# Arguing with Confidential Information

**Nir Oren** and **Timothy J. Norman** and **Alun Preece**[1]

**Abstract.** While researchers have looked at many aspects of argumentation, an area often neglected is that of argumentation strategies. That is, given multiple possible arguments that an agent can put forth, which should be selected in what circumstances. In this paper, we propose a heuristic that implements one such strategy. The heuristic assigns a utility cost to revealing information, as well as a utility to winning, drawing and losing an argument. An agent participating in a dialogue then attempts to maximise its utility. We present a formal argumentation framework in which this heuristic may operate, and show how it functions within the framework. Finally, we discuss how this heuristic may be extended in future work, and its relevance to argumentation theory in general.

## 1 Introduction

Argumentation has emerged as a powerful reasoning mechanism in many domains. One common dialogue goal is to persuade, where one or more participants attempt to convince the others of their point of view. This type of dialogue can be found in many areas including distributed planning and conflict resolution, education and in models of legal argument. At the same time that the breadth of applications of argumentation has expanded, so has the sophistication of formal models designed to capture the characteristics of the domain. Prakken [11] for example, has focused on legal argumentation, and has identified four layers with which such an argumentation framework must concern itself. These are:

- The logical layer, which allows one to represent basic concepts such as facts about the world. Most commonly, this layer consists of some form of non–monotonic logic.
- The dialectic layer, in which argument specific concepts such as the ability of an argument to defeat another are represented.
- The procedural layer governs the way in which argument takes place. Commonly, a dialogue game [17] is used to allow agents to interact with each other.
- The heuristic layer contains the remaining parts of the system. Depending on the form of the underlying layers, these may include methods for deciding which arguments to put forth and techniques for adjudicating arguments.

While many researchers have focused on the lowest two levels (excellent surveys can be found in [3, 11, 12]), and investigation into various aspects of the procedural layer is ongoing (for example, [16, 6]), many open questions remain at the heuristic level.

In this paper, we propose a decision heuristic for an agent allowing it to decide which argument to put forth. The basis for our idea is simple; the agent treats some parts of its knowledge as more confidential than other parts, and, while attempting to win the argument, attempts to reveal as little of the more secret information to others as possible. This heuristic often emerges in negotiation dialogues, as well as persuasion dialogues in hostile setting (such as takeover talks or in some legal cases). Utilising this heuristic in arguments between computer agents can also be useful; revealing confidential information in an ongoing dialogue may damage an agent's chances of winning a future argument.

In the next section, we examine a few existing approaches to strategy selection, after which we discuss the theoretical foundations of our approach. We then present the heuristic, after which we see how it operates by means of an example. We conclude the paper by looking at possible directions in which this work can be extended.

[1] Department of Computing Science, University of Aberdeen, AB24 3UE, Scotland, email: noren,tnorman,apreece@csd.abdn.ac.uk

## 2 Background and related research

Argumentation researchers have recognised the need for argument selection strategies for a long time. However, the field has only recently started receiving more attention. Moore, in his work with the DC dialectical system [7], suggested that an agent's argumentation strategy should take three things into account:

- Maintaining the focus of the dispute.
- Building its point of view or attacking the opponent's one.
- Selecting an argument that fulfils the previous two objectives.

The first two items correspond to the military concept of a strategy, i.e. a high level direction and goals for the argumentation process. The third item corresponds to an agent's tactics. Tactics allow an agent to select a concrete action that fulfils its higher level goals. While Moore's work focused on natural language argument, these requirements formed the basis of most other research into agent argumentation strategies.

In 2002, Amgoud and Maudet [1] proposed a computational system which would capture some of the heuristics for argumentation suggested by Moore. Their system requires very little from the argumentation framework. A preference ordering is needed over all possible arguments, and a level of prudence is assigned to each agent. An argument is assigned a strength based on how convoluted a chain of arguments is required to defend it. An agent can then have a "build" or "destroy" strategy. When using the build strategy, an agent asserts arguments with a strength below its prudence level. If it cannot build, it switches to a destroy strategy. In this mode, it attacks an opponent's arguments when it can. While the authors note other strategies are reasonable, they make no mention of them. Shortcomings of their approach include its basis on classical propositional logic and the assumption of unbounded rationality; computational limits may affect the arguments agents decide to put forth. Finally, no attempt is made to capture the intuition that a fact defended by multiple arguments is more acceptable than one defended by fewer (the so called "accrual of evidence" argument scheme [9]).

Using some ideas from Amgoud's work, Kakas et al. [5] proposed a three layer system for agent strategies in argumentation. The first layer contains "default" rules, of the form $utterance \leftarrow condition$, while the two higher layers provide preference orderings over the rules. Assuming certain restrictions on the rules, they show that only one utterance will be selected using their system, a trait they refer to as determinism. While their approach is able to represent strategies proposed by a number of other techniques, it does require hand crafting of the rules. No suggestions are made regarding what a "good" set of rules would be.

In [2], Bench-Capon describes a dialogue game based on Toulmin's work. He identifies a number of stages in the dialogue in which an agent might be faced with a choice, and provides some heuristics as to what argument should be advanced in each of these cases. Only an informal justification for his heuristics is provided.

## 3 Confidentiality Based Argumentation

In many realms of argument, auxiliary considerations (apart from simply winning or losing the argument) come into play. In many scenarios, one such consideration involves hiding certain information from an opponent. In this section, we describe a utility based heuristic to guide an agent taking part in a dialogue while being careful about what information it reveals. When faced with a number of possible arguments that it can advance, we claim it should put forth the one that minimises the exposure of information that it would like to keep private. The limitations of our current approach, as well as extensions and refinements to it are discussed in Section 5.

Our formalism is based on many ideas from other formal argument systems (e.g. [4, 10, 8, 15]). We use our own formal argumentation system which introduces features not seen in others, and which allows us to study our proposed heuristic in isolation. Our argumentation framework is very simple, and does not contain features such as a preference ordering on arguments which allows one to overrule another. Our system is formalised in two parts. After specifying the argumentation framework, our heuristic is described, in terms of the framework.

### 3.1 The Argumentation Framework

Argumentation takes place over the language $\Sigma$, which contains propositional literals and their negation.

**Definition 1** ***Argument*** *An argument is a pair $(P, c)$, where $P \subseteq \Sigma \cup \{\top\}$ and $c \in \Sigma$ such that if $x \in P$ then $\neg x \notin P$. We define $Args(\Sigma)$ to be the set of all possible arguments derivable from our language.*

$P$ represents the premises of an argument (also referred to as an argument's support), while $c$ stands for an argument's conclusion. Informally, we can read an argument as stating "if the conjunction of its premises holds, the conclusion holds". Facts can be represented using the form $(\top, a)$.

Arguments interact by supporting and attacking each other. Informally, when an argument attacks another, it renders the latter's conclusions invalid.

**Definition 2** ***Attack*** *An argument $A = (P_a, c_a)$ attacks $B = (P_b, c_b)$ if $\exists f \in P_b$ such that $f = \neg c_a$ or $c_a = \neg c_b$. For convenience, we write this as $attacks(A, B)$.*

An argument is only relevant to an instance of argumentation if its premises are true. We call such an argument *justified*. However, a simple definition of this concept can cause problems when it comes to self attacking (or self defending) arguments, as well as circular reasoning, and care must thus be taken when describing this concept. Before doing so, we must (informally) describe the proof theory used to determine which literals and arguments are in effect at any time.

The idea behind determining what arguments and literals are admissible at any time is as follows. We start by looking at the facts, and determining what knowledge can be derived from them by following chains of argument. Whenever a conflict occurs (i.e. we are able to derive both $x$ and $\neg x$), we remove these literals from our derived set. Care must be taken to also get rid of any arguments (and further facts) derived from any conflicting literals. To do this, we keep track of the conflicting literals in a separate set, whenever a new conflict arises, we begin the knowledge determination process afresh, never adding any arguments whose conclusions are in the conflicting set to the knowledge set. The philosophical and practical ramifications of this approach are examined in Section 5.

More formally, an instance of the framework creates two sets $J \subseteq Args(\Sigma)$ and $C \subseteq \Sigma$ representing justified arguments and conflicts respectively.

**Definition 3** ***Derivation*** *An argument $A = (P_a, c_a)$ is derivable from a set $S$ given a conflict set $C$ (written $S, C \vdash A$) iff $c_a \notin C$ and $(\forall p \in P_a : (\exists s \in S$ such that $s = (P_s, p)$ and $p \notin C)$ or $P_a = \{\top\})$.*

Clearly, we need to know what elements are in $C$. Given a knowledge base of arguments $\kappa \subseteq Args(\Sigma)$, this can be done with the following reasoning procedure:

$$J_0 = \{A | A \in \kappa \text{ such that } \{\}, \{\} \vdash A\}$$
$$C_0 = \{\}$$

Then, for $i > 0, j = 1 \ldots i$, we have:

$$C_i = C_{i-1} \cup \{c_A, \neg c_A | \exists A = (P_A, c_A), B = (P_B, \neg c_A) \in J_{i-1} \text{ such that } attacks(A, B)\}$$

$$X_{i0} = \{A | A \in \kappa \text{ and } \{\}, C_i \vdash A\}$$
$$X_{ij} = \{A | A \in \kappa \text{ and } X_{i(j-1)}, C_i \vdash A\}$$

$$J_i = X_{ii}$$

The set $X$ allows us to recompute all derivable arguments from scratch after every increment of $i$[2]. Since $i$ represents the length of a chain of arguments, when $i = j$ our set will be consistent to the depth of our reasoning, and we may assign all of these arguments to $J$. Eventually, $J_i = J_{i-1}$ (and $C_i = C_{i-1}$) which means there are no further arguments to find. We can thus define the conclusions asserted by $\kappa$ as $K = \{c | A = (P, c) \in J_i\}$, for the smallest $i$ such that $J_i = J_{i+1}$. We will use the shorthand $K(\kappa)$ and $C(\kappa)$ to represent those literals which are respectively asserted by, or in conflict with the knowledge base $\kappa$.

We provide an example which illustrates the approach (note that not all steps are shown):

[2] This allows us to get rid of long invalid chains of arguments, as well as detect and eliminate arbitrary loops

**Example 1** $\kappa = \{(\top, s), (s, t), (t, \neg s)\}$
$J_0 = \{(\top, s)\}$, $C_1 = \{\}$,
$J_1 = X_{11} = \{(\top, s), (s, t)\}$
...
$J_2 = (\top, s), (s, t), (t, \neg s)$, $C_3 = \{s, \neg s\}$
$X_{30} = \{\} \ldots J_4 = J_3 = \{\}$

## 3.2 The Heuristic

Agents engage in a dialogue using the argumentation framework described above in an attempt to persuade each other of certain facts. An agent has a private knowledge base ($KB$) as well as a goal literal $g$ and a preference ranking $\rho$ which specifies an agent's reluctance to reveal certain literals. The environment, apart from containing agents, contains a public knowledge base which takes on a role similar to a global commitment store, and we thus refer to it as $CS$ below.

**Definition 4** ***Environment*** *An environment is a pair* $(Agents, CS)$ *where* $Agents$ *is the set of agents participating in the dialogue and* $CS \subseteq Args(\Sigma)$

**Definition 5** ***Agent*** *An Agent* $\alpha \in Agents$ *is a tuple* $(Name, KB, \rho, g, U_{win}, U_{draw}, U_{lose})$ *where* $KB \subseteq Args(\Sigma)$, $g \in \Sigma$. $\rho$ *is a preference ranking function and* $U_{win}, U_{draw}, U_{lose} \in \Re$ *are the utilities gained for winning, drawing or losing an argument.*

The preference ranking expresses the "cost" to an agent of revealing certain information in a specific context. It maps a set of literals $L$ to a real number. The cost of being in a certain environmental state is the result of applying the preference ranking function $\rho$ to the literals present in that state.

**Definition 6** ***Preference Ranking*** *A preference ranking* $\rho$ *is a function* $\rho : L \rightarrow \Re$ *where* $L \subseteq 2^{\Sigma}$.

Agents take turns to put forward a line of argument consisting of a number of individual arguments. For example, an agent could make the utterance $\{(\top, a), (a, b)\}$. Alternatively, an agent may pass (by uttering an empty argument $\{\}$). The dialogue ends when $CS$ has remained unchanged for $n$ turns i.e. after all players have had a chance to modify it, but didn't (this is normally caused by all agents having passed consecutively). Once this has happened, the acceptable set of arguments is computed over the $CS$, and the status of each agent's goal can be determined, allowing one to compute the winners of the game.

**Definition 7** ***Turns and utterances*** *The function*

$$turn : Environment \times Name \rightarrow Environment$$

*takes an environment and an agent label, and returns a new environment containing the result of the utterance* ($utterance : Environment \times Name \rightarrow 2^{Args(\Sigma)}$) *made by the labelled agent during its turn.*

$$turn(Environment, \alpha) = (Agents, \{CS \cup utterance(Environment, \alpha)\})$$

At turn $i$, we set $\alpha = Agent_{i \bmod n}$, where $n$ is the number of agents taking part in the dialogue. The $utterance$ function is dependant on the agent's strategy, and we will describe one such strategy below. Before doing so, we define the dialogue game itself. Each turn in the dialogue game results in a new public commitment store, which is used by agents in later turns.

**Definition 8** ***Dialogue game*** *The dialogue game is defined as*
$turn_0 = turn((Agents, CS_0), Agent_0)$
$turn_{i+1} = turn(turn_i, Agent_{i \bmod n})$

*The game ends when* $turn_i \ldots turn_{i-n+1} = turn_{i-n}$.

$CS_0$ is dependent on the system, and contains any arguments that are deemed to be common knowledge (though these arguments may be attacked like any other argument during later turns in the game). Also, note that the null utterance $\{\}$ is defined to be a pass.

By using the procedure described earlier, agents can

- Determine, by looking at $CS$, what literals are in force and in conflict.
- Determine, by combining $CS$ with parts of their own knowledge base, what literals they can prove (or cause to conflict).

By doing the latter, together with examining which literals are introduced into $K$ and $C$, as well as their cost as computed from $\rho$, an agent will narrow down the range of arguments it will consider submitting, though it may still have multiple arguments to choose from. It should be noted that an agent might be willing to draw or even lose an argument rather than reveal too much information. Winning (or drawing) an argument earns the agent a certain amount of utility. Thus, the final choice about which argument to put forth is based on the effect of the argument in combination with its utility cost. We begin by defining the set of winning and drawing arguments paying no attention to the argument cost.

**Definition 9** ***Winning arguments*** *An agent's set of winning arguments is defined as*
$Win = \{A \in 2^{KB} | \, g \in K(A \cup KB)$ *and if* $A \neq \{\}, \{\} \notin A\}$.

**Definition 10** ***Drawing arguments*** *An agent's set of drawing arguments is defined as*
$Draw = \{A \in 2^{KB} | \, (g \in C(A \cup KB)$ *or* $\{g, \neg g\} \not\subseteq K(A \cup KB))$ *and if* $A \neq \{\}, \{\} \notin A\}$

While many possibilities exist as to how to weigh the cost of information, for reasons discussed in Section 5, we compute the information cost based on the literals present in $CS$ after an argument has been advanced. Currently we make no distinction between whether information is revealed due to an utterance we make, or whether another agent revealed it.

**Definition 11** ***Argument utility*** *Given an agent with a preference ranking* $\rho$*, we define an agent's net utility* $U$ *for advancing an argument* $A$ *as*

$$U(A) = \begin{cases} U_{win} - \rho(L) & \text{if } A \in Win \\ U_{draw} - \rho(L) & \text{if } A \in Draw \\ U_{lose} - \rho(L) & \text{otherwise} \end{cases}$$

*such that* $L = K(CS \cup A) \cup C(CS \cup A)$
*The utterance an agent makes is chosen from the set of arguments that maximise its utility:*

$$utterance \in \{a \subseteq A | \forall a, b \, U(a) \geq U(b)\}$$

At the end of the game, the literals in $K(CS)$ will be those for which undefeated arguments exist.

## 4 Example

To increase readability, we present our example in a somewhat informal manner. The argument consists of a hypothetical dialogue between a government and some other agent regarding the case for, or against, weapons of mass destruction (WMDs) existing at some location.

Assume that our agent ($Agent_0$) would like to show the existence of WMDs, i.e. $g = WMD$. Assume further that $U_{win} = 100, U_{draw} = 50, U_{lose} = 0$ and that the following arguments exist in the agent's private $KB$ (where the context is clear, we omit brackets):

$$(\top, spysat), (\top, chemicals), (\top, news), (\top, factories)$$

$$(\top, smuggling), (smuggling, \neg medicine), (news, WMD)$$

$$(\{factories, chemicals\}, WMD), (spysat, WMD)$$

$$(\{sanctions, smuggling, factories, chemicals\}, \neg medicine)$$

While we will not fully describe the agents preference rating function $\rho$, we set the costs for tuples including literals as follows:

| | |
|---|---|
| $(spysat, 100)$ | $(chemicals, 30)$ |
| $(news, 0)$ | $(\{medicine, chemicals\}, 50)$ |
| $(smuggling, 30)$ | $(factories, 0)$ |

Note that if both medicine and chemicals are present, the agent's utility cost is 50, not 80.

As an example, $\rho$ of an environment state containing both $spysat$ and $chemicals$ will be assigned a cost of 130.

The dialogue might thus proceed as follows:

$$\begin{array}{lll}
(\mathbf{1}) & Agent_0: & (\top, news), (news, WMD) \\
(\mathbf{2}) & Agent_1: & (\top, \neg news) \\
(\mathbf{3}) & Agent_0: & (\top, factories), (\top, chemicals), \\
 & & (\{factories, chemicals\}, WMD) \\
(\mathbf{4}) & Agent_1: & (\top, sanctions), \\
 & & (\{sanctions, factories, chemicals\}, \\
 & & \quad medicine), (medicine, \neg WMD) \\
(\mathbf{5}) & Agent_0: & (\top, smuggling), \\
 & & (\{sanctions, smuggling, factories, \\
 & & \quad chemicals\}, \neg medicine) \\
(\mathbf{6}) & Agent_1: & \{\} \\
(\mathbf{7}) & Agent_0: & \{\}
\end{array}$$

Informally, the dialogue proceeds as follows: $Agent_0$ claims that WMDs exist since the news says they do. $Agent_1$ retorts that he has not seen those news reports. $Agent_0$ then points out that factories and chemicals exist, and that these were used to produce WMDs. In response, $Agent_1$ says that due to sanctions, these were actually used to produce medicine. $Agent_0$ attacks this argument by pointing out that smuggling exists, which means that the factories were not used to produce medicines, reinstating the WMD argument. Both agents have nothing more to say, and thus pass. $Agent_0$ thus wins the game.

It should be noted that while $Agent_0$ is aware that spy satellites have photographed the WMDs, it does not want to advance this argument due to the cost of revealing this information. The final utility gained by $Agent_0$ for winning the argument is 20: 100 for winning the argument, less 30 for revealing $smuggling$, and 50 for the presence of the $chemicals$ and $medicine$ literals. Also, note that the fact that $Agent_1$ revealed the existence of medicines cost $Agent_0$ an additional 20 utility. This is somewhat counterintuitive, and extensions to overcome this behaviour are examined in the next section.

## 5 Discussion

This section examines the argumentation framework and the heuristic, tying it back to the concept of an argumentation strategy as proposed by Moore. We also examine some of the novel features of argument that emerge when dialogue takes place in the framework using the heuristic, and propose avenues for future research.

Our approach seems to share much in common with the "sceptical" approach to argumentation. When arguments conflict, we refuse to decide between them, instead ruling them both invalid. This means that our reasoning procedure is not complete, given the (rather convoluted) set of arguments $(\top, A)$, $(\top, B)$, $(A, \neg B)$, $(B, \neg A)$, $(A, C)$, $(B, C)$, $(\neg A, C)$, $(\neg B, C)$ we can intuitively see that $C$ should hold, but doesn't. Other argumentation systems (namely those utilising the unique–status–assignment approach [12]) are similarly incomplete, leaving this an open area for future research. Our sceptical approach does yield a sound system, as no conflicting arguments will remain in the final set of arguments.

The simplicity of our approach means that only specific types of arguments can be represented (namely, those whose premises are a conjunction of literals, and whose conclusion is a single literal). However, as seen in the example, even with this limitation, useful arguments can still emerge.

We developed our own argumentation framework rather than using an existing one for a number of reasons, including:

- The abstract nature of many frameworks (e.g. [4]) makes arguments atomic concepts. We needed a finer level of granularity so as to be able to talk about which facts are exposed (allowing us to measure the amount of information revealed during the dialogue process). Less abstract frameworks (e.g. [13, 10]), while looking at concepts such as derivability of arguments, still have as their main focus, the interactions between arguments.
- Almost all other frameworks define higher level concepts in terms of arguments attacking, defeating and defending one another. For us, the concept of one argument justifying another is critical, together with the concept of attack.
- Other argumentation systems contain concepts which we do not require, such as a preference ordering over arguments.

Another significant difference between our argumentation framework and most existing approaches is the scope of arguments. In our approach, agents can be aware of and utter arguments of which other agents are unaware. For example, even if no other agent knew of the literals $X$ and $Y$, an agent could make the utterance $(\{X, Y\}, Z)$. An agent arguing for $\neg Z$ would then have no choice but to try obtain a draw result.

While representing the heuristic using one of the other approaches is (probably) not impossible, it appears to be more difficult than by using our own system.

Looking at Moore's three criteria for an agent argumentation strategy, we see that our heuristic fulfils its requirements. If the focus of the argument were not maintained, more information would be given than is strictly necessary to win, thus fulfilling the first requirement. Both the second and third requirements are clearly met by the decision procedure for which argument to advance described in Definition 11.

The way in which we represent the information "leaked" during the dialogue, as well as calculate the agent's net utility, while simple, allows us to start studying dialogues in which agents attempt to hide information. Until now, most work involving utility and argumentation has focused on negotiation dialogues (e.g. [14]). We propose a number of possible extensions to the work presented in this paper.

One simple extension involves the addition of a context to the agent's cost. In other words, given that fact $A, B$ and $C$ are known, we would like to be able to capture the notion that it is cheaper to reveal $D$ and $E$ together than as speech acts at different stages of the dialogue. Without some form of lookahead to allow the agent to plan later moves, this extension is difficult to utilise. Once some form of lookahead exists, the addition of opponent modelling can further enhance the framework. Experimentally, evaluating the effects of various levels of lookahead, as well as different forms of opponent modelling might yield some interesting results.

Currently, we do not differentiate between information which the agent has explicitly committed to, and information that the agent has not yet disagreed with. More concretely, assume that the $CS$ contains the argument $(\top, A)$. If an agent makes use of this argument, perhaps by submitting the argument $(A, B)$, then it is committed to the fact that $A$ is true. If however, it never puts forth arguments making use of the fact, then an opponent cannot know if the agent is actually committed to $A$ or not. We plan to extend our formalism and heuristic to capture this interaction in the near future.

Another extension that emerges from this line of reasoning is the concept of lying. An agent might commit to $A$ to win an argument, even if its knowledge base contains only $\neg A$. How best to deal with this situation is an open question.

The way in which we handle conflicts is also open to debate. At the argumentation framework level, enhancements are required that allow one to present further evidence in support of a literal. By increasing the complexity of the model, methods for retracting literals can be introduced, opening up a whole host of questions at the heuristic level. For example, how does retracting support for a literal influence the information an opponent has of the retracting agent's knowledge base?

## 6 Conclusions

In this paper, we proposed a heuristic for argumentation based on minimising the cost of information revealed to other dialogue participants. While such an argumentation strategy arises in many real world situations, we are not familiar with any application that explicitly makes use of this technique. To study the heuristic, we proposed an argumentation framework that allowed us to focus on it in detail. Several novel features emerged from the interplay between the heuristic and the framework, including the ability of an agent to win an argument that it should not have been able to win (if all information were available to all dialogue participants). While we have only examined a very abstract model utilising the heuristic, we believe that many interesting extensions are possible.

## Acknowledgements

This work is partly funded by the DTI/EPSRC E-Science Core Program and British Telecom, via a grant for the CONOISE-G project (http://www.conoise.org), a multi-university collaboration between Aberdeen, Cardiff and Southampton universities, and BT.

## REFERENCES

[1] Leila Amgoud and Nicolas Maudet, 'Strategical considerations for argumentative agents (preliminary report).', in *NMR*, pp. 399–407, (2002).

[2] Trevor J.M Bench-Capon, 'Specification and implementation of Toulmin dialogue game', in *Proceedings of JURIX 98*, pp. 5–20, (1998).

[3] Carlos Ivan Chesñevar and Ana Gabriela Maguitman, 'Logical models of argument', *ACM Computing Surveys*, **32**(4), 337–383, (2000).

[4] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial Intelligence*, **77**(2), 321–357, (1995).

[5] Antonis C. Kakas, Nicolas Maudet, and Pavlos Moraitis, 'Layered strategies and protocols for argumentation-based agent interaction.', in *ArgMAS*, pp. 64–77, (2004).

[6] Peter McBurney and Simon Parsons, 'Risk agoras: Dialectical argumentation for scientific reasoning', in *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, pp. 371–379, Stanford, USA, (2000).

[7] David Moore, *Dialogue game theory for intelligent tutoring systems*, Ph.D. dissertation, Leeds Metropolitan University, 1993.

[8] John L. Pollock, 'Perceiving and reasoning about a changing world.', *Computational Intelligence*, **14**, 498–562, (1998).

[9] Henry Prakken, 'A study of accrual of arguments, with applications to evidential reasoning', in *Proceedings of the tenth International Conference on Artificial Intelligence and Law*, pp. 85–94, (2005).

[10] Henry Prakken and Giovanni Sartor, 'A dialectical model of assessing conflicting arguments in legal reasoning', *Artificial Intelligence and Law*, **4**, 331–368, (1996).

[11] Henry Prakken and Giovanni Sator, *Computational Logic: Logic Programming and Beyond. Essays In Honour of Robert A. Kowalski, Part II*, volume 2048 of *Lecture Notes in Computer Science*, 342–380, Springer-Verlag, Berlin, 2002.

[12] Henry Prakken and Gerard Vreeswijk, 'Logics for defeasible argumentation', in *Handbook of philosophical logic, 2nd Edition*, eds., D. Gabbay and F. Guenthner, volume 4, 218–319, Kluwer Academic Publishers, (2002).

[13] Guillermo R. Simari and Ronald P. Loui, 'A mathematical treatment of defeasible reasoning and its implementation', *Artif. Intell.*, **53**(2-3), 125–157, (1992).

[14] Katia Sycara, 'Persuasive argumentation in negotiation', *Theory and Decision*, **28**(3), 203–242, (May 1990).

[15] Bart Verheij, 'DefLog: On the logical interpretation of prima facie justified assumptions', *Journal of Logic and Computation*, **13**(3), 319–416, (2003).

[16] Douglas N. Walton, *Legal argumentation and evidence*, Penn State Press, 2002.

[17] Douglas N. Walton and Erik C. W. Krabbe, *Commitment in Dialogue*, State University of New York Press, 1995.