

Assessing debate strategies via computational agents

Tangming Yuan^{a*}, David Moore^b and Alec Grierson^b

^aComputer Science Department, University of York, Heslington, York, YO10 5AT, UK; ^bInnovation North, Leeds Metropolitan University, Beckett Park, Leeds LS6 3QS, UK

(Received 2 July 2009; final version received 24 February 2010)

This paper reports our research concerning dialogue strategies suitable for adoption by a human–computer debating system. We propose a set of strategic heuristics for a computer to adopt to enable it to function as a dialogue participant. In particular, we consider means of assessing the proposed strategy. A system involving two agents in dialogue with each other and a human–agent debate system are constructed and subsequently used to facilitate the evaluations. The evaluations suggest that the proposed strategy can enable the computer to act as an effective dialogue participant. It is anticipated that this work will contribute towards the development of computerised dialogue systems and help to illuminate research issues concerning strategies in dialectical systems.

Keywords: dialogue game; dialogue strategy; computational agents; human–computer debate

1. Introduction

Computer-based learning, aka *e-learning* or *technology-enhanced learning*, is becoming increasingly important in education, in general, and in higher education, in particular. And it does seem to have many advantages. For example, students may learn at their own pace and convenience, remotely from the university campus if necessary. A possible danger with e-learning, however, is that the educational interactions may become unduly didactic (Moore 2000). Computer-based learning systems, that is, tend to adopt what Freire (2000) calls the *banking concept of education* – students are (partially) empty vessels into which suitable knowledge is *deposited*. This concern is very real today; the use, for example, of multiple choice *quizzes* within webCT exemplifies the banking concept: there is a single right answer to each question, the aim is to get students to learn the right answers. While this approach may be valid in some domains, its approach is limited in attempts to encourage critical thinking and reflection, the vital components of higher education.

We argue (e.g. Moore 2000; Naim, Moore, Yuan, Dixon and Grierson 2009) that there are two possible approaches to addressing this problem of untoward didacticism. One is to allow multiple participants in the learning interactions, so that learners are able to use the environment to communicate with each other and their tutors. A different approach is to have the computer itself be a participant in the learning interaction. If the computer is to engage in dialogue with students, it needs a model of dialogue, and such a model is potentially provided by computational dialectics. Computational dialectics is a maturing strand of research that is focused on computational utilisation of the dialogue games developed in the area of Informal Logic, which is an area of philosophy rich in models of communication and discourse, with a heavy focus on argument and “dialogue games”. If the dialectical model (e.g. dialogue game) is, as purported, a model of fair and reasonable dialogue, and if both the computer and students follow the dialogue, then fair and

*Corresponding author. Email: tommy@cs.york.ac.uk

reasonable dialogue will ensue (Moore 1993). Our interest is to use a dialogue game as a vehicle for an educational human–computer dialogue system.

There are many types of dialogue interactions in which people reason together, such as debate, persuasion, inquiry and information-seeking (Walton and 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 a rich educational benefit. A particular concern with our research therefore is to investigate the issues surrounding a computer-based system for educational debate.

Earlier, we developed an amended dialogue model DE (Yuan 2004) 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. 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 (Yuan, Moore and Grierson 2003).

A particular concern with DE, however, 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 a “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 therefore crucial 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.

Appropriate strategic knowledge is, then, essential if the computer is to produce high-quality dialogue contributions. The importance of strategies in dialectical systems has also been stressed elsewhere (e.g. Bench-Capon 1998; Walton 1998; Amgoud and Parsons 2001; Maudet and Moore 2001; Amgoud and Maudet 2002; Rahwan, McBurney and Sonenberg 2003). A set of strategies, enabling a computer to act as a debate participant was therefore proposed based on experimental study of the DC game with human participants (Moore 1993) and subsequently further developed in (Yuan 2004). However, the issue of whether the proposed strategy can in practice provide adequate services for a computer acting as a dialogue participant to produce good dialogue contributions cannot be settled on an *a priori* basis. The aim of this paper is to outline our work seeking to evaluate the proposed strategy.

To assess the appropriateness of a proposed strategy, Maudet and Moore (2001) suggest that the strategic heuristics need to be tested and that a convenient way to do this is via generation of dialogue by the computer itself. There are two possible ways to approach this: (1) allow two computational agents to engage in dialogue with each other and then study the results and (2) enable a human user to debate with a computerised debating system. Both approaches are seen as important to evaluate the proposed strategy from different perspectives. The former approach focuses on assessing whether there are unexpected new situations, requiring new heuristics, which have been missed in the current proposal and assessing whether the computationally generated dialogues are reasonably sound from a dialectical point of view. The latter approach, however, focuses on assessing the usability of the proposed strategy from the users' point of view. An agent–agent assessment is necessary prior to user-based assessment to avoid issues such as missing heuristics and apparent flaws appearing in and interfering with more expensive user studies. Both

approaches are therefore used in this study. A prerequisite for the study is the construction of suitable computational agents.

The remainder of the paper is organised as follows. First, we introduce the game DE and our current set of strategic heuristics. Secondly, we discuss the construction of a set of computational agents that can engage in debate with each other, generating dialogue transcripts for subsequent analysis. Thirdly, we discuss our human–computer debating system with which the user-based evaluations were carried out. We then discuss related work in this area and the significance of this work. We finally draw the conclusion and discuss our intended future work.

2. The dialogue model DE

The DE system is set up with two participants in dialogue with each other. Participants' moves are regulated by a set of rules, which prohibits illegal events. The set of rules is outlined as follows (cf. Yuan et al. 2003).

2.1. 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”.

2.2. 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). Commitments arise purely as a result of moves in the dialogue game. The concept is distinct, therefore, from that of *beliefs*, which are not seen as part of the dialogue games. The use in DE of the commitment stores is adopted from Mackenzie's (1979) DC system, while the concept was originally introduced by Hamblin (1970) in his study of fallacies. The commitment rules are as follows.

- (1) *Initial commitment, CR₀*: 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.

2.3. 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_{RESTAT}*: Mutual commitment may not be asserted unless to answer a question or a challenge. Mutual commitment refers to statements in both participants' stores.
- (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) an 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.

3. Debating strategic heuristics

One of the primary motivations behind the development of our debating system, as argued in Section 1, 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 ultimately intended to be not only a debate competitor but also an intelligent tutor. From an educational point of view, while 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 and Jones 2000) and in the playing of devil's advocate, yet both of these may be educationally valuable (Retalis, Pain and Haggith 1996). 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. How, though, should the computer treat its knowledge base? The computer is required to have the ability to argue either as a proponent or as 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 Appendix 1 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 was proposed in (Yuan 2004) and is outlined below.

In 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. It has been argued (e.g. Moore 1993) that 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 levels 1 and 2.

Levels (1) and (2) refer to strategies which apply only when the computer is facing a statement or withdrawal, since in all other cases the computer must respond to the incoming move. Level (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.

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 that 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 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 priority to a build strategy and make the computer try to open as many subtopics as possible, on the grounds that in an educational debate the aim is to expose the full complexity of the situation and this might be best achieved if the computer seeks to continue until all the knowledge base (henceforth referred as KB) has been explored. Moore also argues that the decisions at levels (1) and (2) heavily depend on the results of level (3) methods. In the current debating system, for example, the computer checks level (3) methods first; if there are level (3) methods available, 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 level (1) decision may be to switch the current focus and 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. Level (3) heuristics for each dialogue situation are given in turn below.

3.1. *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 as argued earlier. 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?”).

- (Q1) If neither P nor $\neg P$ can be found in its KB , then the computer replies with a “no commitment”.
- (Q2) If only one of them (P and $\neg P$) can be found in the KB ,
 - (Q2a) If the computer has previously uttered “no commitment” to the found statement, then it utters “no commitment” to remain consistent.
 - (Q2b) Else the computer utters the found statement.
- (Q3) If both (P , $\neg P$) are found in the computer’s KB , and assuming that one of them (say $\neg P$) supports the computer’s view and the other (say P) supports the user’s view.
 - (Q3a) If the computer has an acceptable support for $\neg P$, then utter $\neg P$.
 - (Q3b) If the computer has no acceptable support for $\neg P$, and the computer has not committed to propositions supporting P , the computer should utter “no commitment”.
 - (Q3c) If the computer has no acceptable support for $\neg P$, and the computer has committed to propositions supporting P , then the computer should utter P .

3.2. 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/her 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 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.

- (C1) If P is a modus ponens consequence of the user’s commitment, then pose a resolution demand.
- (C2) Else if there is only one acceptable support available in the KB , then state the support.
- (C3) Else if there is more than one acceptable support available, then state the one that can be further supported.
- (C4) Else if all the available acceptable supports are equally supported, then randomly choose one of the supports.
- (C5) Else if no acceptable support is available, then withdraw P .

3.3. 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. 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 $\neg P$. In this situation, the computer is required to withdraw one of them to keep consistent. Following Moore (1993), the computer should withdraw the statement which has the smaller

number of grounding statements in the commitment store at the time the resolution demand is made.

The user might invoke another type of resolution demand (i.e. resolve (Q, $Q \supset P$, why P) or resolve (Q, $Q \supset P$, “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 $Q \supset P$ or affirm P. Moore (1993) argues that the 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.

3.4. 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.

- (W1) If the computer is facing a “no commitment” to a statement supporting the user’s thesis
 - (W1a) 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.
 - (W1b) Else check whether the user retains adherence to the thesis.
- (W2) If the computer is facing a “no commitment” to a statement supporting the computer’s thesis
 - (W2a) If the non-committal statement is a modus ponens consequence of the user’s commitments, then pose a resolution demand.
 - (W2b) Else switch the focus.

3.5. 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/her 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.

- (S1a) 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.
- (S1b) Else check whether the user still adheres to his/her 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).

- (S2a) If there is an inconsistency (e.g. (P, $\neg P$)) in the user’s commitment store, then ask for resolution.
- (S2b) Else if there is a piece of hard evidence in support of $\neg P$, then state the piece of hard evidence (where a piece of hard evidence is taken to refer to a statistically or scientifically validated fact, as seen by the creator of the KB).
- (S2c) Else if there is any support of $\neg P$ and the support (say Q) can be further supported, then state $\neg P$ or state Q if $\neg P$ has been uttered, or form a plan of questions making the user accept $\neg P$.

- (S2d) Else if there is any support of $\neg P$ and the support cannot be further supported, then form a plan of questions making the user accept $\neg P$.
- (S2e) Else if P is challengeable, then challenge it.
- (S2f) Else switch the focus.

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 plans of questions in heuristics (c) and (d) are organised. Following the Walton, Reed and Macagno (2008) 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 $A \supset B$, $B \supset C$, . . . , $C \supset P$) towards 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.

- (P1) If a wanted answer is given, then carry on to execute the plan
- (P2) If a non-committal answer is given
 - (P2a) If there is an expressed inconsistency in the user's CS, then pose the appropriate resolution demand
 - (i) If the user affirms the disputed consequence, then continue the plan
 - (ii) Else abandon this line of questions
 - (P2b) Else abandon this line of questions
- (P3) If an unwanted answer (e.g. $\neg P$ rather than P) is given
 - (P3a) If there is an expressed inconsistency in the user's commitment store and the unwanted answer $\neg P$ is an element of the inconsistency, then pose the appropriate resolution demand
 - (i) If the unwanted answer is withdrawn, then continue the plan and re-pose the question.
 - (ii) Else abandon this line of question
 - (P3b) Else if the unwanted statement is challengeable, then challenge the unwanted statement
 - (i) If the unwanted answer is withdrawn, then continue the plan to re-pose the question of P .
 - (ii) Else abandon this line of question
 - (P3c) Else abandon this line of question

This, then, is the set of strategic heuristics currently adopted by our human-computer debating system. The following sections consider the evaluations of such strategy.

4. Agent-based evaluation

This section discusses an agent-based evaluation of the strategy proposed in the previous section. We first outline the construction of the two computational agents that were used for generating dialogue transcripts. We then propose criteria against which the agent-generated dialogue transcripts are analysed. We end this section with the analysis of results.

4.1. Computational agents

Two computational agents (referred to henceforth as Simon and Chris) who conduct debate with each other, operationalising the dialogue model DE and the proposed strategy, have been built by the authors using the Java programming language and deployed on the Internet (<http://staff.unak.is/not/yuan/games/simulation/dialogueSimulationSystem.htm>). The game starts with one agent (say Chris) asking the other agent (say Simon) his opinion on the controversial issue of capital punishment and adopting the opposite position to engage Simon in debate on the issue. Chris can adopt either a proponent or an opponent role. That is, if Simon chooses to support the view of “capital punishment is acceptable”, Chris will adopt the opposite view of “capital punishment is not acceptable”, and vice versa. Both agents then engage in debate on the topic of capital punishment, given these initial positions on the issue.

The agent system contains five main units: the interface unit, the dialogue unit, the planning unit, the commitment unit and the knowledge base unit. The *interface unit* provides the system’s user interface (Figure 1). It provides a dialogue history, which records the debate and commitment stores to show both players’ commitment store contents. In order to control the process of the debate, a *New game* menu item is available to start the debate, a *Pause* button is designed to temporarily stop a debate and a *Continue* button can be used to resume the dialogue. A *Save as* menu item is provided to save the dialogue history and both commitment stores as separate files for subsequent analysis.

The *dialogue unit* can be regarded as the dispatch centre of the dialogue interactions. This unit consists of a dialogue manager, an output manager and a referee. The dialogue manager controls the turn taking of the interaction and is in charge of output manager, referee, the commitment

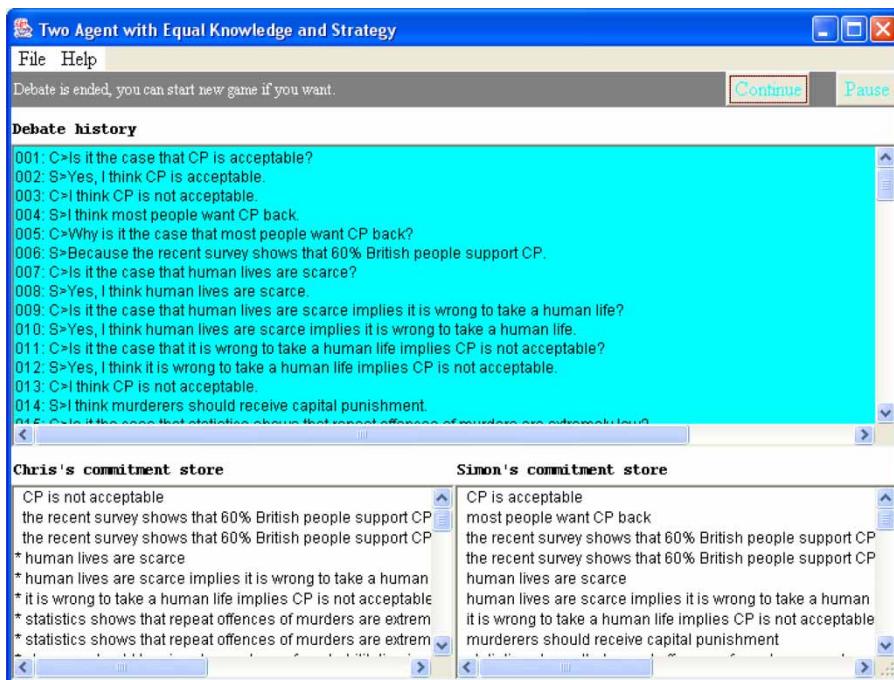


Figure 1. Computational agents user interface.

unit and the planning unit. The output manager is responsible for updating the dialogue history. The referee is responsible for enforcing the DE rule set and for the termination of the debate. The original DE regime makes no stipulation re-winning and losing, but following Moore (1993), one agent will lose the debate when it has given up its thesis or explicitly committed to the opponent's thesis. It might be possible that one agent runs out of strategy but still adheres to its thesis; under this circumstance, the agent hands over its turn to its dialogue partner. If both agents run out of strategy and a winner is still not decided, the referee will call off the game (in effect, the dialogue ends in a *stalemate*).

The *planning unit* is responsible for generating moves in the light of (i) the KB, (ii) the prevailing state of both commitment stores and (iii) the dialogue rules. Each agent is equipped with a planner. A planner manages assertion, challenge, withdrawal, resolution and question *strategists*, each of which is designed to deal with a different dialogue situation following the set of heuristics discussed in Section 3. When the agent planner receives calls from the dialogue manager, it will check the current dialogue situation and schedule the corresponding strategist to produce a move. The agent's planner will then pass the move to the dialogue unit to make the agent'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 managers will be called by the focus shift manager to check whether there are methods available to either build its own positions or attack partner's positions. The plan generator is responsible for generating a set of propositions and forming a line of questions when required by the assertion or withdrawal 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 *commitment unit* is responsible for updating both agents' commitment stores. It contains a commitment manager and two commitment stores, one for each party. The commitment manager will update both agents' commitment stores according to the DE commitment rules. Each commitment store is designed to have two lists of statements, those that have been stated and those that have been merely implicitly accepted. In order to distinguish them from each other, a statement that is only implicitly accepted is marked with an asterisk as shown in Figure 1.

The *knowledge base unit* consists of a KB manager and the KBs of the two agents. When the game starts, the dialogue manager will invoke the KB manager to initialise both agents' KBs. The agents' KBs contain a set of propositions and consequence relationships between these propositions. These relationships are based on a Toulmin-like structure (Toulmin 1958). The domain knowledge is formalised from Moore's (1993) experimental study of DC with human participants, and an example of such a knowledge base (the one used in the current study) can be seen in Appendix 1.

4.2. *Experimental set-up and evaluative criteria*

This, then, is our agent-based system. In this section, we discuss an experiment we conducted to investigate the system and its strategies. There are two component variables in the experiment: the strategy component and the KB component. The strategy component of an agent can adopt the proposed strategy or a random strategy. The KB components for both agents can be the same set or a different set. In order to evaluate the strategies discussed earlier, the agent system was run under three conditions.

- (1) One of the agents adopts the strategy and the other uses random argument, and both have the same KB.
- (2) Both agents adopt the same strategy and share the same KB.
- (3) Both agents adopt the same strategy. One of the agents' KB is a subset of the other.

It is anticipated that using random argument (1) might reveal certain failures of the heuristics (e.g. unexpected new situations) that might be overlooked by manual use of them. Conditions (2) and (3) may reveal whether a high-quality dialogue is generated if both agents adopt the proposed strategy. Condition (3) may also be used to see whether an agent with a smaller KB might turn out to be the loser of the debate given that both agents share the same dialogue strategy.

Three dialogue examples (DE4, DE5 and DE6 – full transcripts can be found from Appendix 2)¹ have been generated under conditions (1), (2) and (3), respectively, for analysis. A prerequisite of such an analysis is a set of evaluation criteria that are independent of the dialogue model and the debating heuristics. Five criteria for evaluation are proposed, based in part on the work by Moore (1993) and Walton (1984). These criteria are outlined as below.

- (1) *Robustness*. The issue here is whether all dialogue situations are dealt with by the proposed strategy. In particular, this concerns whether there are unexpected dialogue situations which have not been considered in the strategy.
- (2) *Equal opportunity*. The issue here is whether each agent has equal opportunity to advocate their point of view. In particular, this criterion concerns whether there are frequent initiative shifts in the process of dialogue, such that the resulting debate transcripts can be described as *mixed initiative* dialogue. A similar point, i.e. fairness, is made by McBurney, Parsons and Wooldridge (2002).
- (3) *Coverage of issues*. The interest here is whether the knowledge in the KB is well revealed and discussed. One of the potential applications of the debate system is to broaden the interaction style of computer-based learning systems (cf. Moore 1993, Yuan, Moore and Grierson 2008). In such an educational setting, it might be expected that the system encourages students to look at an issue from different perspectives, and therefore it would be hoped that as many issues as possible are raised.
- (4) *Argument flow*. The issue here concerns whether the dialogues produced are unreasonably disjointed. If the strategic agents' dialogue contributions are clearly related to its dialogue partner's previous utterances, then the flow of the developing argument can be deemed acceptable. A similar point, i.e. dialectical relevance is made by a number of researchers (e.g. Hitchcock 1992; Walton 1999; Johnson 2000; Maudet 2001; Prakken 2001; Rehg, McBurney and Parsons 2005).
- (5) *Defeatibility*. This concerns whether the strategy is making the agent *too* wise to be beaten, and thus leading to difficulties where a human is one of the participants in the dialogue (cf. Walton 1984). In an educational setting, we would not want students to be frustrated by being constantly defeated by the computer. A useful debating system, that is, should be able to reasonably lose a dialogue as well as to win, and thus avoid demoralising its human interlocutors (Moore 1993).

The agent-generated dialogue transcripts have been analysed against the proposed criteria. Results are given next.

4.3. Analysis of the results

This section contains the results of an analysis of the agent-generated dialogue examples DE4, DE5 and DE6. During the analysis, each utterance of the dialogues is considered, in turn, via the

addition of appropriate annotations in square brackets. The heuristic being invoked is indicated at the beginning of the annotations. For example:

- 044: S>Why is it the case that CP is not acceptable?
[S2e – S challenges for further reasons]
- 045: C>Because innocent people may get killed.
[C3 – C gives reasons supporting its thesis]

This approach to the analysis makes it possible to examine the data under the evaluative criteria discussed in the previous section, and thus to assess whether the proposed strategy can provide adequate services to enable the computer to act as a dialogue participant and produce good dialogue contributions.

4.3.1. Evaluation criterion 1: robustness

The interest here is whether all dialogue situations generated by the agents are successfully dealt with by the proposed strategy. In total, the agent runs have generated 133 dialogue situations for the strategic agents to deal with. Among them 74 are *assertions*, 11 *withdrawals*, 37 *questions*, 11 *challenges* and zero *resolution demands*. How various strategists deal with these dialogue situations are analysed in turn below.

4.3.1.1 *Assertion strategist.* Seventy-four assertions were generated and are summarised in Table 1. Their move contents are classified into six categories: move maker’s thesis, statements supporting the move maker’s thesis, opponent’s thesis, statements supporting the opponent’s thesis, statements handing over its turn and unrecognised statements. The right most column of the table contains row number to facilitate the discussion. The second right column of the table includes a summary of each strategic response to these dialogue situations.

The 18 instances of mover maker’s thesis and 25 instances of supporting the move maker’s thesis are considered as statements standing on the side of the move maker, and it might be expected that the strategic agent should attack them. It can be seen from Table 1 that the assertion strategist does provide various means of either attacking opponent’s view or building its own view, with seven exceptions of giving up this opportunity (four occasions summarised at row 5 and three occasions at row 12 in Table 1). In these circumstances, the assertion strategist runs out of methods and therefore hands over its turn to the opponent. This might be seen as reasonable since the opponent may have something more to say, but on the other hand, more sophisticated means might be needed if the strategic agent constantly runs out of methods and therefore overuses the handover strategy.

For statements standing on the side of the opponent, the assertion strategist is expected to use them rather than to attack them (cf. Walton 1998). It is shown at rows 13–16 of Table 1 that the assertion strategy does provide some means of handling this, e.g. using the strategy to build its thesis or continuing its plan execution or checking the opponent’s thesis adherence. On the occasion of DE6-053, the game ends since the move maker has committed to the opponent’s thesis. It is interesting to see that on five occasions, the assertion strategist has to decide what to do when its dialogue partner hands over its turn. This is not specified in the heuristics. The current implementation provides some mechanisms for this. On one of the occasions (at row 18 of Table 1), the referee calls off the game since both parties have run out of methods. Concerning the remaining four instances (summarised at row 17 of Table 1), the assertion strategist checks its opponent’s thesis adherence.

There are four unrecognised statements generated in DE6 since one agent has only a subset of the KB of the other. The strategic agent currently responds with a question to check partner’s

Table 1. Dialogue situation 1 (response to statement): summary.

Dialogue situation: assertion (74)			
Move content	Turn	Strategic response	Row no.
Move maker's thesis (18)	DE4-002, DE5-002, DE6-002	Adopt the opposite view	1
	DE4-016, DE4-024, DE5-034, DE6-006, DE6-021, DE6-031	Build its thesis	2
	DE5-003, DE5-046, DE6-003	Issue a direct thesis support	3
	DE6-014, DE6-043	Challenge	4
	DE5-043, DE6-018, DE6-028, DE6-040	Hand over the turn	5
Supporting move maker's view (25)	DE5-004	Issue contradictory evidence	6
	DE5-020, DE5-047	Issue an objection	7
	DE4-052, DE5-009, DE5-014, DE6-009	Demolish plan	8
	DE4-036, DE5-006, DE5-021, DE5-025, DE5-048, DE6-024, DE6-045, DE6-047	Challenge	9
	DE5-005, DE5-008, DE5-019, DE6-008	Switch the current focus and issue a direct thesis support	10
	DE4-026, DE5-023, DE5-027	Switch the current focus and build its thesis	11
	DE4-042, DE4-054, DE5-050	Hand over the turn	12
Opponent's thesis (1)	DE6-053	End the game	13
Supporting the opponent's thesis (21)	DE4-012, DE4-038, DE4-040, DE4-048	Use to build its thesis	14
	DE5-011, DE5-013, DE5-016, DE5-018, DE5-029, DE5-031, DE5-033, DE5-036, DE5-038, DE5-040, DE5-042, DE6-033, DE6-035, DE6-037, DE6-039	Continue its plan execution	15
	DE4-056, DE4-022	Check partner's thesis adherence	16
Handing over turn (5)	DE5-044, DE6-019, DE6-029, DE6-041	Check partner's thesis adherence	17
	DE5-051	End the game	18
Unrecognised statements (4)	DE6-004, DE6-012, DE6-016, DE6-026	Check partner's thesis adherence	19

thesis adherence. There might be a need for more sophisticated means to handle this kind of situation, particularly in a human–computer debate setting where users are allowed to input fresh propositions.

Generally speaking, the assertion strategist successfully handles most dialogue situations except unrecognised statements and the situation of running out of moves.

4.3.1.2 Withdrawal strategist. Eleven withdrawals (or “no commitment”) are present in the transcripts. They are categorised as follows: withdrawal of move maker's thesis, withdrawal of a statement supporting the move maker's view and withdrawal of a statement supporting the opponent's view. These are summarised in Table 2 and discussed in turn below.

On one occasion (at row 1 of Table 2), the move maker is withdrawing its thesis. The game is therefore ended since the move maker has given up its view. On 2 occasions (at rows 2 and 3

Table 2. Dialogue situation 2 (response to withdrawal): summary.

Dialogue situation: withdrawal (11)			
Withdrawn statement	Turn	Strategic response	Row no.
Move maker's thesis (1)	DE4-058	End the game	1
Supporting move maker's thesis (2)	DE6-049	Further challenge	2
	DE6-051	Check partner's thesis adherence	3
Supporting opponent's thesis (8)	DE4-014, DE4-018, DE4-028, DE6-011, DE6-023	Drop the plan and build your own thesis.	4
	DE4-030, DE4-050, DE4-044	Hand over the turn	5

of Table 2), the move maker is withdrawing statements supporting its thesis. The response of the strategic agent is to challenge the statement supported by the withdrawn statement, or assess whether the dialogue partner still adheres to his thesis.

There are eight instances in which an agent replies "no commitment" to statements that support its opponent's thesis. On five occasions (at row 4 of Table 2), the withdrawal strategist deals with this by starting another line of argument. However, for the remaining three instances (at row 5 of Table 2), the withdrawal strategist fails to do so. The explanation here is that the withdrawal strategist has run out of methods, and therefore hands over its turn. This needs further consideration if the strategic agent constantly faces this kind of situation.

Given the above analysis, the withdrawal strategist seems to be working satisfactorily with the exception of needing more sophisticated strategies when running out of moves.

4.3.1.3 Challenge strategist. There are 11 challenges generated. It is shown in Table 3 that on nine occasions (at row 1 of Table 3), the challenge strategist provides a suitable ground following its knowledge structure. There are two occasions (DE6-048 and DE6-050 at row 2 of Table 3), on which the challenge strategist gives a non-committal answer. Concerning the first of these, the strategic agent cannot find a support for the statement in its KB and therefore responds with a non-committal answer. Regarding the latter, the strategic agent does have a support in its KB for the statement being challenged; however, the support is not an acceptable ground since the partner of the strategic agent has challenged the support and the strategic agent had withdrawn this support from its commitment store during the earlier stage of dialogue. The strategic agent, then, would beg the question were it to answer the challenge with this unacceptable support (cf. Yuan et al. 2003). It is therefore reasonable for the challenge strategist to give a non-committal answer rather than to commit the fallacy of question begging.

Table 3. Dialogue situation 3 (response to challenge): summary.

Dialogue situation: challenge (11)			
Challenged statement	Turn	Strategic response	Row no.
Opponent's thesis or statement supporting opponent's thesis (11)	DE4-046, DE5-007, DE5-022, DE5-026, DE5-049, DE6-015, DE6-025, DE6-044, DE6-046	Give a suitable ground	1
	DE6-048, DE6-050	No commitment	2

Table 4. Dialogue situation 4 (response to question): summary.

Dialogue situation: question (37)			
Questioned statement	Turn	Strategic response	Row no.
Game start (2)	DE5-001, DE6-001	Give a positive answer	1
Supporting move maker's view (24)	DE4-006, DE4-010, DE4-032, DE5-010, DE5-012, DE5-015, DE5-017, DE5-028, DE5-030, DE5-032, DE5-035, DE5-037, DE5-039, DE5-041, DE6-032, DE6-034, DE6-036, DE6-038	Give a positive answer	2
	DE4-004, DE5-024, DE6-007	Give a negative answer	3
	DE4-034, DE6-010, DE6-022	No commitment	4
Opponent's thesis (9)	DE5-045, DE6-005, DE6-013, DE6-017, DE6-020, DE6-027, DE6-030, DE6-042	Give a positive answer	5
	DE6-052	Give a negative answer	6
Supporting opponent's thesis (2)	DE4-008, DE4-020	Give a positive answer	7

In sum, then, the challenge strategist seems to successfully deal with all dialogue situations in this category.

4.3.1.4. Question strategist. In total, 37 questions are generated and summarised in Table 4. They fall into four categories according to the nature of their move contents: game start, move content supporting move maker's thesis, opponent's thesis, move content supporting opponent's thesis. These are in turn discussed below.

Only the strategic agent in DE5 and DE6 needs to respond to the game starting question (at row 1 of Table 4). In DE4, it is the random agent responding and therefore not in need of analysis. Heuristics for responding to the starting question are not specified in the current strategy. The current implementation uses a random approach and agent choosing either view would be considered as reasonable.

The situation of a proposition being questioned that supports the questioner's view (24 instances at rows 2–4 of Table 4) might be interpreted as the questioner asking the opponent to commit to that proposition, according to (Walton and Krabbe 1995). The opponent might be expected not to commit to it if it has an alternative because the opponent knows it is a proposition not on his side. It is shown in Table 4 that the strategic agents respond 18 times with a *Yes*, 3 times with a *No* and 3 times with a “no commitment”. Each of the *Yes* responses is accounted for as the only answer found in its KB, and the opponent speaking the truth. For the *No* responses, the opponent finds the negative answer (which mitigates against the opponent's view) from its KB, and the answer can be further supported, the opponent therefore gives the negative answer. For the “no commitment” responses, there are three instances (DE4-034, DE6-010 and DE6-022); on the first situation, the opponent does have a negative answer (which mitigates against the opponent's view) in its KB, but the opponent finds no further support for the negative answer if being challenged, giving a positive response might harm its own positions, the opponent therefore issues a “no commitment” response; on the latter two situations, the opponent does not recognise the questioned statement (it cannot be found in its KB), and therefore answers with a “no commitment”.

There are nine instances (at rows 5 and 6 of Table 4) that can be seen as one agent merely checking whether the opponent still adheres to its thesis. On eight occasions, the opponent insists on its view since it still has acceptable grounds for its view. On one occasion (DE6-052), the opponent accepts the opposite view. The dialogue fragment for this instance is reproduced below.

- 032: S>Is it the case that nobody is willing to die?
- 033: C>Yes, I think nobody is willing to die.
- 034: S>Is it the case that “nobody is willing to die” is a prima facie reason for “CP makes people less likely to commit serious crimes”?
- 035: C>Yes, I think “nobody is willing to die” is a prima facie reason for “CP makes people less likely to commit serious crimes”.
- 036: S>Is it the case that “CP makes people less likely commit serious crimes” is a prima facie reason for “CP is a good deterrent”?
- 037: C>Yes, I think “CP makes people less likely commit serious crimes” is a prima facie reason for “CP is a good deterrent”.
- 038: S>Is it the case that “CP is a good deterrent” is a prima facie reason for “CP is acceptable”?
- 039: C>Yes, I think “CP is a good deterrent” is a prima facie reason for “CP is acceptable”.
- 040: S>I think CP is acceptable.
-
- 051: C>I don’t know why innocent people may get killed.
- 052: S>Is it the case that CP is not acceptable?
- 053: C>No, I think CP is acceptable.

In the above dialogue fragment, the opponent (agent Chris) has no acceptable ground for its thesis since its support has been withdrawn in turn 051. Further, agent C has explicitly committed to the set of propositions and conditionals which implies its dialogue partner’s thesis at turn 32–29. Agent C therefore makes a concession and accepts the opposite view in turn 053, thus losing the debate. The nine instances of questions involving thesis adherence checking can be seen as being reasonably answered, given the above analysis.

It is interesting to see that two questions of statements supporting the opponent’s thesis (at row 7 of Table 4) were generated. As expected, the strategic agent takes advantage of this and gives positive responses.

Overall, most dialogue situations that can occur in DE are arguably successfully handled, and the strategy can largely be regarded as satisfying the *robustness* criterion. Certain dialogue situations, i.e. unrecognised statements and the situation of running out of moves, however, do need more sophisticated heuristics.

4.3.2. Evaluation criterion 2: equal opportunity

Of concern here is the issue of initiative. Initiative is relevant because if one dialogue participant is constantly starved of the initiative, he/she cannot fully or freely advocate his/her point of view (cf. Walton 1989; Moore 1993, p. 229).

In DE4, the strategic agent hands over its initiative nine times to the random agent during the 54-turn dialogue. There are seven instances of initiative shift during the 52-turn dialogue in DE5, and there are four instances of initiative shift during the 54-turn dialogue in DE6. The longest duration of one agent retaining the initiative is from turn DE5-022 to turn DE5-034, in that the agent made two challenges, four questions and one statement to hand over the initiative.

On the whole, both agents have had opportunities to express their point of view, and the strategy therefore appears to satisfy the *equal opportunity* criterion.

4.3.3. Evaluation criterion 3: coverage of issues

Of interest here is whether points implemented in the KB are revealed and discussed. The KB used by the agents discussed in this paper (partly shown in Appendix 1) can be seen as containing three subtopics (deterrent issue, popularity issue and moral issue) which support the view of “CP is acceptable” and two subtopics (moral issue and consequence issue) which support the contrary view.

It can be seen that all the issues in the KB are raised during the evolving dialogue. Further, these issues are discussed to the maximum depth possible given the KB in the DE5 and DE6 dialogues (DE4 is not considered here since one of the agents uses random argument).

In sum, the dialogue generated by the strategy is acceptable in respect of the *coverage of issues* criterion.

4.3.4. Evaluation criterion 4: argument flow

The analysis here will consider whether the arguments generated by the agents as they follow the strategy are flowing well. If participants’ contributions are clearly related to each other or appear logical, then, the flow will be deemed acceptable (cf. Moore 1993).

In order to inspect the protocols for disjointedness, the process of analysis adopted includes a manual insertion of missing premises of incomplete arguments into the machine generated transcripts; this approach is in line with that adopted by (Walton and Reed 2002). An example of this is the following extract from DE5 (in which presumed missing premises are indicated in italics).

003: C>I think CP is not acceptable.

004: S>I think CP is a good deterrent, and “*CP is a good deterrent*” is a *prima facie* reason for “*CP is acceptable*”.

By inserting the missing premises, it is easy to see the argument flow in this example – agent S is addressing agent C’s utterance.

In DE4, DE5 and DE6, the strategic agents made 133 dialogue contributions. Ninety-three of them involve directly addressing the previous utterance. There are 11 instances of the strategic agent’s contributions which involve handing over its turn and asking the dialogue partner to continue its line of argument. There are seven instances of checking whether the dialogue partner still adheres to its thesis given its previous statement. All these can be seen as engendering reasonable argument flow. However, there are 22 apparent instances of disjointedness, 10 of which involve building a line of questions toward the negation of the dialogue partner’s previous statement and 12 of which involve a shift of the current focus. The apparent examples of disjointedness might in fact be reasonable in terms of the underlying strategy of the move maker. Consider, for example, the following extract from DE5.

014: S>I think it is not always wrong to take a human life.

015: C>Is it the case that human lives are scarce?

016: S>Yes, I think human lives are scarce.

017: C>Is it the case that “human lives are scarce” is a *prima facie* reason for “it is wrong to take a human life”?

018: S>Yes, I think “human lives are scarce” is a *prima facie* reason for “it is wrong to take a human life”.

019: C>I think it is wrong to take a human life.

020: S>I think murderers should receive capital punishment.

In the above dialogue fragment, turn 015 might appear at first sight to be disjointed from the previous utterance. However, turn 015 can be seen as agent C starting its distance plan to build the case for “it is wrong to take a human life” (culminating in turn 019) in order to rebut agent S’s view “it is not always wrong to take a human life”. Turn 20 might be seen as abrupt shift, but is perhaps, justified as opening a new focus (punishment). Part of the reason for the feeling of abruptness is the lack in DE (indeed all dialectic models) of linguistic means of introducing transitions between foci.

Generally speaking, though, the arguments can be seen as flowing well – the total of 22 instances of disjointedness represents a relatively small proportion of the 133 strategic agents’ dialogue contributions. The apparent examples of disjointedness are arguably reasonable in terms of the underlying strategy of the move maker with the exception of the absence of explicit linking for a focus shift.

4.3.5. Evaluation criterion 5: defeatability

The interest here is whether the agent adopting the strategy is defeatable. In DE4, the strategic agent wins over the random agent. In DE5, the two agents with the same strategy and same KBs end up in a stalemate. In DE6, the strategic agent with a subset of the KB of the other loses the game. The proposed strategy seems more intelligent than a random strategy given its winning result in DE4. On the other hand, the strategy is defeatable as shown in DE6 in which the strategic agent Chris does lose the game and does so in a manner which might be considered reasonable, as opposed to a mere *surrender*. In the process of arriving at this defeat, agent Chris’s thesis supports are removed by agent Simon, agent Simon asserts the *prima facie* reasons for its thesis in turn DE6-004, DE6-012, DE6-016 and DE6-024, and agent Chris is therefore implicitly committed to them.

The evidence therefore suggests that an agent adopting the strategy can both win and lose a game in an artificial setting. However, the fact that the agent can be defeated in the special case when the KB is a subset of the other participant KB is not sufficient to prove the defeatability criteria in the more general case of human–agent argumentation. The defeatability criteria will be further discussed in the user-based evaluation section below.

To summarise the agent-based evaluation, the qualitative assessment suggests that, generally speaking, the proposed strategy can provide good services enabling the computer to act as a dialogue participant. The assessment also reveals several weaknesses in respect to the *robustness* criterion, e.g. the absence of heuristics to handle the situations like the game start, unrecognised statements and running out of methods. The assessment also reveals a potential weakness with respect to the argument flow criterion where the strategic agent sometimes switches focus without giving a clear indication of this. Current work involves amending the strategy to cater for these concerns.

5. User-based evaluation

A prerequisite for a user-based evaluation is to construct a human–computer debating system. A fully functional human–computer debate prototype, operationalising the dialogue model DE and the proposed strategy (currently able to debate the issue of capital punishment) has been built and deployed on the Internet (<http://staff.unak.is/not/yuan/games/debate/system.htm>). An example user system interface can be seen in Figure 2. The user system interface provides a debate history to record the debate for subsequent analysis, commitment stores for both the user and the computer and input facilities for the user to make a move. 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

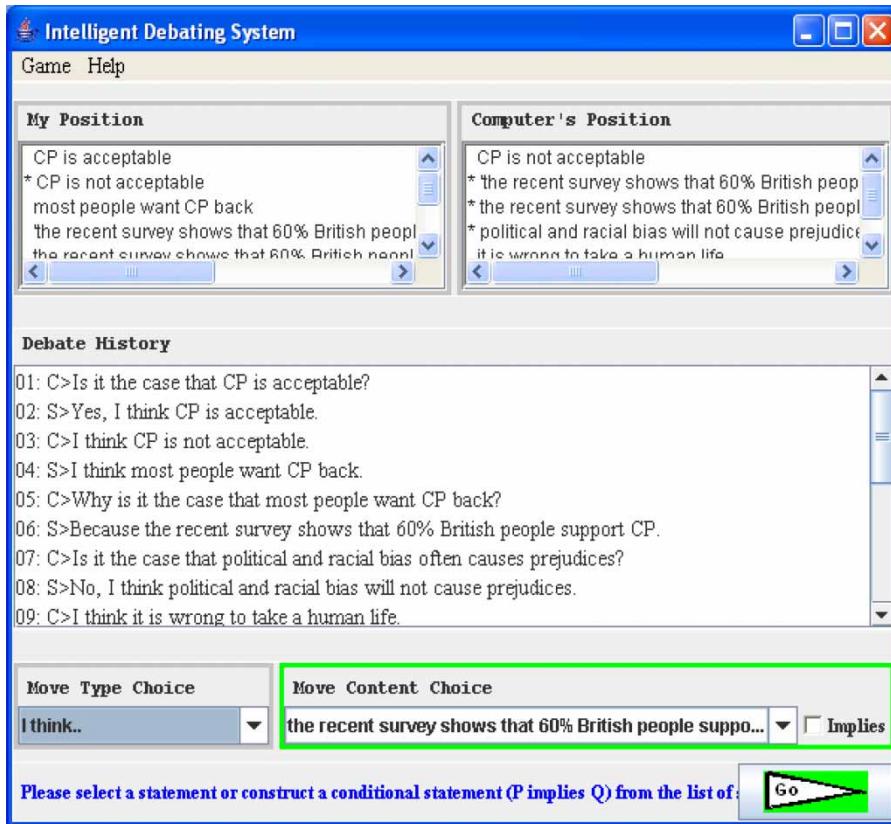


Figure 2. Human–computer debating system user interface.

have been merely implicitly accepted. In the current system, a statement that is only implicitly accepted is marked with an asterisk, as shown in Figure 2. The commitment stores are updated during the debate according to the DE commitment rules.

Turning to the user input facilities, a menu-based approach is adopted. Under this approach, the user needs to make a double selection, choosing from the available move types and then from the list of prescribed propositions from the domain under discussion. To prevent the user from constantly breaking the rules and to increase the learnability of the system, only the legally available move types are provided by the system before the student makes a move. 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. PQ). 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.

The system enables the user (S) to conduct a debate with it on the controversial issue of capital punishment. The computer (C) 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 of “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. Further details of the system can be found from (Yuan, Moore and Grierson 2007a).

Two types of evaluations have been carried out: expert evaluation and user-based evaluation. The aim was to assess how acceptable, usable and potentially valuable this innovation was, prior to greater exploitation and subsequent further evaluations of its educational value. 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 and 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.

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” and “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 rather uncomfortable when the computer constantly hands over its turn after a period of debate. She further suggested 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 fails to 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 documented in (Ævarsson 2006). Ten university students from the University of Akureyri, Iceland, who are interested in debate participated in the study. 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 researcher briefing the participants about the purpose and procedures of the study. In the debate session, the participant was asked to conduct debate with the computer for 15 min in the pilot studies; this was extended to 20 min 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 summarised in Table 5.

Table 5. Summary of dialogue transcripts.

Participant No.	Number of turns	Number of rule breaks	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 10 participants successfully conducted debate with the system without difficulty. In total, 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, eight of the debates ended with a stalemate and none of the participants made the computer concede.

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 the debates with the system, and they particularly liked the non-deterministic 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 debate with the system again were it available on the Internet. The value of the system 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 commitment store, 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 programme rather than the current three-window set-up. 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 an * prefix in the commitment stores.

To address these user interface issues, the user interface of the system is being redesigned 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: available propositions, user commitment store and the computer commitment store. Thus the user interface becomes a one-window set-up, 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.

6. Discussion

We have outlined our research aimed at the development and evaluation of suitable strategies for an interactive system offering dialogue involving competitive debate. Before discussing the significance of the evaluations reported above, there are three possible difficulties with the methodology that ought to be discussed. The first difficulty concerns the general approach of this study. Dynamic testing has been used (as opposed to a static analytical approach). For the dynamic approach, by its very nature, completeness is not possible (Dijkstra 1972). This nature is reflected in this study, for example, no *resolution demand* dialogue situations are generated for analysis, and the experimental set-up of the agent configurations does not consider the situation where both agents have completely different KBs. Dynamic testing, however, shows one step forward than static analysis by enabling us to assess the quality of the dialogue generated by the strategy. Any concern of the limitation of dynamic testing has been alleviated to some extent by (1) the combination of the static analysis of the strategies when they were developed in (Yuan 2004) and in (Moore 1993) and (2) the representative data generated for this study.

Second, it might be argued that only a small number of dialogue transcripts (three in total from three pairs of agents) have been generated for analysis. However, this study is intended not as a statistical enquiry, but rather as an investigation into the detail of the argument generated by the strategy. Further, 165 utterances are generated (DE4: 59; DE5: 52; DE6: 54). Each utterance needs to be considered in depth, and as a result this study does, it is held, provide sufficient data for the purpose of this assessment of debating strategies.

The third difficulty may be that there is a heavy reliance on judgements of quality by the author of the heuristics and the systems and that the criteria of quality are themselves intuitively formulated. The judgement issue may be endemic to the field, and similar criticisms could perhaps be levelled against much of the dialectics literature (Moore 1993). Further, computationally generating dialogues from dialectical theories may represent a step forward, and making the various criteria clear and explicit may well localise the issues to relatively narrow concerns at any one time, and this may detract from the judgemental element. In addition, these criteria have enabled us to provide a thorough analysis of the data collected, and to leave the results, and the data itself, available for independent inspection.

We argue, then, that the methodology adopted is sound. We believe that the work reported makes a valuable contribution to the fields of human–computer dialogue in general, and of strategies in dialectical systems in particular. Concerning the latter, we have proposed a set of strategies to be utilised with the dialectical system DE and provide a means to assess the appropriateness of the strategy. Further, since the agent systems and the human–computer system we have built can potentially be adapted to function with a different set of strategies, it potentially provides people working in the field of dialectical strategies with a test bed within which they can experiment with new strategies they develop (Maudet and Moore 2001; Amgoud and Maudet 2002).

Considerable research effort has been devoted to the formulation of strategies for dialectical systems. Grasso et al. (2000), for example, adopt, for their nutritional advice-giving system, schemas derived from Perelman and Olbrechts-Tyteca's (1969) *New Rhetoric*, and Ravenscroft and Pilkington (2000) have "a repertoire of legitimate tactics available for addressing common conceptual difficulties" (p. 283) that can be selected based on a strategic pedagogy developed through modelling effective tutor behaviour (Ravenscroft 2000). Amgoud and Maudet (2002) suggest *meta-preferences*, such as "choose the smallest argument", to drive the choice of move, and Freeman and Farley (1996) delineate ordering heuristics as guidelines for selecting argument moves. Prakken (2001, 2005) suggests, in line with his dialogue framework, that the dialogue focus and dialectical relevance can be maintained by restricting dialogue participants to replying and backtracking to the previous move. Yuan, Svansson, Moore and Grierson (2007b) implement a probability-utility-based strategy, which enables a software agent to compute all the possible

dialogue sequences and then select a legal move with the highest probability of winning an abstract argumentation game. Similarly, Oren, Norman and Preece (2006a) present a strategy for argumentation based on the principle of revealing as little information as possible, and Oren, Norman and Preece (2006b) discuss strategic possibilities when an agent's information varies in confidentiality. Noting that information might be so highly confidential that the cost of revealing it in the course of an argumentation dialogue would outweigh the benefit accrued from winning the dialogue, the authors present a heuristic to help an agent take account of such confidentiality-related costs in argumentation.

Our work adds to this body of research and is unique in that it concerns specifically strategies for computer-based debate. More generally, our work contributes to the field of human-computer dialogue. We have proposed a set of strategies for an educational human-computer debating system and our evaluations of the strategy in use reveal that the strategy is able to provide a good service for a computer to act as a dialogue participant. Our 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. Further, given the usefulness of a dialectical approach to interactive computer systems (Moore 1993; Moore and Hobbs 1996; Yuan 2004; Ravenscroft, Wegerif and Hartley 2007), any development of dialectical strategy per se potentially has a pay-off in terms of human-computer dialogue.

7. Conclusion and further work

We have proposed a set of strategic heuristics for a human computer debating system regulated by the dialogue game DE. An agent-based dialogue system and a human-computer debating system have been constructed to facilitate the evaluation of the proposed strategy. Both agent-based and user-based evaluations of the strategy in use have been carried out. The evaluations provide evidence that the proposed strategy can provide good service for a computer to act as a dialogue participant. The results are essentially favourable in demonstrating the innovation's acceptability and usability. The evaluations also provide evidence for the educational and entertainment value of the system. Several weaknesses of the dialogue strategy and user interface are revealed and discussed, and our immediate further work is to address these concerns.

There is also a variety of additional interesting ways in which the research can be carried forward. A limitation of the current study is that it is restricted to the consideration of a single strategy. The comparison is with moves generated in a random manner. This was necessitated by the absence of alternative strategies. When there is a refined version of the strategy available in the future, further comparative experimental studies of the strategy can be conducted, e.g. by comparing the current set and the refined set. A similar weakness is that our evaluative criteria need to be judgementally applied. Further work will involve refinement of the criteria, in line perhaps with Amgoud and Dupin de Saint-Cyr (2008).

Turning to our basic human-computer debating system, it can be enhanced to allow the user to question or challenge a conjunction of statements (e.g. $P \wedge Q$) or conditional. Currently, the DE system is perhaps over-zealous at 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 and Grierson (1996), can also be implemented. 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 are also planning to permit free user input for the debating system, 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.

A further problem with our basic debating system, and indeed the underlying DE model, is that the move set, and in particular the question move, is restricted to bi-polar questions. Specifically, the absence of a move which enables debate participants to seek explanation of the substantive points being made, assumes that each participant fully understands all the points. In an educational context in particular, this is likely to be an undesirable restriction. Given this, our current research is seeking to extend the dialogue beyond debate per se and Naim et al. (2009) have addressed this.

The dialogue model and the strategy are currently represented informally using structured English and are hard-wired into the program logic by the developer of the system. Ideally, the dialogue model, strategy or even the knowledge base of the system could be written and modified by the users of the system and then directly translated to the program code. A formal representation is required for this. This is a general software engineering issue in the field of computational dialectics. A plausible investigation might be along the lines of (Bench-Capon, Geldard and Leng 2000; Reed 2006; Yuan, Moore, Reed, Ravenscroft and Maudet 2011) where each move act is specified using a pre- and post-condition pair.

Note

1. DE1, DE2 and DE3 were generated for a separate study of the dialogue model DE in (Yuan 2004).

References

- Amgoud, L., and Dupin de Saint-Cyr., F. (2008), 'Measures for Persuasion Dialogs: A Preliminary Investigation', in *2nd International Conference on Computational Models of Argument (COMMA'08)*, Toulouse, France, 28–30 May 2008, pp. 13–24.
- Amgoud, L., and Maudet, N. (2002), 'Strategical Considerations for Argumentative Agents', *NMR'2002 Special Session on Argument, Dialogue, and Decision*, Toulouse, France.
- Amgoud, L., and Parsons, S. (2001), 'Agent Dialogue with Conflicting Preferences', in *Proceedings of the 8th International Workshop on Agent Theories Architectures and Languages*, Seattle.
- Bench-Capon, T.J.M. (1998), 'Specification and Implementation of Toulmin Dialogue Game', in *Proceedings of the 11th International Conference on Legal Knowledge Based Systems (JURIX)*, Nijmegen: Gerard Noodt Institute (GNI), pp. 5–20.
- Bench-Capon, T.J.M., Geldard, T., and Leng, P.H. (2000), 'A Method for the Computational Modelling of Dialectical Argument with Dialogue Games', *Artificial Intelligence and Law*, 8, 233–254.
- Dijkstra, E.W. (1972), 'Notes on Structured Programming', in *Structured Programming*, ed. Dahl, O.J., Dijkstra, E.W., Hoare, C.A.R., London: Academic Press, 1972.
- Dix, A., Finlay, J., Abowd, G., and Beale, R. (2004), *Human-Computer Interaction*, (3rd ed.), New Jersey: Prentice-Hall.
- Ævarsson, K. (2006), 'Human-Computer Debating System, Evaluation and Development', unpublished BSc. dissertation, University of Akureyri, Iceland.

- Freeman, K., and Farley, A. (1996), 'A Model of Argumentation and Its Application to Legal Reasoning', *Artificial Intelligence and Law*, 4, 163–197.
- Freire, P. (2000), *Pedagogy of the Oppressed* (reprinted), New York: Continuum.
- Grasso, F., Cawsey, A., and 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.
- Hamblin, C. (1970), *Fallacies*, London: Methuen.
- Hitchcock, D. (1992), 'Relevance', *Argumentation* 6, 251–270.
- Johnson, R. (2000), *Manifest Rationality: A Pragmatic Theory Of Argument*, Mahwah, NJ: Lawrence Erlbaum Associates.
- Mackenzie, J. D. (1979), 'Question-Begging in Non-Cumulative Systems', *Journal of Philosophical Logic*, 8, 117–133.
- Maudet, N. (2001), 'Notes on Relevance in Dialectical Systems', in *Proceeding of the ECSQARU'2001 Workshop Adventures in Argumentation*, Toulouse, pp. 36–42.
- Maudet, N., and Moore, D. (2001), 'Dialogue games as Dialogue Models for Interacting with, and via, Computers', *Informal Logic*, 21(3), 219–243.
- McBurney, P., Parsons, S., and Wooldridge, M. (2002), 'Desiderata for agent argumentation protocols', in *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2002)*, Bologna, Italy, pp. 402–409.
- 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., and Hobbs, D.J. (1996), 'Computational Use of Philosophical Dialogue Theories', *Informal Logic*, 18 (2), 131–163.
- Naim, R., Moore, D., Yuan, T., Dixon, M., and Grierson, A. (2009), 'Computational Dialectics for Computer Based Learning', in *Proceedings of the 2009 International Conference on the Current Trends in Information Technology (CTIT'09)*, Dubai.
- Oren, N., Norman, T., and Preece, A. (2006a), 'Loose Lips Sink Ships: A Heuristic for Argumentation', in *Proceedings of the Third International Workshop on Argumentation in Multi-Agent Systems*, Hakodate, Japan.
- Oren, N., Norman, T., and Preece, A. (2006b), 'Arguing with Confidential Information', in *Proceedings of the 2006 European Conference on Artificial Intelligence (ECAI-06)*, Riva del Garda, Italy.
- Perelman, C., and Olbrechts-Tyteca, L. (1969), *The New Rhetoric: a Treatise on Argumentation*, Notre Dame, IN, USA: University of Notre Dame Press.
- Pilkington, R.M., and Grierson, A. (1996), 'Generating Explanations in a Simulation-Based Learning Environment', *International Journal of Human-Computer Studies*, 45 (5), 527–551.
- Pilkington, R.M., and Mallen, C. (1996), 'Dialogue Games to Support Reasoning and Reflection in Diagnostic Tasks', in *Proceedings of the European Conference on Artificial Intelligence in Education*, ed. P. Brna, A. Paiva, and J. Self, Lisbon, Portugal: Fundacao Calouste Gulbenkian.
- Prakken, H. (2001), 'Relating Protocols for Dynamic Dispute with Logics for Defeasible Argumentation', *Synthese, Special Issue on New Perspectives in Dialogical Logic*, 127, 187–219.
- Prakken, H. (2005), 'Coherence and Flexibility in Dialogue Games for Argumentation', *Journal of Logic and Computation*, 15, 1009–1040.
- Rahwan, I., McBurney, P., and Sonenberg, E. (2003), 'Towards A Theory of Negotiation Strategy (A Preliminary Report)', in *Proceedings of an AAMAS-2003 Workshop*, Melbourne, Australia, pp. 73–80.
- Ravenscroft, A. (2000), 'Designing Argumentation for Conceptual Development', *Computers and Education*, 34, pp. 241–255.
- Ravenscroft, A., and Pilkington, R. (2000), 'Investigate by Design: Dialogue Models to Support Reasoning and Conceptual Change', *International Journal of Artificial Intelligence in Education*, 11, 237–298.
- Ravenscroft, A., Wegerif, R.B., and Hartley, J.R. (2007), 'Reclaiming Thinking: Dialectic, Dialogic and Learning in the Digital Age', *Special Issue of British Journal of Educational Psychology: Psychological Insights into the Use of New Technologies in Education*, 11 (5), 39–57.

- Reed, C. (2006), Representing Dialogic Argumentation, *Knowledge Based Systems*, 19 (1), 22–31.
- Rehg, W., McBurney, P., and Parsons, S. (2005), ‘Computer Decision-Support Systems for Public Argumentation: Assessing Deliberative Legitimacy’, *AI and Society*, 19 (3), 203–228.
- Retalis, S., Pain, H., and Haggith, M. (1996), ‘Arguing with the Devil: Teaching in Controversial Domains’, in *Intelligent Tutoring Systems, Third International Conference, ITS’96*, Montreal, Canada.
- Toulmin, S. (1958), *The Uses of Argument*, Cambridge, UK: Cambridge University Press.
- Walton, D. (1984), *Logical Dialogue Games and Fallacies*, Lanham, MD: University Press of America.
- Walton, D. (1989), *Question–Reply Argumentation*, Westport, CT: Greenwood Press.
- Walton, D. (1998), *The New Dialectics: Conversational Context of Argument*, Toronto: University of Toronto Press.
- Walton, D. (1999), ‘Dialectical Relevance in Persuasion Dialogue’, *Informal Logic*, 19, 119–143.
- Walton, D., and Krabbe, E. (1995), *Commitment in Dialogue: Basic Concept of Interpersonal Reasoning*, Albany, NY: State University of New York Press.
- Walton, D., and Reed, C. (2002), ‘Argumentation Schemes and Defeasible Inference’, in *Proceedings of ECAI’2002 Workshop on Computational Models of Natural Argument*, Lyon, France.
- Walton, D., Reed, C., and Macagno, F. (2008), *Argumentation Schemes*, Cambridge: Cambridge University Press.
- Yuan, T. (2004), ‘Human–Computer Debate, a Computational Dialectics Approach’, unpublished doctoral dissertation, Leeds Metropolitan University.
- Yuan, T., Moore, D., and Grierson, A. (2003), ‘Computational Agents as a Test-Bed to Study Philosophical Model “DE”, A Development of Mackenzie’s “DC”’, *Journal of Informal Logic*, 23(3), 263–284.
- Yuan, T., Moore, D., and Grierson, A. (2007a), ‘A Human–Computer Debating System and Its Dialogue Strategies’, *International Journal of Intelligent Systems, Special Issue on Computational Models of Natural Argument*, 22 (1), 133–156.
- Yuan, T., Svansson, V., Moore, D., and Grierson, A. (2007b), ‘A Computer Game for Abstract Argumentation’, in *Proceedings of IJCAI’2007 Workshop on Computational Models of Natural Argument*, Hyderabad, India, pp. 62–68.
- Yuan, T., Moore, D., and Grierson, A. (2008), ‘A Human–Computer Dialogue System for Educational Debate, A Computational Dialectics Approach’, *International Journal of Artificial Intelligence in Education*, 18, 3–26.
- Yuan, T., Moore, D., Reed, C., Ravenscroft, A., and Maudet, N. (2011), ‘Informal Logic Dialogue Games in Human–Computer Dialogue’, *Knowledge Engineering Review*, 26 (3).

Appendix 1. System knowledge base

The example shown in Figure A1 below is part of the KB adopted by the strategic agents system discussed in the paper (i.e. the agents in DE4 and DE5 and agent S in DE6). A horizontal line shows a rebuttal relationship, a vertical line with an arrow shows an upward supporting relationship and an italicised proposition indicates that it is a piece of hard evidence, i.e. statistical evidence.

Appendix 2. Dialogue transcripts

The following dialogue transcripts (DE4, DE5 and DE6) were generated by two computational agents operationalising the DE dialogue model and the proposed strategy. The annotations in square brackets were added by the author to illustrate the proposed strategic heuristics in action. The heuristic being invoked is indicated at the beginning of the annotations. In the annotations that follow, *Setup* stands for heuristics for the start of the game, *Handover* refers to arrangement when running out of strategy, *Random* stands for random selection of the available choices and others stands for various heuristics in the main text. In addition, presumed missing premises are indicated in italics.

DE4

C = DE strategic agent Chris
S = DE random agent Simon

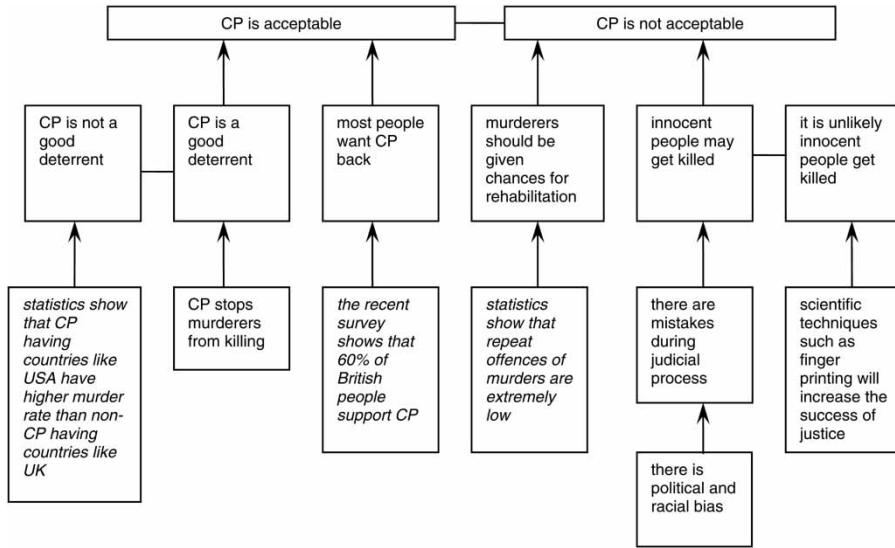


Figure A1. An example system knowledge base.

CP = Capital Punishment

C and S use the same knowledge base as depicted in Appendix 1.

001: C>Is it the case that CP is acceptable?

.....[Setup – C starts the game by asking S’s view on the issue]

002: S>No, I think CP is not acceptable.

003: C>I think CP is acceptable.

.....[Setup – C automatically adopts the opposite view]

004: S>Is it the case that CP is not a good deterrent?

005: C>No, I think CP is a good deterrent.

.....[Q3a – gives an answer which supports its own view for the sake of argument, since both answers “CP is a good deterrent” and “CP is a good deterrent” are found in C’s knowledge base]

006: S>Is it the case that mistakes rarely happen during judicial process is a prima facie reason for it is unlikely innocent people get killed?

007: C>Yes, I think mistakes rarely happen during judicial process is a prima facie reason for it is unlikely innocent people get killed.

.....[Q2b – C gives the only answer found in its knowledge base]

008: S>Is it the case that suicide bombers want to die is a prima facie reason for some people are willing to die?

009: C>Yes, I think suicide bombers want to die is a prima facie reason for some people are willing to die.

.....[Q2b – C gives the only answer found in its knowledge base]

010: S>Is it the case that CP definitely stops murderers commit further crimes is a prima facie reason for CP is acceptable?

011: C>Yes, I think CP definitely stops murderers commit further crimes is a prima facie reason for CP is acceptable.

.....[Q2b – C gives the only answer found in its knowledge base]

012: S>I think CP definitely stops murderers commit further crimes.

013: C>Is it the case that CP definitely stops murderers commit further crimes is a prima facie reason for CP is acceptable?

.....[S1a – C uses S’s previous statement to build its thesis as a response to a statement supporting its thesis]

014: S>I am not sure about it.

015: C>I think murderers should receive capital punishment.

.....[SP2b – C builds its view by giving a direct support of its thesis]

016: S>I think CP is not acceptable.

- 017: C>Is it the case that the recent survey shows that 60% British people support CP?
[S2c-C builds its view by asking S to accept a statement which supports C's thesis]
- 018: S>I am not sure about it.
- 019: C>I think most people want CP back.
[SP2b – C builds its view by giving a direct support of its thesis]
- 020: S>Is it the case that chances should be given to murderers for rehabilitation is a prima facie reason for murderers should not receive capital punishment?
- 021: C>Yes, I think chances should be given to murderers for rehabilitation is a prima facie reason for murderers should not receive capital punishment.
[Q2b – C gives the only answer found in its knowledge base]
- 022: S>I think scientific techniques will increase the success of justice is a prima facie reason for mistakes rarely happen during judicial process.
- 023: C>Is it the case that CP is not acceptable?
[S1b – C checks whether S still adheres to its thesis, given that S's previous statement is on C's side]
- 024: S>Yes, I think CP is not acceptable.
- 025: C>Is it the case that execution of murderers is fair for the people being murdered?
[S2c – C builds its view by asking S to accept a statement which supports C's thesis]
- 026: S>No, I think execution of murderers is not fair for the people being murdered.
- 027: C>Is it the case that nobody is willing to die?
[P3c – C builds its view by asking S to accept a statement which supports C's thesis]
- 028: S>I am not sure about it.
- 029: C>Is it the case that CP definitely stops murderers commit further crimes?
[P2b – C builds its view by asking S to accept a statement which supports C's thesis]
- 030: S>I am not sure about it.
- 031: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of methods and therefore hands over its turn to S]
- 032: S>Is it the case that political and racial bias often causes prejudices is a prima facie reason for there are mistakes during judicial process?
- 033: C>Yes, I think political and racial bias often causes prejudices is a prima facie reason for there are mistakes during judicial process.
[Q2b – C gives the only answer found in its knowledge base]
- 034: S>Is it the case that some people are willing to die?
- 035: C>I am not sure about it.
[Q3b – C finds both answers "some people are willing to die" and "nobody is willing to die" in its knowledge base, a positive response would support S's thesis, and there is no further support for a negative response. C therefore gives a non-committal response for the sake of argument]
- 036: S>I think murderers should not receive capital punishment.
- 037: C>Why is it the case that murderers should not receive capital punishment?
[S2e – C challenges S's statement]
- 038: S>Because the recent survey shows that 60% British people support CP.
- 039: C>Is it the case that the recent survey shows that 60% British people support CP is a prima facie reason for most people want CP back?
[S1a – C uses S's previous statement to build its thesis as a response to a statement supporting its thesis]
- 040: S>Yes, I think the recent survey shows that 60% British people support CP is a prima facie reason for most people want CP back.
- 041: C>Is it the case that most people want CP back is a prima facie reason for CP is acceptable?
[P1 – C carries on to execute its plan]
- 042: S>No, I think it is not the case that most people want CP back is a prima facie reason for CP is acceptable.
- 043: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of methods and therefore hands over its turn to S]
- 044: S>I don't think mistakes rarely happen during judicial process is a prima facie reason for it is unlikely innocent people get killed.
- 045: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of methods and therefore hands over its turn to S]
- 046: S>Why is it the case that most people want CP back?

- 047: C>Because the recent survey shows that 60% British people support CP.
[C2 – C answers a challenge with a ground]
- 048: S>I think CP definitely stops murderers commit further crimes.
- 049: C>Is it the case that CP definitely stops murderers commit further crimes is a prima facie reason for CP is acceptable?
[S1a – C uses S’s previous statement to build its thesis as a response to a statement supporting its thesis]
- 050: S>I am not sure about it.
- 051: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of methods and therefore hands over its turn to S]
- 052: S>I think there are mistakes during judicial process.
- 053: C>Is it the case that scientific techniques will increase the success of justice?
[S2c – C builds a plan aiming to demolish the previous statement]
- 054: S>No, I think scientific techniques will not increase the success of justice.
- 055: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of methods and therefore hands over its turn to S]
- 056: S>I think a state has the right to deliberately execute the wrong people.
- 057: C>Is it the case that CP is not acceptable?
[S1b – C checks whether S stills adheres to its view given the previous statement, which goes against S’s thesis]
- 058: S>I am not sure about it.
- 059: R>Simon, you seem to have given up your view, and so the game ends.
[Referee – Referee stops the game since S has given up its view]

DE 5

C = DE strategic agent Chris

S = DE strategic agent Simon

CP = Capital Punishment C and S use the same knowledge base as depicted in Appendix 1.

- 001: C>Is it the case that CP is acceptable?
[Setup – C starts the game by asking S’s view on the issue of capital punishment]
- 002: S>Yes, I think CP is acceptable.
[Random – S chooses to support CP]
- 003: C>I think CP is not acceptable.
[Setup – C automatically adopts the opposite view]
- 004: S>I think CP is a good deterrent, and “CP is a good deterrent” is a reason for “CP is acceptable”.
[S2c – S issues a direct support of its thesis, a subtopic is raised]
- 005: C>But CP having countries like USA has higher murder rate than UK, and “CP having countries like USA has higher murder rate than UK” is a reason for “CP is not a good deterrent”.
[S2b – C issues a piece of direct contradictory evidence to S’s statement “CP is a good deterrent”]
- 006: S>I think most people want CP back.
[S2f – S issues a direct support of its thesis, another subtopic is raised]
- 007: C>Why is it the case that most people want CP back?
[S2e – C has doubt on but no objection to the popularity point, C therefore elicits further grounds]
- 008: S>Because the recent survey shows that 60% British people support CP.
[C2 – S provides an evidential support]
- 009: C>I think it is wrong to take a human life.
[S2f – C changes the focus and raises a subtopic – ethical issue]
- 010: S>Is it the case that a state has the right to deliberately execute the wrong people?
[S2c – S has objections to C’s statement “it is wrong to take a human life”, and then plans a line of questions to make C accept the objection as in turn 12 and 14]
- 011: C>Yes, I think a state has the right to deliberately execute the wrong people.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 012: S>Is it the case that a state has the right to deliberately execute the wrong people is a prima facie reason for it is not always wrong to take a human life?
[P1 – S continues its line of questions]

- 013: C>Yes, I think a state has the right to deliberately execute the wrong people is a prima facie reason for it is not always wrong to take a human life.
[Q2b – C answers truthfully and responds with the only found answer in its knowledge base]
- 014: S>I think it is not always wrong to take a human life.
[P1 – S ends its line of questions by stating its counter-view on C’s point “it is wrong to take a human life”]
- 015: C>Is it the case that human lives are scarce?
[S2d – C has something more to say on the point of whether it is wrong to take a human life, and uses a line of questions to attack S’s view “it is not always wrong to take a human life”]
- 016: S>Yes, I think human lives are scarce.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 017: C>Is it the case that human lives are scarce is a prima facie reason for it is wrong to take a human life?
[P1 – C continues its line of questions]
- 018: S>Yes, I think human lives are scarce is a prima facie reason for it is wrong to take a human life.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 019: C>I think it is wrong to take a human life.
[Q2b – C ends its line of questions by stating its counter-view to S’s point “it is not always wrong to take a human life”]
- 020: S>I think murderers should receive capital punishment, and “murderers should receive capital punishment” is a reason for “CP is acceptable”.
[S2f – S issues a direct support of its thesis, and the moral issue is raised]
- 021: C>I think chances should be given to murderers for rehabilitation, and “chances should be given to murderers for rehabilitation” is a reason for “murderers should not receive capital punishment”.
[S2C – C has an objection to S’s moral point, and directly states that objection]
- 022: S>Why is it the case that chances should be given to murderers for rehabilitation?
[S2e – S has doubt on but no objection to C’s point “chances should be given to murderers for rehabilitation”, S therefore elicit further grounds]
- 023: C>Because statistics shows that repeat offences of murders are extremely low.
[C2 – C provides an evidential support]
- 024: S>Is it the case that nobody is willing to die?
[S2f – S switches the focus and builds a distance plan (asking C to accept statement not directly supporting S’s thesis) and tries to make C accept its thesis gradually]
- 025: C>No, I think some people are willing to die.
[Q3a – unfortunately, C gives an answer which is not the one S hopes for]
- 026: S>Why is it the case that some people are willing to die?
[P3b – S tries to remove the obstacles and put its plan on track]
- 027: C>Because suicide bombers want to die.
[C2 – C provides a support for its view “some people are willing to die”]
- 028: S>Is it the case that execution of murderers is fair for the people being murdered?
[P3bii – S fails to remove S’s point “some people are willing to die”, which suggests the plan is becoming a blind alley. S therefore switches the focus and starts another distance plan to build its thesis]
- 029: C>Yes, I think execution of murderers is fair for the people being murdered.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 030: S>Is it the case that execution of murderers is fair for the people being murdered is a prima facie reason for murderers should receive capital punishment?
[P1 – S continues to execute its plan]
- 031: C>Yes, I think execution of murderers is fair for the people being murdered is a prima facie reason for murderers should receive capital punishment.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 032: S>Is it the case that murderers should receive capital punishment is a prima facie reason for CP is acceptable?
[P1 – S continues to execute its plan]
- 033: C>Yes, I think murderers should receive capital punishment is a prima facie reason for CP is acceptable.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]

- 034: S>I think CP is acceptable.
[P1 – S successfully ends its plan by stating its thesis]
- 035: C>Is it the case that political and racial bias often causes prejudices?
[S2f – C has its own point of view concerning whether “CP is acceptable”, then builds its thesis by making a plan as well]
- 036: S>Yes, I think political and racial bias often causes prejudices.
[Q2b – S answers truthfully and responds with the only answer found in its knowledge base]
- 037: C>Is it the case that political and racial bias often causes prejudices is a prima facie reason for there are mistakes during judicial process?
[P1 – C continues to execute its plan]
- 038: S>Yes, I think political and racial bias often causes prejudices is a prima facie reason for there are mistakes during judicial process.
[Q2b – S answers truthfully and responds with the only answer found in its knowledge base]
- 039: C>Is it the case that there are mistakes during judicial process is a prima facie reason for innocent people may get killed?
[P1 – C continues to execute its plan]
- 040: S>Yes, I think there are mistakes during judicial process is a prima facie reason for innocent people may get killed.
[Q2b – S answers truthfully and responds with the only answer found in its knowledge base]
- 041: C>Is it the case that innocent people may get killed is a prima facie reason for CP is not acceptable?
[P1 – C continues to execute its plan]
- 042: S>Yes, I think innocent people may get killed is a prima facie reason for CP is not acceptable.
[Q2b – S answers truthfully and responds with the only answer found in its knowledge base]
- 043: C>I think CP is not acceptable.
[P1 – C successfully ends its plan by stating its thesis]
- 044: S>I think if you have anything more to say, you can go on.
[Handover – S runs out of methods and therefore hands over its turn]
- 045: C>Is it the case that CP is acceptable?
[S1b – C checks whether S still adheres to its thesis]
- 046: S>Yes, I think CP is acceptable.
[Q3a – S still adheres to its thesis though run out of moves in turn 044]
- 047: C>I think innocent people may get killed, and “*innocent people may get killed*” is a reason for “*CP is not acceptable*”.
[S2c – C raises a further point – the consequence issue]
- 048: S>I think mistakes rarely happen during judicial process, and “*mistakes rarely happen during judicial process*” is a reason for “*it is unlikely innocent people get killed*”.
[S2c – S has an objection to C’s consequence point, and directly states that objection]
- 049: C>Why is it the case that mistakes rarely happen during judicial process?
[S2e – C has doubt but no objection to S’s point “mistakes rarely happen during judicial process”, C therefore elicit for further reasons]
- 050: S>Because scientific techniques will increase the success of justice.
[C2 – S provides an evidential support]
- 051: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of methods]
- 052: R>Ok, you both are doing very well. We need an audience vote on who wins.
[Referee – Referee calls off the debate since both agents have run out of methods]

DE6

C = DE strategic agent Chris

S = DE strategic agent Simon

S uses the knowledge base as depicted Appendix 1.

C uses a subset of the knowledge base as depicted in Appendix 1.

- 001: C>Is it the case that CP is acceptable?
[Setup – C starts the game by asking S’s view on the issue of capital punishment]
- 002: S>Yes, I think CP is acceptable.
[Random – S chooses to support CP]

- 003: C>I think CP is not acceptable.
[Setup – C automatically adopts the opposite view]
- 004: S>I think most people want CP back.
[S2c – S issues a direct support of its thesis, a subtopic is raised]
- 005: C>Is it the case that CP is acceptable?
[S1b – C does not recognise the previous statement, and then assesses whether S still adheres to its thesis]
- 006: S>Yes, I think CP is acceptable.
[S3a – S adheres to its thesis]
- 007: C>Is it the case that there are mistakes during judicial process?
[S2c – S makes a plan and asks C to accept a statement that supports S's thesis]
- 008: S>No, I think mistakes rarely happen during judicial process.
[Q3a – S gives an answer which supports its own view for the sake of argument, since both answers “there are mistakes during the judicial process” and “mistakes rarely happen during judicial process” are found in S's knowledge base]
- 009: C>I think innocent people may get killed.
[P3c – S issues a direct support of its thesis, and a subtopic is raised]
- 010: S>Is it the case that scientific techniques will increase the success of justice?
[S2d – S has objections to C's statement “innocent people may get killed”, and then plans a line of questions to make C accept the objection]
- 011: C>I am not sure about it.
[Q1 – C does not recognise the statement being questioned, and therefore answer truthfully]
- 012: S>I think murderers should receive capital punishment.
[P2b – S issues a direct support of its thesis, and a subtopic is raised]
- 013: C>Is it the case that CP is acceptable?
[S1b – C does not recognise the previous statement, then, assesses whether S still adheres to its thesis]
- 014: S>Yes, I think CP is acceptable.
[Q3a – S adheres to its thesis]
- 015: C>Why is it the case that CP is acceptable?
 ...[S2e – C hopes to find the attacking point behind it, and therefore challenges for further reasons]
- 016: S>Because murderers should receive capital punishment.
[C3 – S issues a support which can be further supported]
- 017: C>Is it the case that CP is acceptable?
[S1b – C does not recognise the previous statement, and then assesses whether S still adheres to its thesis]
- 018: S>Yes, I think CP is acceptable.
[Q3a – S adheres to its thesis]
- 019: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of moves and therefore hands over the turn]
- 020: S>Is it the case that CP is not acceptable?
[S1b – S checks whether C still adheres to its thesis]
- 021: C>Yes, I think CP is not acceptable.
[Q3a – C adheres to its thesis since it has a support “innocent people may get killed” which is uttered in turn 09]
- 022: S>Is it the case that execution of murderers is fair for the people being murdered?
[S2d – S makes a plan and asks C to accept a statement that supports S's thesis]
- 023: C>I am not sure about it.
[Q1 – C does not recognise the statement being questioned, and then answers truthfully]
- 024: S>I think CP is a good deterrent.
[P2b – S issues a direct support of its thesis]
- 025: C>Why is it the case that CP is a good deterrent?
[S2e – C has doubt on but no objection to S's deterrent point, C therefore elicits for further reasons]
- 026: S>Because statistics shows an increase in murder rate since the abolition of CP.
[C2 – C provides evidential support]

- 027: C>Is it the case that CP is acceptable?
[S1b – C does not recognise the previous statement, and then assesses whether S still adheres to its thesis]
- 028: S>Yes, I think CP is acceptable.
[Q3a – S adheres to its thesis]
- 029: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of moves and therefore hands over the turn]
- 030: S>Is it the case that CP is not acceptable?
[S1b – S checks whether C still adheres to its thesis]
- 031: C>Yes, I think CP is not acceptable.
[Q3a – C adheres to its thesis since it has a support “innocent people may get killed” which was uttered in turn 09]
- 032: S>Is it the case that nobody is willing to die?
[S3c – S builds a distance plan by asking C to accept a statement not directly supporting S’s thesis, and tries to make C accept its thesis gradually]
- 033: C>Yes, I think nobody is willing to die.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 034: S>Is it the case that nobody is willing to die is a prima facie reason for CP makes people less likely commit serious crimes?
[P1 – S continues to execute its plan]
- 035: C>Yes, I think nobody is willing to die is a prima facie reason for CP makes people less likely commit serious crimes.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 036: S>Is it the case that CP makes people less likely commit serious crimes is a prima facie reason for CP is a good deterrent?
[P1 – S continues to execute its plan]
- 037: C>Yes, I think CP makes people less likely commit serious crimes is a prima facie reason for CP is a good deterrent.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 038: S>Is it the case that CP is a good deterrent is a prima facie reason for CP is acceptable?
[P1 – S continues to execute its plan]
- 039: C>Yes, I think CP is a good deterrent is a prima facie reason for CP is acceptable.
[Q2b – C answers truthfully and responds with the only answer found in its knowledge base]
- 040: S>I think CP is acceptable.
[P1 – S successfully ends its plan by stating its thesis]
- 041: C>I think if you have anything more to say, you can go on.
[Handover – C runs out of moves and hands over the turn]
- 042: S>Is it the case that CP is not acceptable?
[S1b – S checks whether C still adheres to its thesis]
- 043: C>Yes, I think CP is not acceptable.
[Q3a – C adheres to its thesis since it has a support “innocent people may get killed” which was uttered in turn 09]
- 044: S>Why is it the case that CP is not acceptable?
[S2e – S challenges for further reasons]
- 045: C>Because innocent people may get killed.
[C3 – C gives reasons supporting its thesis]
- 046: S>Why is it the case that innocent people may get killed?
[S2e – S challenges for further reasons]
- 047: C>Because there are mistakes during judicial process.
[C2 – C provides a support]
- 048: S>Why is it the case that there are mistakes during judicial process?
[S2e – S challenges for further reasons]
- 049: C>I don’t know why there are mistakes during judicial process.
[C5 – C has no further support available in its knowledge base, and therefore speaks the truth]
- 050: S>Why is it the case that innocent people may get killed?
[W1a – S challenges the statement being supported by the withdrawn statement]

- 051: C>I don't know why innocent people may get killed.
.....[C5 – C withdraws the point since its support “there are mistakes during judicial process”
has been withdrawn in turn 049]
- 052: S>Is it the case that CP is not acceptable?
.....[W1b – S assesses whether C still adheres to its thesis since C has withdrawn a direct
support of its thesis “innocent people may get killed”]
- 053: C>No, I think CP is acceptable.
.....[Q3c – C accepts S's view since C's only thesis support has been withdrawn in turn 051]
- 054: R>Chris, you lose and Simon wins.
.....[Referee – The debate ends since C's view has been converted to S's view]