "DE" system can be outlined as follows (cf. Yuan, Moore, & Grierson, 2003b).

Available Move Types

The DE model makes the following move types available to both participants in the dialogue:

1) Assertions. The content of an assertion is a statement P, Q, etc. or the truth-functional compounds of statements: "Not P", "If P then Q", "P and Q".
2) Questions. The question of the statement P is "Is it the case that P?".
3) Challenges. The challenge of the statement P is "Why is it supposed that P?" (or briefly "Why P?").
4) Withdrawals. The withdrawal of the statement P is "no commitment P".
5) Resolution demands. The resolution demand of the statement P is "resolve whether P".

Commitment Rules

Each participant in a dialogue using the DE model owns a commitment store. Each commitment store contains two lists of statements: the "assertion list" contains the statements a participant has explicitly stated and the "concession list" contains the statements a participant has implicitly accepted (i.e. statements made by their interlocutor and against which they have raised no objection). The commitment rules are as follows:

1) Initial commitment, CR_0: The initial commitment of each participant is null.
2) Withdrawals, CR_W: After the withdrawal of P, the statement P is not included in the move maker's store.
3) Statements, CR_S: After a statement P, unless the preceding event was a challenge, P is included in the move maker's assertion list and the dialogue partner's concession list, and 'Not P' will be removed from the move maker's concession list if it is there.
4) Defence, CR_{YS}: After a statement P, if the preceding event was "Why Q?", "P" and "If P then Q" are included in the move maker's assertion list and the dialogue partner's concession list, and 'Not P' and 'Not (If P then Q)'are removed from the move maker's concession list if they are there.
5) Challenges, CR_Y: A challenge of P results in P being removed from the store of the move maker if it is there.

Dialogue Rules

Participants in a dialogue using the DE model are required to adopt the following rules:

1) R_{FORM}: Participants may make one of the permitted types of move in turn.
2) $R_{REPSTAT}$: Mutual commitment may not be asserted unless to answer a question or a challenge.
3) R_{QUEST}: The question P may be answered only by P, "Not P" or "no commitment P".
4) R_{CHALL}: "Why P?" must be responded to by a withdrawal of P, a statement acceptable to the challenger, or a resolution demand of any of the commitments of the challenger which immediately imply P. A statement S is acceptable to participant A at a stage n, just in case that S is at stage n (i) a commitment or (ii) a de facto commitment (e.g. participant A commits to Q, $Q \supset S$ or (iii) a new commitment of A.
5) $R_{RESOLVE}$: Resolution demands may be made only if the dialogue partner has in his commitment store an immediately inconsistent conjunction of statements, or withdraws or challenges an immediate consequent of his commitments.
6) $R_{RESOLUTION}$: a resolution demand must be followed by withdrawal of one of the offending conjuncts, or affirmation of the disputed consequent.
7) $R_{LEGALCHALL}$: "Why P?" may not be used unless P is on the assertion list of the dialogue partner.

A HUMAN-COMPUTER DEBATING SYSTEM

The approach, then, is to use the DE dialogue model outlined above as the basis for a human-computer debate system. A fully functional prototype, operationalising the dialogue model DE, has been built by the authors using the Java programming language. The system asks the user his opinion on the controversial issue of capital punishment, adopts the opposite position and engages the user in debate on the issue. The computer can adopt either a proponent or an opponent role. That is, if the user chooses to support the view of "capital punishment is acceptable", the computer will adopt the opposite view "capital punishment is not acceptable", and vice versa. The system then engages the user in debate on the topic of capital punishment, given these initial positions on the issue.

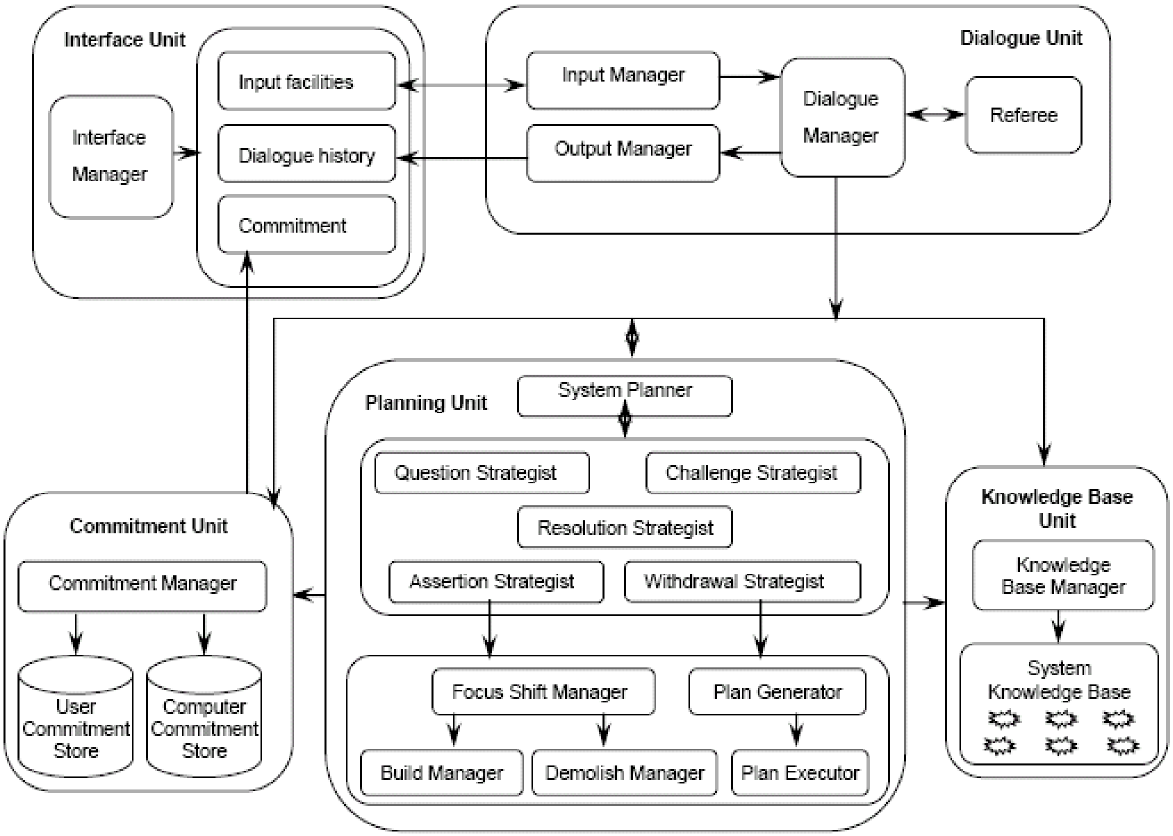

Fig.1. System Architecture.

The system architecture is shown in Figure 1. The remainder of this section provides details of the main components, input facilities and knowledge representation of the system.

System Components

There are five main units of the system: the interface unit, the dialogue unit, the commitment unit, the planning unit and the knowledge base unit. The *interface unit* provides the system's user interface (see Figure 2). It provides a dialogue history, which records the debate, and commitment stores to show both the user's and the computer's commitment store contents. Input facilities enable the user to select the move type and move content from a menu. The interface manager enables the user to save the debate history and to change the background colour of the interface, and provides help facilities.

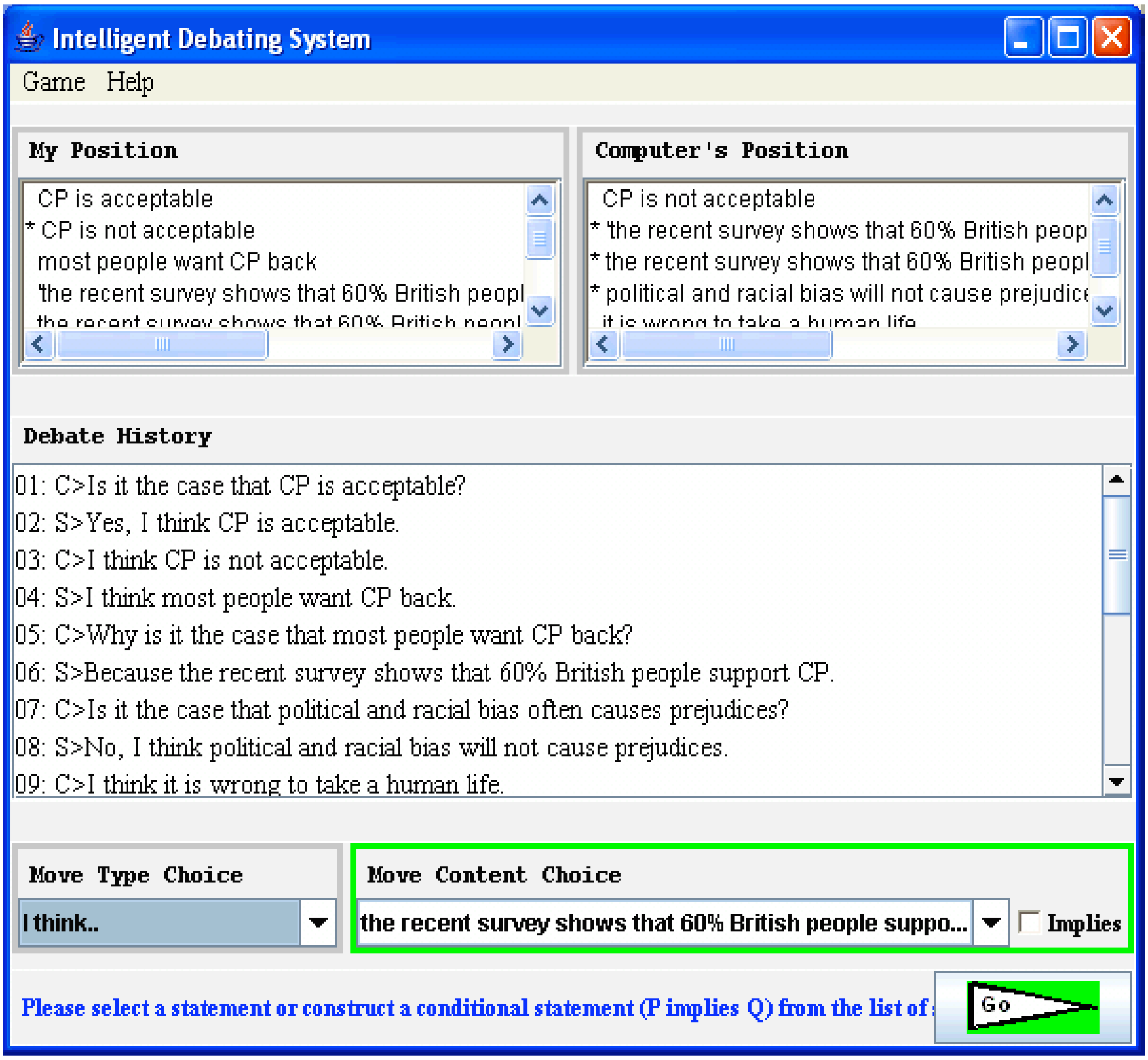

Fig.2. Human-Computer Debating System User Interface.

The *dialogue unit* can be regarded as the dispatch centre of the dialogue interactions. This unit consists of an input manager, a dialogue manager, a referee, and an output manager. The input manager provides dynamic support for the user's input, in that it makes available to the user only those move types permissible under the DE rules given the prevailing state of the dialogue. It then delivers the user input to the dialogue manager. The dialogue manager controls the turn taking of the interaction and is in charge of the input manager, output manager, referee, the commitment unit and the planning unit. Each user's move will be passed to the DE referee for judgement. If the move is legal, the commitment manager will be called to update the commitment stores, the output manager will be called to update the dialogue history and the planning unit will be called to make a move on behalf of the computer. In the event of an illegal move, the referee will post a message and request the user to make another move.

The *commitment unit* is responsible for updating the user's and the computer's commitment stores. It contains a commitment manager and two commitment stores, one for each party. The commitment manager will update both parties' commitment stores according to the DE commitment rules. Each commitment store is designed to have two lists of statements, those that have been explicitly stated by the owner of the store and those that have been merely implicitly accepted. In the current prototype, a statement that is only implicitly accepted is marked with an asterisk, as shown in Figure 2.

The *planning unit* is responsible for generating the computer's dialogue moves. To do so, it takes into account the system knowledge base, the prevailing state of both commitment stores and the dialogue rules. The system planner is key to the planning unit, and manages assertion, challenge, withdrawal, resolution and question "strategists", which are designed to deal with different dialogue situations following the set of heuristics discussed below. When the system planner receives calls from the dialogue manager, it will check the current dialogue situation and schedule the corresponding strategist to produce a move. Then the system planner will pass the move to the dialogue unit to make the computer's contribution. In addition, there are five components (focus shift manager, build manager, demolish manager, plan generator and plan executor) that are designed to provide special services to the assertion and the withdrawal strategists. The focus shift manager will be called by the assertion or withdrawal strategist to decide whether to change the current focus. The build and demolish manager will be called by the focus shift manager to check whether there are methods available to either build its own positions or attack the user's positions. The plan generator is responsible for generating a set of propositions and forming a line of questions when required by the assertion strategist, the build manager or the demolish manager. The plan executor is responsible for executing a plan. The assertion and withdrawal strategists will constantly look up whether there is a plan under execution, if there is, then they call the plan executor to carry on its execution.

The *knowledge base unit* consists of a knowledge base manager and a dedicated system knowledge base. The knowledge base contains a set of propositions and consequence relationships between these propositions; in the current prototype these relationships are based on a structure as shown in Figure 3. The knowledge base manager provides means of appropriately accessing the knowledge base given requirements from the planning unit. The capital punishment domain knowledge shown in Figure 4 is taken from (Moore, 1993), an experimental study of the dialogue game DC with human participants.

Input Facilities

The ambiguity of natural language and the precise input required by the computer renders the use of a general natural language interface currently impossible for the debating system. Moore (1993) argues that a menu driven interface may suffice to yield interesting human-computer debate. In addition, there are several precedents for the use of a menu approach in a computational dialogue game (e.g. Hartley & Hintze, 1990; Bench-Capon, 1998). Although it may be argued that the prescribed menu may not offer the line of argument the user would like to pursue (Moore, 1993), this may in practice be alleviated by enlarging the computer's knowledge base and hence providing more available options from which the user may select. In the light of the above arguments, a menu based approach is adopted in the current system.

Under this menu based approach, the user needs to make a double selection, choosing from the available move types and then from the list of prescribed propositions. Moore (1993) suggests that the legally available move type can be identified before the student makes a move, hence largely preventing the user from breaking the rules. Further, this arrangement may increase the learnability of the game since users are not required to remember the dialogue rules. This approach is therefore adopted by the current prototype.

Once the user has selected a move type, they need to select some propositional content. The system provides a number of means for doing this, depending on the nature of the move type. The details are as follows: (1) the move contents for resolution demand and challenge move types can be selected from the computer's commitment store; (2) the move contents for a withdrawal can be selected from the user's commitment store; (3) the move contents for assertion and question move types can be selected from the list of propositions (with the aid of the "implies" checkbox shown in Figure 2, the user may construct a conditional, e.g. $P \supset Q$). The location of propositions on the screen is highlighted with a green-coloured border. In addition, the message bar at the bottom of the user interface provides dynamic instructions to support user input (see Figure 2).

Knowledge Representation

Several ways of representing knowledge for use in argumentation systems can be found in the literature, e.g. Ravenscroft and Pilkington (2000) use a rhetorical structure theory (RST) approach and Bench-Capon (1998) and Freeman and Farley (1996) use revised versions of Toulmin's (1958) argument formalism. Moore (1993) argues that the knowledge base should be able to provide answers to questions, support for statements, and rebuttals of statements. To provide such a service for the debating system, the system knowledge base architecture is developed as illustrated in Figure 3.

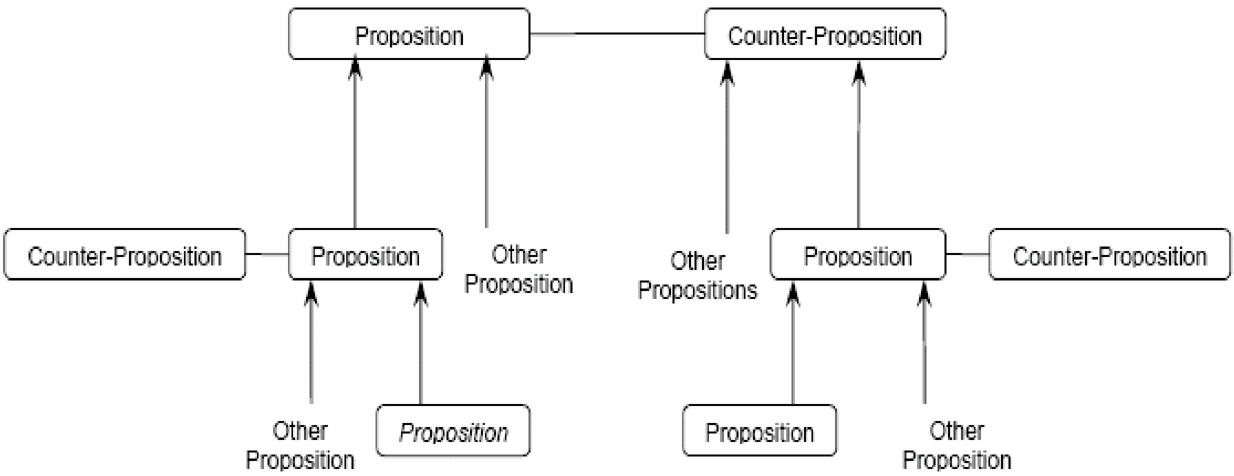

Fig.3. Knowledge Base Architecture.

Chaining of argument is allowed by this representation. Each proposition may have any number of supporting propositions and a counter proposition. The knowledge base distinguishes hard evidence from opinion; the former is represented by italics in Figures 3 and 4; the distinction is important in certain aspects of the system's dialogue strategy. This simple model is sufficient for the strategic needs of the current debating system, to which we now turn.

STRATEGIC HEURISTICS

A particular concern with DE, especially from a computational perspective, is that it leaves much to the discretion of the user of the model. For example, after a challenge (why P?), various options are open: one can respond with no commitment to P, or a resolution demand (in some circumstances) or a support for P. Further, there is no guidance within the rules as to the content of the support. Similarly, after a withdrawal or a statement, there are no restrictions on the move types or move contents. All DE does is to legitimise a set of move types given the prevailing circumstances, and occasionally give some indication of the semantic possibilities. In a human-computer debate setting, it is crucial therefore that the computer is given some means of selecting between available possibilities, e.g. to maintain focus after a statement or a withdrawal, so that the produced moves are appropriate at the pragmatic level. This choice must be based on some suitable strategy (Maudet & Moore, 2001). Appropriate strategic knowledge is, then, essential if the computer is to produce high quality dialogue contributions. Given this, this section will discuss the development of a set of strategic heuristics usable by a computer as a dialogue participant.

Philosophical investigation of dialogue, e.g. Walton (1996) and Krabbe (1999), has revealed rich sets of strategies that are used when people argue with each other. Similarly, in the area of computational dialectics, there are investigations of strategies for human-computer interaction (e.g. Moore, 1993), and agent communications (e.g. Amgoud & Maudet, 2002). In addition, there has been work on the notion of dialectical relevance (e.g. Walton, 1999; Prakken, 2000; Maudet, 2001). Each of these investigations has a different focus, e.g. Walton (1996) focuses on schemas for the generation of arguments; Krabbe's (1999) dialogue profile may serve as an instrument for exploring various strategic possibilities according to a "dialogue tree"; Moore (1993) and Amgoud and Maudet (2002) focus on decision making during the process of argument; dialectical relevance research tries to define the requirements for a move to be deemed relevant. All are potentially useful toward a theory of debating heuristics for a human-computer debate. Indeed, the proposal for strategy made in this section takes a hybrid approach drawing on each of these investigations.

System Profile

It is argued that some of the strategic decisions are highly dependent on the truth seeking nature of the system (Moore, 1993), e.g. after a question, a challenge or a resolution demand, in that a truth seeking system would always speak the truth while a non-truth seeking system may speak falsehood for the sake of argument (as in Maudet and Moore's (2001) "devil's advocate" discussion and Grasso, Cawsey and Jones' (2000) "partially cooperative" agent). Similarly, Amgoud and Maudet (2002) argue that the "profile" of an agent may influence its choice between the available possibilities. It is therefore necessary to consider the profile that might be adopted by an educational human-computer debating system.

Amgoud and Parsons (2001) make some proposals for a broad class of agent profiles: agreeable (accept whenever possible), disagreeable (accept only when no reason not to), open-minded (challenge only when necessary), argumentative (challenge whenever possible) and elephant child (question whenever possible). In addition, Amgoud and Maudet (2002) propose strategies for a prudent agent, that is one which is more reluctant to expose his argument than a non-prudent agent. In a similar vein, Grasso, Cawsey and Jones' (2000) system Daphne is defined as a "partially cooperative" agent, which is allowed to tell lies about her own beliefs only.

One of the primary motivations behind the development of our debating system, as argued in the *introduction* section earlier, is the expectation that it can be used to educational advantage – to develop students' debating and reasoning skills and domain knowledge. In the context of an educational human-computer debate, the computer is intended to be not only a debate competitor but also an intelligent tutor. From an educational point of view, whilst intuitively one may wish the system to speak the truth, on the other hand, it could be argued that some sort of deception may be inherent in the definition of dialectical argumentation (Grasso, Cawsey, & Jones, 2000) and in the playing of devil's advocate, yet both of these may be educationally valuable. A balance between trust and deception might therefore be required. It can be argued that the computer should be honest with respect to the publicly inspectable stores, since the system should be seen to be trustworthy (Grasso, Cawsey, & Jones, 2000). How, though, should the computer treat its knowledge base? The computer is required to have the ability to argue either as a proponent or an opponent of the topic under discussion, and this implies that the computer's knowledge base can support both the opponent view and proponent view (see Figure 4 for an example of the system knowledge base in the domain of capital punishment). As a result, the computer may constantly face inconsistent knowledge while making decisions (for example it can find both support for and objection to the notion that capital punishment acts as a deterrent). In this situation, it is suggested that the computer is allowed to insist on its own view for the sake of argument even though it may have more reasons in favour of the user's view. Given the above discussions, the system is currently configured as what can be described as a "partially honest" agent. Against this profile of the debating system, a set of debating heuristics can now be proposed.

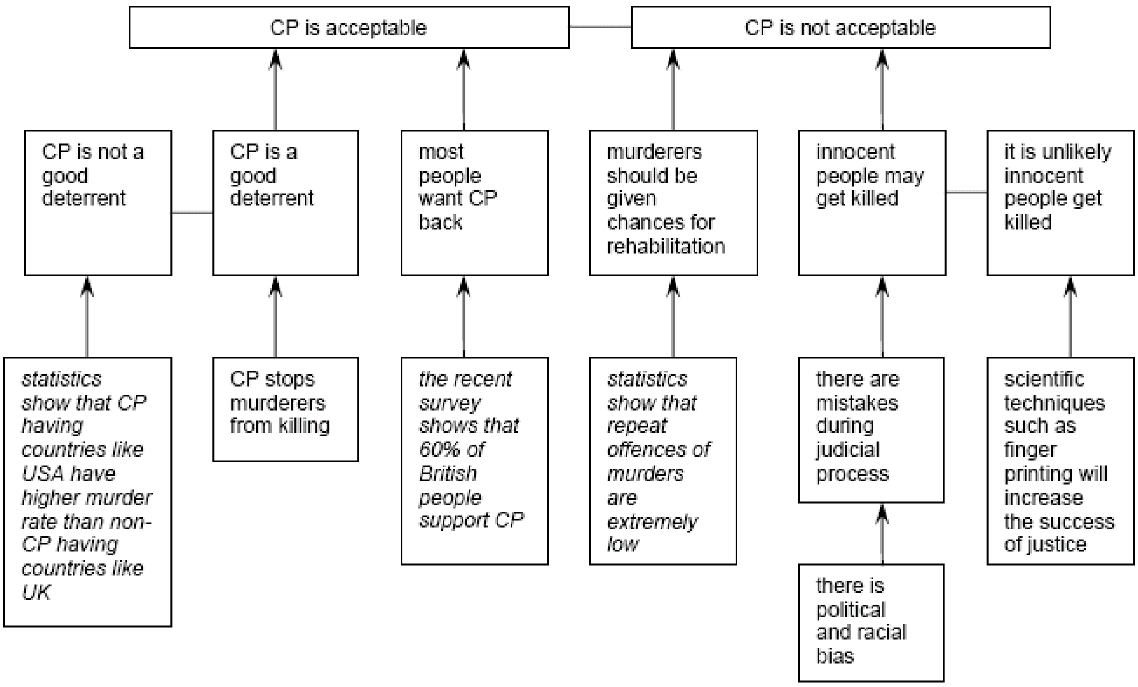

Fig.4. An Example of the System's Knowledge Base.

A Set of Strategic Heuristics

Given the nature of the DE model, there are five dialogue situations that the computer might face, defined by the previous move type made by the user: a challenge, a question, a resolution demand, a statement or a withdrawal. Each therefore needs to be considered in relation to the strategic decisions the computer might need to make. Following Moore (1993), these decisions are best captured at three levels.

1) Retain or change the current focus.
2) Build own view or demolish the user's view.
3) Select method to fulfil the objective set at level (1) and (2).

Moore argues that (1) and (2) refer to strategies which apply only when the computer is facing a statement or withdrawal, while (3) refers to tactics used to reach the aims fixed at levels (1) and (2), and applies in every game situation. These levels of decisions are discussed in turn below.

The level (1) decision concerns whether to retain the current focus or to change it. The decision, that is, involves whether to continue the attempt to substantiate or undermine a particular proposition. Moore (1993) argues that continuing to execute a plan of questions or addressing the previous move will guarantee the current focus is retained but that it is possible not to directly address the user's latest utterance yet still retain focus. Moore further suggests that there is a presumption in favour of addressing the previous move, but that this presumption may be broken when the line of questioning is deemed a blind alley, or if a successful removal of the user's support has been made, or if, on regaining the initiative after a period without it, a resolution demand can legally be made.

The decision at level (2) considers whether to adopt a build or a demolish strategy. A build strategy involves seeking acceptance of propositions that support the computer's own thesis while a demolish strategy seeks to remove the user's support for his thesis. The decision is needed only at the beginning of games and when the level (1) decision involves a shift in focus. A demolish strategy could possibly be part of a broader build strategy, e.g. a goal directed plan of questions building the computer's own view might involve removing some unwanted responses from the user. A building attempt might also be part of a broader demolish strategy, e.g. the computer is using a line of questions to build the case for P in order to attack the user's view $\neg$P. Moore found no evidence to suggest a priority between the build and demolish strategy. In the current debating system, we give the priority to a build strategy and make the computer try to open as many subtopics as possible, based on the view that in an educational debate the aim is to expose the full complexity of the situation and the desire might be to continue until all the knowledge base has been explored. Moore also argues that the decisions at level (1) and (2) heavily depend on the results of level (3) methods. In the current debating system, for example, the computer checks the level 3 methods first; if there are level (3) methods available, the level (1) and (2) decisions do not need to be applied; however, if there is no level (3) method available, level (1) and (2) decisions will come into play, in that the level (1) decision may be to switch the current focus and the level (2) decision is to build the computer's thesis if there are build methods available.

The third level of decisions applies to each of the dialogue situations. Each dialogue situation is discussed in turn below. To illustrate the proposed strategic heuristics in action, an example dialogue (generated by the current version of the debating system as it follows the knowledge base depicted in Figure 4) is provided. A full set of example dialogues is documented elsewhere (Yuan, 2004).

A Question Raised By the User

Questions asked may involve questioning an individual statement, e.g. "Is it the case that P?" or a conditional, e.g. "Is it the case that Q implies P?". In these situations, the computer is allowed by the DE rules to answer "Yes", "No" or "no commitment". Moore (1993) suggests that the decision must be based on the truth seeking nature of the game. In the current proposal, the system is required to be a partially honest agent, which may tell lies only for the sake of argument. In addition, Moore suggests one should give an answer in such a way as to avoid unwelcome commitment. Given this, heuristics for a situation in which the computer is facing a question can be proposed as follows (assume the question is "is it the case that P?").

(1) If neither P nor ¬P can be found in knowledge base (KB), then the computer speaks the truth with a "no commitment".

(2) If only one of them (P and ¬P) can be found in KB
   a. If the computer has previously uttered "no commitment" to the found statement, then it utters "no commitment" to remain consistent
   b. Else the computer speaks the truth and utters the found statement.

(3) If both (P, ¬P) are found in the computer's KB, and assuming that one of them (say ¬P) supports the computer's view and the other (say P) supports the user's view
   a. If the computer has an acceptable support for ¬P, then utter ¬P.
   b. If the computer has no acceptable support for ¬P, and the computer has not committed to propositions supporting P, the computer should utter "no commitment".
   c. If the computer has no acceptable support for ¬P, and the computer has committed to propositions supporting P, then the computer should utter P.

Heuristic (1) suggests that the computer should be honest if it does not know the answer. To make this decision, the computer needs to consider its own knowledge base only. Heuristic (2a) suggests that the computer should avoid an unwelcome commitment which has previously been withdrawn from its store. Heuristic (2b) suggests the computer should utter the only found answer truthfully. To make this decision, the computer needs to consider its knowledge base and its own commitment store.

Heuristic (3) will be used when the computer is facing inconsistent knowledge in that both answers are found in its KB. Under this situation, it was argued earlier, the computer may "tell lies" for the sake of argument. Heuristic (3a) suggests the computer should insist on its view ¬P for the sake of argument if it has an acceptable support for ¬P. Heuristic (3b) suggests the computer should utter "no commitment" if the computer has no acceptable reason for ¬P and the user has not provided a prima facie reason for P. The "no commitment" option in heuristic (3b) arguably prevents the computer from committing the fallacy of ignorance (Walton, 1998) which consists of accepting a proposition to be true simply on the basis that it has not been proved false, or that it is false because it has not been proved true (Copi & Cohen, 1994). Heuristic (3c) suggests the computer should accept the user's view P if the computer has no acceptable reason for ¬P and the user has provided prima facie reasons for P.

To illustrate the proposed strategic heuristics in action, an example dialogue is given below. The heuristic being invoked is indicated in parenthesis at the end of the appropriate line. In the transcripts that follow, CP stands for capital punishment, C stands for the computer and S stands for the student using the system (in the transcript provided, S's input is provided by one of the paper authors):

01: C>Is it the case that CP is acceptable?
02: S>No, I think CP is not acceptable.
03: C>I think CP is acceptable.
04: S>Is it the case that innocent people may get killed?
05: C>No, I think it is unlikely innocent people get killed. …………………. ……………………...............(3a)
06: S>Why do you think it is unlikely innocent people get killed?
07: C>Because scientific techniques such as fingerprinting will increase the success of justice.
08: S>Why do you think scientific techniques such as fingerprinting will increase the success of justice?
09: C>I don't know why scientific techniques such as fingerprinting will increase the success of justice.
10: S>Is it the case that innocent people may get killed?
11: C> I am not sure about it. ……..……………………………………………………………......…………(3b)
12: S>Is it the case that there are mistakes during the judicial process?
13: C>Yes, I think there are mistakes during the judicial process. ………………………...……................(2b)
14: S>Is it the case that there are mistakes during the judicial process is a prima facie reason for innocent people may get killed?
15: C>Yes, I think there are mistakes during the judicial process is a prima facie reason for innocent people may get killed. ………………..………………………………………………………………………(2b)
16: S>Is it the case that innocent people may get killed?
17: C>Yes, I think innocent people may get killed. …………………………...………………................(3c)
18: S>Is it the case that scientific techniques such as fingerprinting will increase the success of justice?
19: C>I am not sure about it. ……...…………………………………………………………………................(2a)
20: S>Is it the case that it is wrong to take a human life?
21: C> I am not sure about it. …….……………………………………………………………….…………..(1)

The game is set up in turn 01-03. The user adopts the view of "CP is not acceptable" and the computer automatically adopts the opposite view of "CP is acceptable". In turn 05, the computer faces a question and finds both answers for that question {"innocent people may get killed", "it is unlikely innocent people get killed"} in its KB. The computer is also aware that the proposition "innocent people may get killed" is a support of the user's thesis and the computer has an acceptable support "scientific techniques such as fingerprinting will increase the success of justice" for the proposition "it is unlikely innocent people get killed". The computer therefore fires the heuristic (3a) and insists on its view "it is unlikely innocent people get killed".

In turn 11, the computer faces the same question as in turn 05. The computer now has no acceptable support for the proposition "it is unlikely innocent people get killed" since the proposition "scientific techniques such as fingerprinting will increase the success of justice" is neither a commitment nor a de facto commitment nor a new commitment of the user's store due to the computer's withdrawal of the statement in turn 09. However, the computer has not committed to the reasons for the user's view "innocent people may get killed", the heuristic (3c) is therefore fired and the computer pretends to have no idea about it, for the sake of argument.

In turn 13 and 15, the computer again faces questions from the user. In both situations, only one of the answers can be found in the computer's KB and in neither case has the computer previously issued a "no commitment", the computer therefore fires the heuristic (2b) and speaks the truth, leaving the computer committed to the reasons for the user's view "innocent people may get killed".

In turn 17, the computer faces the same question as in turn 05 and in turn 11. The computer now has no acceptable support for the proposition "it is unlikely innocent people get killed", and has committed to the reasons for the user's view "innocent people may get killed". The computer therefore fires the heuristic (3b) and accepts the user's view "innocent people may get killed".

In turn 19, the computer is facing a question of the proposition "scientific techniques such as fingerprinting will increase the success of justice". The computer finds only one answer in its KB and the answer has been withdrawn by the computer itself in turn 09, the computer therefore fires the heuristic (2a) to avoid the unwelcome commitment.

In turn 21, neither answer to the user's question of the proposition "it is wrong to take a human life" can be found in the computer's KB, the computer therefore fires the heuristic (1) and honestly admits having no idea about it.

A Challenge Made By the User

There are three DE legal options available in response to a challenge: a resolution demand, a support, or a withdrawal. The first option concerns an inconsistency when the user is challenging a modus ponens consequence of his own commitments. From an educational point of view, it can be argued that the computer should point out this inconsistency and make the user aware of this kind of inconsistency in a debate. For the latter two options, Moore's (1993) experimental analysis suggests that one would normally reply with a carefully chosen support if available. In DE, there is no guidance within the rules as to the content of the support. The selection between alternative supports may be influenced by the profile of the agent. Given the definition of the profile of a partially honest agent, the computer should give a support according to its knowledge structure honestly rather than invent one which may not be a suitable support. In addition, it can be suggested that a support which can be further supported is preferred over one which cannot be further supported, since a further challenge might be expected from the user. Given this, the heuristics after a challenge of P can be proposed as follows:

(1) If P is a modus ponens consequence of the user's commitment, then pose a resolution demand
(2) Else if there is only one acceptable support available in the knowledge base, then state the support
(3) Else if there is more than one acceptable support available, then state the one that can be further supported
(4) Else if all the available acceptable supports are equally supported, then randomly choose one of the supports
(5) Else if no acceptable support is available, then withdraw P.

A Resolution Demand Made By the User

A resolution demand made by the user concerns an allegation that the computer has committed to an inconsistency in its commitment store. As a partially honest agent, the computer is required to be honest regarding the publicly inspectable stores. In the most likely event, the computer would face a resolution demand of the type 'resolve $\{\neg P, P\}$', in that the computer has committed to both P and ¬P. In this situation, the computer is required to withdraw one of them to keep consistent. Following Moore (1993), the computer should withdraw the one which is less supported according to its commitment store.

The user might invoke another type of resolution demand (i.e. resolve (Q, Q⊃P, why P) or resolve (Q, Q⊃P, no commitment P)) in the event of the computer's challenging or withdrawing a modus ponens consequence of its commitments. In this situation, the computer is required, by the game DE, to withdraw either Q or Q⊃P or affirm P. Moore (1993) argues that use of such a resolution demand would suggest that, in the user's view at least, the computer has challenged or withdrawn a proposition to which it ought to be committed given the remainder of its commitment store. In such a case, given the partially honest agent profile argued for earlier, the computer takes the option of affirming the disputed consequent P.

A "No Commitment" Made By the User

After a "no commitment", DE places no restrictions on either move type or contents. The computer's decisions are therefore more open. Following Moore (1993), the heuristics after a "no commitment" are proposed as follows:

(1) If the computer is facing a "no commitment" to a statement supporting the user's thesis
   a. If the withdrawn statement is a unique support of the user's asserted proposition Q, and Q is not the user's thesis, then challenge Q
   b. Else check whether the user retains adherence to the thesis.

(2) If the computer is facing a "no commitment" to a statement supporting the computer's thesis
   a. If the non-committal statement is a modus ponens consequence of the user's commitments, then pose a resolution demand
   b. Else switch the focus.

Heuristic (1) will be used when the computer is facing a "no commitment" to a statement supporting the user's thesis. Moore (1993) suggests in the event of withdrawal of a proposition at a lower level of a support hierarchy, seeking withdrawal of propositions uniquely supported by that proposition. Heuristic (1a) is therefore suggested to challenge the statement supported by that proposition being withdrawn. If no such statement supported by the withdrawn proposition exists, heuristic (1b) can be invoked to check whether the user has given up his thesis given the situation that, from the computer's point of view, the user might surrender since he has lost some part of the debate.

Heuristic (2) can be used if the computer is facing a "no commitment" to a statement supporting the computer's thesis. Heuristic (2a) suggests using a resolution demand if available. Walton (1998) argues that this kind of inconsistency, where the statement being withdrawn is a modus ponens consequence of the remainder of the speaker's commitments, cannot be passed unchallenged. If there is no such inconsistency in the user's store, it is suggested that the computer should switch the current focus since there is nothing to attack arising directly from the "no commitment".

A Statement Made By the User

After a statement, there is no restriction on either move types or move contents in DE. Intuitively, one would expect the user to assert a statement which supports his view or opposes the computer's view. However, it is possible that the user may unwisely make a statement which supports the computer's view or goes against his own view. The computer may need to deal with these two kinds of statement differently. When the computer is facing a statement (say P) which supports the computer's thesis or militates against the user's view, two heuristics are proposed as follows:

a. If P is a support of the computer's thesis, then use P as the starting point to build a case for the computer's thesis
b. Else check whether the user still adheres to his thesis.

When the computer is facing a statement (say P) which supports the user's view or militates against the computer's view, a set of heuristics is proposed as follows, in line with Moore (1993):

a. If there is an inconsistency (i.e. (P, ¬P)) in the user's commitment store, then ask for resolution
b. Else if there is a piece of hard evidence support of ¬P, then state the piece of hard evidence
c. Else if there is any support of ¬P and the support (say Q) can be further supported, then state ¬P or state Q if ¬P has been uttered, or form a plan of questions making the user accept ¬P
d. Else if there is any support of ¬P and the support cannot be further supported, then form a plan of questions making the user accept ¬P
e. Else if P is challengeable, then challenge it.

Heuristic (a) suggests the computer should point out the user's obvious inconsistency. Walton (1998) argues that this kind of inconsistency should not be passed unchallenged. Moore (1993) also suggests that this heuristic can be regarded as the strongest attack of the opponent's positions. Therefore, this heuristic has highest priority.

Heuristic (b) suggests the computer should state a piece of hard evidence which directly contradicts the user's view. There are two options in heuristic (c), the first option can be considered as a rebuttal of the user's view, and is borrowed from Amgoud and Maudet (2002); the second option considers forming a plan of questions. Heuristic (d) suggests the computer should form a plan of questions only for a less supported point. The priority between heuristics (b), (c) and (d) concerns how the computer would select the strongest argument. Moore (1993) suggests that one would like to state a piece of hard evidence (e.g. official statistics) rather than to pose it as a question. Heuristic (b) is therefore considered as the first option. An argument whose support can be further supported is considered as more secure than an argument whose support cannot be further supported since the former can provide reasons for the user's potential forthcoming challenge while the latter cannot. Heuristic (c) is therefore considered to have higher priority than heuristic (d). The selection from the two options within heuristic (c) is currently made on an arbitrary basis.

Heuristic (e) could be used if the user's view P is challengable. To decide whether a statement is challengeable, the computer needs to consider the nature of that statement (e.g. whether it is a piece of hard evidence) and the relevant DE dialogue rules. If the computer arrives at option e and the statement in question is not challengeable, the computer reverts to level 1 of the strategic decision-making process.

A further concern is how the plan of questions in heuristics (c) and (d) are organised. Following Walton's (1996) scheme of "argument from gradualism", the plan can be started by asking a question of a proposition (say A), followed by a series of connected conditionals (say $A \supset B$, $B \supset C \ldots . C \supset P$) toward the conclusion (say P). Moore (1993) argues that the computer should hand over the initiative by stating the conclusion P at the end if the plan is executed successfully, with a view to avoiding a one-sided dialogue.

During a plan execution, the user might give unwanted answers (i.e. answers not favourable to the computer's plan). The approach taken here is that the computer tries to remove the obstacles (unwanted answers) and put the plan back on track while the initiative is still held. The plan execution process is as follows:

1) If a wanted answer is given, then carry on to execute the plan
2) If a non-committal answer is given

(2.1) If there is an expressed inconsistency in the user's CS, then pose the appropriate resolution demand
   a) If the user affirms the disputed consequence, then continue the plan
   b) Else abandon this line of questioning
(2.2) Else abandon this line of questioning.

3) If an unwanted answer (e.g. ¬P rather than P) is given
(3.1) If there is an expressed inconsistency in the user's commitment store and the unwanted answer ¬P is an element of the inconsistency, then pose the appropriate resolution demand
   a) If the unwanted answer is withdrawn, then continue the plan and re-pose the question
   b) Else abandon this line of questioning
(3.2) Else if the unwanted statement is challengeable, then challenge the unwanted statement
   a) If the unwanted answer is withdrawn, then continue the plan to re-pose the question of P
   b) Else abandon this line of questioning
(3.3) Else abandon this line of questioning.

EVALUATION

Two types of evaluations have been done: expert evaluation and user based evaluation. A preliminary usability evaluation of the human computer debating prototype has been carried out and documented in (Yuan, 2004). Three HCI experts were invited to evaluate the human computer debating system. One expert preferred to evaluate the system cooperatively with the system author, in that the system author noted down the pertinent issues while the evaluator operated on the system (in effect, adopting a cooperative evaluation approach (Dix, Finlay, Abowd, & Beale, 2004)). In addition, the expert agreed to take part in a short interview after the cooperative evaluation session. After the evaluation, the notes of this evaluation were formalised by the system author and emailed to the evaluator to check their accuracy. The two other HCI experts preferred to evaluate the system at their own convenience. The debating system was emailed to these experts. Formal feedback was emailed back to the system author after their evaluations.

Full details of this evaluation can be found in (Yuan, 2004). Essentially, the expert evaluations give positive evidence concerning the usability of the system in general, and of the DE dialogue model and the proposed strategy in particular. This is supported by the evaluators' views on their experiences of operating on the system, such as "definitely easy for students who are familiar with computers", "very straightforward to use it", "no procedures annoyed me while operating on the system", "the system's overall performance is acceptable".

There are, however two weaknesses concerning the proposed strategy that were revealed. One participant reported that she found it a bit uncomfortable when the computer constantly hands over its turn after a period of debate. She further suggests that "this is fine, to make me to explore more argument. I would say it depends on personality of the debate participants". The second weakness is that the system should make a concession at the right time. The evaluator wrote: "after two long debates with the computer, it seemed to let me win. Though it is not clear why at that point it changed its mind. During these debates I thought I had the computer agree to a series of propositions that would lead it to change its initial position but it seemed to hold these incompatible ideas, without difficulty. When it did concede, it was a surprise to me.". This reflects the issue that there is no heuristic available for the computer to voluntarily concede a debate except when the user checks its thesis adherence. At some point, the computer should concede the debate voluntarily, i.e. when its thesis supports have all been removed from its commitment store and the user's thesis support added into its store, whereas currently the computer, without a heuristic for voluntary concession, simply hands over its turn to the user. Current work involves amending the system to cater for this concern.

A user based evaluation of the debating system has also been carried out and is documented in (Ævarsson, 2006). Ten students interested in debate were drawn from the University of Akureyri. Three were from the computer science department, two from the education faculty and five from the business faculty. No participant took part in more than one study. There were two pilot studies, in order to determine the set up of the experiment, followed by eight further studies. Each study was conducted in three stages: introduction, debate and interview. Prior to each study, the debating system was set up on the screen by the researcher. An English-Icelandic online dictionary was provided since the participants were native Icelandic and English was their second language. The introduction session involved the briefing of the purpose and procedures of the study by the researcher. In the debate session, the participant was asked to conduct a debate with the computer for 15 minutes in the pilot studies; this was extended to 20 minutes for the subsequent studies. After the debate session, a semi-structured interview was carried out concerning users' experience of using the system. The dialogue transcripts were saved and the interviews were taped for subsequent analysis.

The results are summarized in Table 1. The ten participants successfully conducted a debate with the system without difficulties. 462 turns were generated. The longest debate took 79 turns and the shortest 29 turns. Incidents of users breaking the DE rules were very rare. Six participants did not break the rules at all. Two participants seemed to have changed their original view on capital punishment, 8 of the debates ended with a stalemate, and none of the participants made the computer concede.

Table 1
Summary of Dialogue Transcripts

Participant No.	Number of turns	Rule break	Debate result
1	50	2	user concedes
2	39	0	stalemate
3	62	0	user concedes
4	29	2	stalemate
5	49	0	stalemate
6	79	4	stalemate
7	42	1	stalemate
8	40	0	stalemate
9	42	0	stalemate
10	30	0	stalemate
Total	462	9	

The interview transcripts were analysed under four headings: system intelligence, user enjoyment, value of the system and the user interface issues. All of the participants agreed that the system is intelligent and a worthy debate component though some participants would like to see the system more aggressive and more attacking. Eight participants said they enjoyed playing the game, and they particularly liked the nondeterministic nature of the system's dialogue contributions. Two participants said they felt frustrated a little with the input facilities though they managed to debate with the system.

All participants claimed they would like to play the game again were it available as an Internet game. The value of the game was affirmed by the participants. All participants agreed that the system can be used to help them practise argumentation. One participant recommended that the system could be used as an aid to the *Dialectics* course. Participants from the education faculty said they would like to see the system tailored for child education in many areas.

The user based evaluation revealed several concerns with the user interface. First, in the beginning, the participants had to spend some time in figuring out where the debate took place, whether it was in the student's commitment store, the computer's commitment store or in the debate history window. Once they had worked this out, most participants stayed focused on the debate history rather than the commitment stores. Participants expressed that they would prefer a one-window arrangement like the MSN communication program rather than the current three-window setup. Second, the participants found it confusing sometimes being guided to select move contents from the commitment stores on the top of the screen while the major input facilities were located at the bottom of the screen. Thirdly, participants were not happy with constantly clicking the *move content choice* then scrolling down in order to find a suitable proposition. Finally, participants were sometimes confused with the statement with a * prefix in the commitment stores.

To address these user interface issues, the user interface of the system is being rearranged and tested further. The first three weaknesses are being addressed together by moving the commitment stores to the *Move Content Choice* panel which has been redesigned as a tabbed panel with three tabs: prescribed propositions, user commitment store and the computer commitment store. Thus the user interface becomes a one-window setup, and at the same time the space for the list of visible move content choice is increased. The final weakness is being addressed by colouring different types of commitments rather than simply using an * prefix.

CONCLUSION AND FURTHER WORK

A new dialogue model DE has been proposed, using a human-computer debating prototype which has been constructed. In order to enable the computer to act as a dialogue participant, a set of computational strategic heuristics has been developed based on literature in the area of informal logic and computational dialectics, and incorporated into the debating system. The system has been evaluated by HCI experts and potential users of the system. The results are essentially favourable. The evaluations suggest that the novice users can successfully operate the system without training. The evaluations also provide evidence for the educational and entertainment value of the system.

We believe that the work reported makes a valuable contribution to the fields of dialectics and of human-computer dialogue. Concerning the former, we have developed a robust dialogue model, DE (Yuan, Moore, & Grierson, 2003b) and proposed a set of strategies to be utilised with the model. Further, because the computer system we have built can readily be adapted to function with a different dialogue model and/or a different set of strategies, it potentially provides people working in the field of dialectics with a test bed within which they can experiment with new models and new strategies they develop (Maudet & Moore, 2001; Amgoud & Maudet, 2002).

Turning to the contribution of this work to human-computer dialogue design, it can be argued that the study of such dialogue encompasses two broad issues. The first concerns the manner in which the user will interact with the machine, whether, for example, they will be presented with a menu of options, a command line interface, or some form of question-answer interaction. The second area concerns the attempt to emulate dialogue of the sort engaged in by human conversants, in which topics of mutual concern are discussed, and in which each conversant has substantive points to make; the interest is in the content and form of the dialogue, rather than the manner of its physical implementation. It is the second area to which our work chiefly contributes. It does this, we argue, in two ways. It indirectly contributes via the contribution to dialectics we have just outlined. Given the usefulness of a dialectical approach to interactive computer systems (Moore, 1993; Moore & Hobbs, 1996; Yuan, 2004), any development of dialectics per se potentially has a pay-off in terms of human-computer dialogue. Our work also makes a more direct contribution to human-computer dialogue, in that the debate system is a unique system and therefore makes a contribution to the broadening of the human-computer interaction "bandwidth" in general, and to the development of computer-based educational debate in particular.

There is a variety of interesting ways in which the research can be carried forward. In the short term, the basic system can be enhanced to allow the user to question or challenge a conjunction of statements (e.g. P ^ Q), or conditional. Currently the DE system is perhaps over-zealous re the challenge move, in that saying "why P" removes commitment to P, whereas a dialogue participant may wish to remain committed to P but hear his interlocutor's reasons for P. Relaxing this rule could be implemented and tested. Further features, such as the system explaining its tactics and reasons for choosing one move over another to the user, along the lines of (Pilkington & Grierson, 1996), can also be implemented and tested further. The system could then be evaluated with a number of different domains of debate, e.g. abortion, politics, terrorism, to test the extent to which the design and knowledge representation are generic. This evaluation might be extended to encompass the use of the system to investigate pedagogic issues, such as the educational value of one to one debate, and how learners make inferences about the knowledge domain (Moore, 1993). The evaluation could also be used to chart, through and across dialogues, how the way in which students engage in dialogue evolves. Ultimately, the system can perhaps be enhanced to keep track of student learning as such.

We also plan to permit freer user input, initially via an option to enter fresh propositional content in addition to selecting from those made available by the system. This will enable us to build up the system knowledge base by adding new claims to it and, more importantly, to experiment with the extent to which the strategic heuristics can cope with such new input.

REFERENCES

Amgoud, L., & Maudet, N. (2002). Strategical Considerations for Argumentative Agents. In S. Benferhat and E. Giunchiglia (Eds.) *Proceedings of the Ninth International Workshop on Non-Monotonic Reasoning* (*NMR'2002*) *Special Session on Argument, Dialogue, and Decision* (pp. 409-417). Toulouse.

Amgoud, L., & Parsons, S. (2001). Agent Dialogue with Conflicting Preferences. In J.-J. C. Meyer & M. Tambe (Eds.) *Intelligent Agents VIII: 8th International Workshop, ATAL 2001* (pp. 190-205). Berlin: Springer.

Baker, M. (2000). The Role and Models in Artificial Intelligence and Education Research: a Prospective View. *International Journal of Artificial Intelligence in Education,* 11, 122-143.

Beck, J. E., & Stern, M. K. (1999). Bring Back the AI to AI & ED. In S.P. Lajoie & M. Vivet (Eds.) *Artificial Intelligence in Education – Open Learning Environments: New Computational Technologies to Support Learning, Exploration, Collaboration* (pp. 233-240). Amsterdam: IOS Press.

Bench-Capon, T. J. M. (1998). Specification and Implementation of Toulmin Dialogue Game. In J.C. Hage, T.J.M. Bench-Capon, A.W. Koers, C.N.J. de Vey Mestdagh & C.A.F.M. Grutters (Eds.) *Legal Knowledge Based Systems: JURIX: The Eleventh Conference Proceedings* (pp. 5-20). Nijmegen: Gerard Noodt Institute (GNI).

Bench-Capon, T. J. M., Geldard, T., & Leng, P. H. (2000). A Method for Computational Modelling of Dialectical Argument with Dialogue Games. *Artificial Intelligence and Law,* 8, 233-254.

Bouwer, A. (1999). ArgueTrack: Computer Support for Educational Argumentation. In R. Pilkington, J. McKendree, H. Pain & P. Brna (Eds.) *Proceedings of Workshop on Analysing Educational Dialogue Interaction* (pp. 1-8). International Conference on Artificial Intelligence in Education 1999, Le Mans, France.

Copi, I. M., & Cohen C. (1994). *Introduction to Logic*. Prentice Hall.

Cumming, G., & McDougall, A. (2000). Mainstreaming AIED into Education. *International Journal of Artificial Intelligence in Education,* 11, 197-207.

Cunningham-Atkins, H., Powell, N., Moore, D., Hobbs, D., & Sharpe, S. (2004). The Role of Cognitive Style in Educational Computer Conferencing. *British Journal of Educational Technology,* 35(1), 69-80.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (2004). *Human Computer Interaction*, third edition. Prentice Hall.

Ævarsson, K. (2006). *Human-Computer Debating System, Evaluation and Development*. Unpublished BSc. Dissertation. University of Akureyri, Iceland.

Freeman, K., & Farley, A. (1996). A Model of Argumentation and Its Application to Legal Reasoning. *Artificial Intelligence and Law,* 4, 163-197.

Garrison, D. R. (1991). Critical Thinking and Adult Education: A Conceptual Model for Developing Critical Thinking in Adult Learners. *International Journal of Lifelong Education*, 10(4), 287-304.

Girle, R. A. (1992). Eliza and the Automata. In *Proceedings of the Third Annual Conference on AI Simulation and Planning in High Autonomy Systems* (pp. 287-293). Perth.

Gordon, T. (1996). Computational Dialectics. In P. Hoschka (Ed.) *Computers as Assistants – A New Generation of Support Systems* (pp. 186-203). Mahwah, NJ: Lawrence Erlbaum Associates.

Grasso, F., Cawsey, A., & Jones, R. (2000). Dialectical Argumentation to Solve Conflicts in Advice Giving: a Case Study in the Promotion of Healthy Nutrition. *International Journal of Human Computer Studies,* 53, 1077-1115.

Grasso, F., & Leng, P. (2003). Issues of Quality in Online Degree Programmes. In C. Ghaoui (Ed.) *Usability Evaluation of Online Learning Programs* (pp. 198-217). Information Science Publishing.

Hamblin, C. (1971). Mathematical Models of Dialogue. *Theora,* 37, 130-155.

Hartley, J. R., & Hintze, D. (1990). Dialogue and Learner Modelling. In S. A. Cerri (Ed.) *Student Model Acquisition in a Natural Laboratory (NATLAB),* GEC DELTA Project D-1016 Final Report, Brussels.

Hatcher, D. (1999). Why Formal Logic is Essential for Critical Thinking. *Informal Logic*, 19(1), 77-89.

Krabbe, E. (1999). Profile of Dialogue. In J. Gerbrandy, M. Marx, M. de Rijke & Y. Venema (Eds.) *JFAK: Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday, III* (pp. 25-36). Amsterdam: Amsterdam University Press.

Laurillard, D. (1995). Multimedia and the Changing Experience of the Learner. *British Journal of Educational Technology*, 26(3), 179-189.

Lodder, A. R., & Herczog, A. (1995). Dialaw, A Dialogue Framework for Modelling Legal Reasoning. In *Proceedings of the Fifth International Conference on Artificial Intelligence and Law* (pp.146-155). New York: ACM.

Mackenzie, J. (1979). Begging the Question in Non-Cumulative Systems. *Journal of Philosophical Logic*, 8, 117-133.

Mackenzie, J., & Staines, P. (1999). Hamblin's Case for Commitment: A Reply to Johnson. *Philosophy and Rhetoric*, 32(1), 14-39.

Maudet, N. (2001). Notes on Relevance in Dialectical Systems. In A. Hunter & S. Parsons (Eds.) *Proceeding of the ECSQARU'2001 Workshop Adventures in Argumentation* (pp. 36-42). Toulouse.

Maudet, N., & Chaib-Draa, B. (2002). Commitment-Based and Dialogue-Game-Based Protocols: New Trends in Agent Communication Languages. *The Knowledge Engineering Review*, 17(2), 157-179.

Maudet, N., & Moore, D. (2001). Dialogue games as Dialogue Models for Interacting with, and via, Computers. *Informal Logic,* 21(3), 219-243.

McBurney, P., & Parsons, S. (2001). Agent Ludens: Games for Agent Dialogues. In P. Gmytrasiewicz & S. Parsons (Eds.) In *Proceedings of the 2001 AAAI Spring Symposium on Game Theoretic and Decision Theoretic Agents* (pp.70-77). Menlo Park, CA: AAAI Press.

Moore, D. (1993). *Dialogue Game Theory for Intelligent Tutoring Systems.* Unpublished Doctoral Dissertation, Leeds Metropolitan University.

Moore, D. (2000). A Framework for Using Multimedia within Argumentation Systems. *Journal of Educational Multimedia and Hypermedia,* 9(2), 83-98.

Moore, D., & Hobbs, D. (1996). Computational Use of Philosophical Dialogue Theories. *Informal Logic*, 18(2), 131-163.

National Curriculum Council (1990a). *Curriculum Guidance 7 – Environmental Education*. NCC, York.

National Curriculum Council (1990b). *Curriculum Guidance 8 – Education for Citizenship*. NCC, York.

Pilkington, R. M. (1998). Dialogue games in Support of Qualitative Reasoning. *Journal of Computer Assisted Learning*, 14, 308-320.

Pilkington, R. M. (1999). *Analysing Educational Discourse: the DISCOUNT Scheme, Version 3.* Technical Report No. 99/2, Computer Based Learning Unit, The University of Leeds.

Pilkington, R. M., & Grierson, A. (1996). Generating Explanations in a Simulation-Based Learning Environment. *International Journal of Human-Computer Studies*, 45(5), 527-551.

Pilkington, R. M., & Mallen, C. (1996). Dialogue Games to Support Reasoning and Reflection in Diagnostic Tasks. In P. Brna, A. Paiva & J. Self (Eds.) *Proceedings of the European Conference on Artificial Intelligence in Education* (pp. 213-219). Lisbon: Editicões Colibri.

Prakken, H. (2000). On Dialogue Systems with Speech Acts, Arguments, and Counterarguments. In *Proceedings of the Seventh European Workshop on Logic for Artificial Intelligence* (JELIA'2000) (pp. 224-238). Berlin: Springer Verlag.

Quinn, V. (1997). *Critical Thinking in Young Minds*. David Fulton Publishers.

Ravenscroft, A., & Matheson, P. (2001). Carpe Diem: Models and Methodologies for Designing Engaging and Interactive e-learning Discourse. In T. Okamoto, R. Hartley, Kinshuk & J. P. Klus (Eds.) *Proceedings of IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges* (pp.74-77). IEEE Computer Society.

Ravenscroft, A., & Matheson, P. (2002). Developing and Evaluating Dialogue Games for Collaborative E-learning. *Journal of Computer Assisted Learning,* 18, 93-101.

Ravenscroft, A., & Pilkington, R. (2000). Investigate by Design: Dialogue Models to Support Reasoning and Conceptual Change. *International Journal of Artificial Intelligence in Education,* 11, 237-298.

Reed, C. (1998). Dialogue Frames in Agent Communication. In *Proceedings of the Third International Conference on Multi Agent Systems (ICMAS98)* (pp.246-253). Paris: IEEE Press.

Self, J. (1992). Computational Viewpoints. In R. Moyse & M. T. Elsom-Cook (Eds.) *Knowledge Negotiation*. Academic Press.

Toulmin, S. (1958). *The Uses of Argument.* Cambridge University Press.

Veerman, A., Andriessen, J., & Kanselaar, G. (2002). Collaborative Argumentation in Academic Education. *Instructional Science,* 30, 155-186.

Walton, D. (1984). *Logical Dialogue-Games and Fallacies*. Lanham, Maryland: University Press of America.

Walton, D. (1996). *Argumentation Schemes for Presumptive Reasoning*. Mahwah, N. J: Erlbaum Associates.

Walton, D. (1998). *The New Dialectics: Conversational Context of Argument*. University of Toronto Press.

Walton, D. (1999). Dialectical Relevance in Persuasion Dialogue. *Informal Logic*, 19, 119-143.

Walton, D. (2000). The Place of Dialogue Theory in Logic, Computer Science and Communication Studies. *Synthese,* 123, 327-346.

Walton, D., & Krabbe, E. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning.* State University of New York Press.

Yuan, T., Moore, D., & Grierson, A. (2003a). A Conversational Agents System as a Test-Bed to Study Philosophical Model "DC". In *Proceedings of IJCAI'2003 Workshop on Computational Models of Natural Argument* (pp. 45-50). Acapulco, Mexico.

Yuan, T., Moore, D., & Grierson, A. (2003b). Computational Agents as a Test-Bed to Study Philosophical Model "DE", A Development of Mackenzie's "DC". *Informal Logic,* 23(3), 263-284.

Yuan, T. (2004). *Human-Computer Debate, a Computational Dialectics Approach*. Unpublished Doctoral Dissertation, Leeds Metropolitan University.