SOURCE SENSITIVE ARGUMENTATION SYSTEM

Chee Fon Chang, Peter Harvey and Aditya Ghose
Decision Systems Lab
University of Wollongong, NSW 2522, Australia

Keywords: Multi-agent Argumentation, Multi-agent Negotiation.

Abstract: There exist many approaches to agent-based conflict resolution. Of particular interest are approaches which adopt argumentation as their underlying conflict resolution machinery. In most argumentation systems, the argument source plays a minimal role. We feel that ignoring this important attribute of human argumentation reduces the capabilities of current argumentation systems. This paper focuses on the importance of sources in argumentation and extends this to the notion of credibility of agents.

1 INTRODUCTION

There have been many recent developments on logical systems for defeasible argumentation and the use of argumentation in multi-agent interaction such as negotiation, group decision making and dispute mediation. Systems such as (Prakken, 1993; Dung, 1995; Bondarenko et al., 1997; Vreeswijk, 1997; Dung and Son, 2001), provides formalisation to defeasible or non-monotonic reasoning. These systems focus on representation and general interaction that exist between arguments. (Jennings et al., 1998; Sierra et al., 1997; Kraus et al., 1998; Parsons et al., 1998), studies argumentation as a conflict resolution machinery of negotiation, where arguments for an offer should persuade the other party to accept the offer. These recent applications of argumentation in multi-agent systems have drawn great interest. If one was to use defeasible argumentation as the conflict resolution machinery, modifications are required.

We view argumentation as begin (simultaneously) a process for information exchange, a process for conflict and an approach to knowledge representation/reason. In (Kraus et al., 1998), the authors stated that argumentation is a process that has the structure of a logical proof however does not hold the strength of one. Consistent with this view, we believe that multi-agent argumentation focuses not on logical "truth" but on convincing/persuading other agents of a particular view or position. As such, how an individual is perceived in their community will effect the acceptance of their arguments.

In the next section, we will motivate this work with a simple example. In section 2, a formalisation of source sensitive argumentation system is provided. Section 3 provides the procedures and worked examples. Section 4 describes the results and behavioural impacts.

1.1 Motivation

We will first motivate our work using an adaptation of an example from (Verheij, 1994) and then extend the example to capture real-life situations. Let's assume that Bill is a juvenile and has committed a crime. We pose a question: "Should Bill be punished?". Assuming that you are given the statements below:

Bill has robbed someone, so he should be jailed.(α)
Bill is a juvenile, therefore he should not go to jail.(β)

In the above statement, if we assume that the second statement is stronger, then we can say that β *defeats* α. However, the above form of argumentation is insufficient in capturing real-life argumentation. Existing argumentation systems (Jennings et al., 1998; Parsons et al., 1998; Sierra et al., 1997; Kraus et al., 1998; Dung, 1995; Verheij, 1994; Pollock, 1991; Vreeswijk, 1997) separate the source of the argument from the argument when evaluating the defeat. We believed that the validity and strength of an argument cannot be capture by the argument alone. It is common in reality for one to evaluate the strength

Fon Chang C., Harvey P. and Ghose A. (2006).
SOURCE SENSITIVE ARGUMENTATION SYSTEM.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - AIDSS*, pages 39-46
DOI: 10.5220/0002464800390046

and validity of arguments with respect to the provider of the argument. Furthermore, we claim that "*undercut*"(Pollock, 1970) (a form of *attack*) cannot be performed without this meta information.

Consider the following modification of the above example:

Tom: *Bill has robbed someone, so he should be jailed.*(α)
Dick: *Bill is a juvenile, therefore he should not go to jail.*(β)

Previously, we assumed that the second statement is stronger and hence we said that β defeats α. If we knew that Tom is a known liar it would further strengthen the acceptability of this *defeat*. However, if you were told or knew that Tom is a police officer we would be less willing to accept the *defeat*. We would most likely rule in favour of Tom.

We will now extend this example further. Let us assume that Tom is a police officer, however over a period of time allegations of wrongful arrest, brutality and violation of arrest procedures against Tom were made by an independent source. These allegations would diminish Tom's credibility. Now, if we were asked to evaluate subsequent arguments between Tom and Dick or to re-evaluate the defeat between α and β the original defeat may no longer hold. From the above examples, it is clear that the defeat relation cannot simply be between arguments.

Proposals by (Carbogim, 2000) capture the affect of the underlying knowledge base by arguments, however the proposal does not deal with changing defeat relation. In this paper, we will show that by simply by tagging arguments with sources, the resulting defeat relation is more fluent and changes the dynamics of argumentation systems.

2 FORMAL DEFINITION

As stated in (Prakken and Vreeswijk, 2002), any defeasible argumentation system consists of several important components: a language, a definition of arguments, a definition of inconsistency/conflict/attack, a defeat relation and finally, the overall status of the argument. We will now provide the components of our system closely matching these key components.

For simplicity, we will take any finitely generated propositional language $\mathcal{L}$ with the usual punctuation signs, and logical connectives $\neg$ (not) and $\supset$ (implies). For any set of wffs S $\subseteq \mathcal{L}$ and any $\alpha \in \mathcal{L}$, S $\vdash \alpha$ means α is provable from premises S. For any set of wffs S $\subseteq \mathcal{L}$, Cn_L(S) = $\{ \alpha \mid \text{S} \vdash \alpha \}$.

Definition 1 (Argument) An argument α is a triple $\langle F,\ A,\ C\rangle$ where $F,\ A$ and C denote the sets of facts, assumptions and conclusion respectively.

Definition 2 (Well-founded Argument) An argument α is a *well-founded* argument iff it satisfies the following conditions:

- $F, A, C \subseteq \mathcal{L}$
- $F \cap A = \emptyset$
- $F \cup A \vdash C$
- $Cn\,(F \cup A \cup C) \not\vdash \bot$

We will write F_α, A_α and C_α to respectively denote the facts, assumptions and conclusions associated with an argument α. One can view a well-founded argument as a *premise-conclusion* pair where the *premise* is $\langle F,\ A\rangle$ and the *conclusion* is C. By allowing assumptions, this generalise the arguments to allow for representation of weak facts. Note that since we are interested in *rational* agents, we have eliminated self-defeating (Prakken and Vreeswijk, 2002; Pollock, 1991) arguments. If one was interested in creating *irrational* agents or paradoxes, one can make the appropriate changes by removing the last condition.

Definition 3 (Conflict) A pair of well-founded arguments β and γ are said to be in *conflict* iff $Cn\,(C_\beta \cup C_\gamma) \vdash \bot$

Definition 4 (Attack) For a pair of well-founded arguments β and γ we say that β *attacks* γ iff:

1. β and γ are in conflict
2. $Cn\,(F_\beta \cup F_\gamma \cup A_\gamma) \vdash \bot$

There exists great variation of the definition for conflict/attack in the defeasible argumentation literature, and are often defined a way that makes them interchangeable. In our system, these notions are instead assigned specific meanings.

Conflict represents an inconsistency between the conclusion and is symmetrical. Attack represents an inconsistency on the premise. It is our view that if the conclusions are non-conflicting then there are no issues to be resolved. Generally, if we are unaware of a difference in opinion, we do not argue about it. Later we will combine these two notions to define *rebuttal*, *assumption attack* and *undercut*. Note that our definitions will differ slightly from those described in (Prakken and Vreeswijk, 2002).

Definition 5 (*Tagged* Arguments) Given a set of unique identifiers $\mathcal{I}$, we define $\mathcal{A}$ as a set of *tagged* arguments of the form $\langle S, A\rangle$ where

- $S \in \mathcal{I}$ represents the *tagged* arguments' *source*.
- A is a set of *well-founded* arguments.

We will write S_ϕ and A_ϕ to respectively denote the source and well-founded argument associated with a tagged argument ϕ.

Definition 6 (Credibility Function) Given a set of unique identifiers $\mathcal{I}$, we say C is a credibility function if it maps all values of $\mathcal{I}$ into a some totally ordered set.

The notion of credibility provides the agent with a measure of strength belief per argument source. This notion also provides a simplistic measure of trust. For simplicity, we have define it as a function that maps a set of unique identifiers into a total ordered set. However, one could define an arbitrary complex measure taking into account of the context in which the argument is situated. For example, if the agent is a stock-broker, then any arguments related to the share market from this agent will be more credible than any agent that is not a stock-broker. One could also define a mapping to an arbiter set as long as there exist operators that are transitive and asymmetric defined over the set.

Definition 7 (Defeat) Given a set of *tagged* arguments $\mathcal{A}$ and the credibility function C, a relation $D \subseteq \mathcal{A} \times \mathcal{A}$ is said to be a *defeat* relation on $\mathcal{A}$. We will write $\phi D \psi$ iff at least one of the following if true:

- A_ϕ attacks A_ψ and A_ψ does not attack A_ϕ
- A_ϕ attacks A_ψ and $C(S_\phi) > C(S_\psi)$
- A_ϕ and A_ψ are in conflict, neither A_ϕ attacks A_ψ nor A_ψ attacks A_ϕ and $C(S_\phi) > C(S_\psi)$

The tagging of arguments allows us to uniquely identify the argument source, and so we make use of this in our definition of defeat. Also, note that our notion of defeat is not defined as a global relation but as a per-agent defeat relation and is determined by a credibility function C (see later definition 8).

Our definition of defeat also encapsulates various types of defeat. For example, (Prakken and Vreeswijk, 2002) states that *assumption attack* occurs when one argument proves what was assumed unprovable by the first (in other words, when a conclusion of one argument attacks the assumption of another). We will say that assumption attack occurs when facts of one argument attacks the assumption of another argument. This is captured by $attack(\phi, \psi) \wedge \neg attack(\psi, \phi)$.

Similarly, (Prakken and Vreeswijk, 2002) states that *rebuttal* occurs when the conclusion of one argument attacks the premise of another. We deviate slightly and say that rebuttal occurs when facts of one argument attack facts of another argument. This is captured by $attack(\phi, \psi) \wedge C(S_\phi) > C(S_\psi)$.

Finally, (Prakken and Vreeswijk, 2002) states that undercut occurs when an argument attacks an inference rule used in another argument. In our system undercut occurs when the conclusion of the arguments contradict but neither the facts nor assumptions of both arguments contradict, as this implies that the contradiction occurs during inference. This notion is capture by $conflict(\phi, \psi) \wedge \neg\left(attack(\phi, \psi) \vee attack(\psi, \phi)\right) \wedge C(S_\phi) > C(S_\psi)$

Definition 8 (Agent) Given a set of unique identifiers $\mathcal{I}$ and a set of tagged arguments $\mathcal{A}$, an agent is represented as a 4-tuple of the form $\langle I,\ A,\ C\rangle$ where

- $I \in \mathcal{I}$.
- $A \subseteq \mathcal{A}$ s.t. $\forall \phi : \phi \in \mathcal{A}$, if $S_\phi = I$ then $\phi \in A$
- C is a credibility function. This represents the credibility (higher values are better) of other agents in the system as evaluated by the agent.

Note that the credibility function is subject to revision by the agent during the execution of the system as each agent adjusts it's view of fellow agents. Note also that the set of tagged arguments A is subjected to change as individual agent discover new arguments during the argumentation process. We do not require the agent to know all arguments nor do we require that the set of arguments be *conflict-free*. The notion of conflict-free set is simply no two arguments in the set defeat each other.

The uniqueness of this system is that there exists no global consensus on the credibility value on each agent. This measure is recorded from individual agent's perspective and is stored by individual agents. Given that there is no requirement for global consensus on individual agent's credibility, consensus on the amount of adjustment on the credibility value is not required either. However, for agents to be productive, we believed that there should be a consensus on when and the kind of adjustment that should be performed. A simple rule would be that if an observation is made of an agent winning an argument, then that agent's credibility should be adjusted upwards and the converse holds. One could also extend this to capture situation importance. For example, if an agent is observed to have won an important argument, then it's credibility is revised upwards by a greater value to that of an argument with less important. Similar to human debates, this provides a notion of a "career making win". This also provides incentive for agents to win the argument. We have provided directions in which one could formula a true-to-life function, we leave the function details to designers.

For convenience we will write D_ρ to denote the defeat relation (from definition 7) as determined by the A and C held by agent ρ.

Definition 9 (Stable) For an agent ρ, we say a set $S \subseteq A$ is a stable set of arguments for that agent iff it is a maximal (w.r.t. set inclusion) set of arguments such that:

- $\forall \psi \in A_\rho - S,\ \exists \phi \in S$ that conflicts with ψ.
- $\forall \phi \in A_\rho,\ \nexists \psi \in S$ where $\psi D_\rho \phi$.

A stable set provides a particular view which the agent has adopted. To select a particular view is to side with some set of agents in the system hence providing support to that group of agents, in essence forming parties or coalitions. We will write $\mathcal{S}_\rho$ to denote a stable set of arguments adopted by an agent ρ.

Definition 10 (Source Sensitive Argumentation System) A source sensitive argumentation system is defined as:

$$\mathcal{SAS} = \langle Agt, \mathcal{A} \rangle$$

where

- Agt is a set of agents.
- $\mathcal{A}$ is a set of tagged arguments.

Definition 11 (Consensus) Given a source sensitive argumentation system $\mathcal{SAS} = \langle Agt, \mathcal{A} \rangle$, we say a set $\mathcal{S} \subseteq \mathcal{A}$ is a *consensus* iff for all agents $\rho \in Agt$, there exist a state S_ρ such that no argument in S_ρ D_ρ an argument in $\mathcal{S}$.

A consensus is a set of arguments that cannot be defeated by any agent.

Definition 12 (G-consensus) Given a source sensitive argumentation system $\mathcal{SAS} = \langle Agt, \mathcal{A} \rangle$ and $G \subseteq Agt$, we say a set $\mathcal{S} \subseteq \mathcal{A}$ is a *G-consensus* iff for all agents $\rho \in G$, there exist a state S_ρ such that no argument in S_ρ D_ρ an argument in $\mathcal{S}$.

We understand that it is generally very hard to reach a consensus. G-consensus can be viewed as group consensus where a predefined set of agents are consulted when determine consensus. So if G is the complete set of the agents in the system then G-consensus is equal to consensus.

Definition 13 (N-acceptable) Given a source sensitive argumentation system $\mathcal{SAS} = \langle Agt, \mathcal{A} \rangle$, we say a set $\mathcal{S}_\mathcal{N} \subseteq \mathcal{A}$ is a *n-acceptable* iff there exists some $G \subseteq Agt$ such that $|G| = n$ and $\mathcal{S}_\mathcal{N}$ is a G-consensus.

With n-acceptable, we simply denote the number of agents that we will seek consensus and not any specify a group.

Definition 14 (Majority) Given a source sensitive argumentation system $\mathcal{SAS} = \langle Agt, \mathcal{A} \rangle$, we say a non-empty set $\mathcal{S}_\mathcal{M} \subseteq \mathcal{A}$ is a *majority* if it is $\frac{|Agt|}{2}$-acceptable.

Majority is simply a set of arguments that more than half the agent population are happy with.

3 PROCEDURE

Informally, an agent's belief consists of a set of arguments, and a credibility ordering derived from the credibility function. Each agent will revise their credibility function as the argumentation process progresses, and so we assume that each agent has a credibility revision operator. We leave the details of the revision operator to the individual agent's designer, though we note that it is possible for an agent to assign other agents higher credibility than itself.

For the purpose of our discussion, we propose the use of the following basic revision rules:

- If an agent is observed to have lost an argument, its credibility is reduced.
- If an agent is observed to have won an argument, its credibility is increased.
- We revise only when a new observation is made, and so do not repeatedly reduce/increase the credibility of an agent.

We note that the above general rules still permit agents to participate in more than one argumentation session.

We will also assume that there exists a communication protocol used by agents to transmit and receive arguments. For simplicity, we will assume that the arguments transmitted are transparent to all agents. As a simplistic protocol to the argument exchange, we suggest that agents transmit their argument in half-duplex mode and synchronously, hence taking turns in presenting their argument or counter-argument. At the end of each round of exchange, individual agents will determine the defeat of relationship between arguments. The agents will then present their outcome to so that the consensus and majority can be computed. Following this simple procedure we will now provide an example of multi-agent dialogue that is in line with our argumentation system.

We introduce three agents to frame our example; Tom, Dick and Harry. Tom (a police officer) holds the argument α, which is 'Bill has robbed someone, so he should be jailed'. Dick will hold the argument β, which is 'Bill is a juvenile, therefore he should not go to jail'. Harry holds no argument yet, but will participate in the argumentation process. Therefore initially:

Beliefs At Start	
Tom	C(Tom) $>$ C(Dick) $>$ C(Harry)
	$\{\alpha\}$ is the outcome
Dick	C(Dick) $>$ C(Tom) $>$ C(Harry)
	$\{\beta\}$ is the outcome
Harry	C(Harry) $>$ C(Tom) $>$ C(Dick)
	$\{\}$ is the outcome

At this initial point there is no consensus between the agents, and so $\mathcal{S} = \emptyset$. Similarly, there is no majority, and so $\mathcal{S}_\mathcal{M} = \emptyset$. Note that Harry sees Tom (the police officer) as more credible than Dick.

We assume now that Tom and Dick communicate their arguments to the other agents, and so all agents know the arguments $\langle \text{Tom}, \alpha \rangle$ and $\langle \text{Dick}, \beta \rangle$. This gives the following situation:

Beliefs After Communicating Arguments	
Tom	$C(\text{Tom}) > C(\text{Dick}) > C(\text{Harry})$
	$\langle \text{Tom}, \alpha \rangle$ defeats $\langle \text{Dick}, \beta \rangle$
	$\{\alpha\}$ is the outcome
Dick	$C(\text{Dick}) > C(\text{Harry}) > C(\text{Tom})$
	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$\{\beta\}$ is the outcome
Harry	$C(\text{Harry}) > C(\text{Tom}) > C(\text{Dick})$
	$\langle \text{Tom}, \alpha \rangle$ defeats $\langle \text{Dick}, \beta \rangle$
	$\{\alpha\}$ is the outcome

There remains no consensus between the agents, and so $\mathcal{S} = \emptyset$. However, there is now a majority, and so $\mathcal{S}_{\mathcal{M}} = \{\langle \text{Tom}, \alpha \rangle\}$.

We now consider what is to happen if the reputation of Tom was tarnished. Formally, this would mean that Harry changes his assessment of the credibility of each agent so that $C(\text{Harry}) > C(\text{Dick}) > C(\text{Tom})$. Note that this could happen through some other argumentation process involving Dick and Harry.

Beliefs After Decreasing Tom's Credibility	
Tom	$C(\text{Tom}) > C(\text{Dick}) > C(\text{Harry})$
	$\langle \text{Tom}, \alpha \rangle$ defeats $\langle \text{Dick}, \beta \rangle$
	$\{\alpha\}$ is the outcome
Dick	$C(\text{Dick}) > C(\text{Harry}) > C(\text{Tom})$
	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$\{\beta\}$ is the outcome
Harry	$C(\text{Harry}) > C(\text{Dick}) > C(\text{Tom})$
	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$\{\beta\}$ is the outcome

Finally, we consider the situation where Tom himself is 'ashamed' and lowers his own credibility below that of Dick's. As shown below, a consensus is established, with all agents agreeing on β. From this example we can see that consensus will only be established if there was no disagreement, or one agent 'gives ground' and lowers their own credibility. We will explore this more completely below.

Beliefs After Tom 'Gives Ground'	
Tom	$C(\text{Dick}) > C(\text{Tom}) > C(\text{Harry})$
	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$\{\beta\}$ is the outcome
Dick	$C(\text{Dick}) > C(\text{Harry}) > C(\text{Tom})$
	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$\{\beta\}$ is the outcome
Harry	$C(\text{Harry}) > C(\text{Dick}) > C(\text{Tom})$
	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$\{\beta\}$ is the outcome

3.1 Cycles

There exists two forms of cycles. Firstly, cycles on the defeat relation between arguments. This is caused by agents ranking themselves as definitive source of credibility. Below is an example of a cycle.

Cycle in Defeat Relation	
Tom	$\langle \text{Tom}, \alpha \rangle$ defeats $\langle \text{Dick}, \beta \rangle$
	$C(\text{Tom}) > C(\text{Dick}))$
	$\{\alpha\}$ is the outcome
Dick	$\langle \text{Dick}, \beta \rangle$ defeats $\langle \text{Tom}, \alpha \rangle$
	$C(\text{Dick}) > C(\text{Tom})$
	$\{\beta\}$ is the outcome

This form of cycle results in no consensus and, as in this case, may result in no majority. This form of cycle is not a serious issue as it indicates to the operators that no agreement can be reached, but does not hinder the termination of the program. Note also that we do expect such cycles to occur, as they result from underlying cycles in the attacks relation.

The second form of cycle occurs during the computation of consensus and majority. This is caused by agents inability to commit to a particular credibility ordering. This is directly attributed to a bad revision operator. We will term this type of cycle *oscillation*. For the purpose of this discussion, we will assume that the revision operators defined by designers are well behaved. With a well-behaved revision operator, cycles of random raising and lowering credibility never occur. Furthermore, after winning an argument an agent is more likely to win future arguments. This self-supporting nature further ensures that cycles of raising and lowering credibility never occur. This ensures that a fix-point is achieved, and hence a termination condition for the program exists.

4 RESULTS AND BEHAVIOURAL IMPACTS

This section will outline the contributions made by this system. Agent behaviours resulting from this system falls into the categories of: agent reasoning, agent autonomy and group behaviours.

4.1 Agent Reasoning

In (Jennings et al., 1998; Sierra et al., 1997; Parsons et al., 1998), the authors proposed the use of argumentation to assist in agent negotiation. Information that are naturally embedded in arguments are use to determine or eliminate infeasible proposals. With the additional notion of credibility, our system further assist in this process. With an embedded defeat relation, agents are in the position to evaluate their own arguments before communicating it to other agents. This creates the notion of a *minimal* winning argument. A minimal winning argument increases its' strength by minimising the possible attack on it. For example, by reducing the number of assumptions in an argument, minimises the possible assumption attack. The notion of a minimal argument allows for the arguments to be ranked according to strength and provides an order in which arguments are to be presented to fellow agents. The introduction of credibility also introduces a notion of self-preservation. If an agent wishes to be in a position to win future arguments, then it must keep its' credibility value high. The self-preservation notion forces the agent into evaluate arguments more carefully so as not to randomly propose weak arguments. Essentially agents be will force into a position in which they only "fight in battles that they believe they can win". This improves the efficiency of the argumentation process.

Additionally, agents are now equipped with additional rationale. The use of credibility provides another means to support their decisions.

4.2 Agent Autonomy

The lack of a global defeat relation and consensus on the credibility provides this system the ability to distribute and scale. This feature do away with the need for a central repository/location where conflicts are resolved. This system also allows for arguments to propagate hence conflicts can be isolated and resolved independently. This approach also removes the notion of a static top argument.

4.3 Agent Group Behaviour

Group decision making and coalition formation have received huge amount of interest recently. Our system provides agents with the ability to support fellow agents. This is archived through the consensus and majority operation. If an agent's argument is deem strong, it would persuade fellow agents of its' view. This exercise causes parties to form and hence coalition to form leading to dynamic coalition formation.

5 RELATED WORKS

This section will outline the related works. The related works falls into three categories: defeasible argumentation, argumentation-based negotiation and distributed constraints satisfaction.

5.1 Defeasible Argumentation

Although the works such as (Dung, 1995; Prakken, 1993; Bondarenko et al., 1997; Dung and Son, 2001; Dung, 1995; Prakken, 1993; Vreeswijk, 1997) focuses on representation and logical reasoning, we have shown that by introducing the notion of sources and credibility as well as a corresponding notion of defeat, we are able to provide a mapping for a theoretical to a practical system. These systems provide a strong basis for our proposal. We have modelled our system around these collection of defeasible argumentation systems. We feel that our work will complement those advances already achieved.

Additionally, recent work (Prakken, 2001; Carbogim, 2000; Brewka, 2001) in defeasible argumentation have focused on the notion of dynamic argumentation. (Prakken, 2001) focus on the fairness and soundness of dynamic argumentation protocol. (Carbogim, 2000) focused on addressing issues of associated to change in the underlying knowledge base caused by arguments. (Brewka, 2001) deals with meta-level argumentation, providing undercut via a notion of preference on which defeat rule holds. Components of these work have similarities to ours, however the notion of sources and credibility forcing a change of preference on the defeat rules are not investigated. In fact these works do not have explicit notion of sources and credibility.

5.2 Argumentation-Based Negotiation

The exchange of arguments and counterarguments has also been studied in the context of multi-agent interaction. In (Kraus et al., 1998) and (Parsons et al., 1998) have studied argumentation as a component of negotiation protocols, where arguments for an offer should persuade the other party to accept the offer. Generally, these works focus on the negotiation protocol and agent behaviours, leaving the representation and internal reasoning to the designer. In (Kraus et al., 1998), a range of agents and behaviours were prescribed. We feel that our work further enhance the internal reasoning of these agents. Although some notion of sources and credibility may have existed in (Jennings et al., 1998; Sierra et al., 1997), these notions were not explicitly utilised when evaluating arguments.

5.3 Distributed Constraint Satisfaction

Recent work in distributed constraint satisfaction algorithms (Harvey et al., 2005b; Harvey et al., 2005a; Harvey et al., 2006) is built upon the theoretical underpinnings described in this paper. Support-Based Distributed Search (SBDS) is a distributed constraint satisfaction algorithm in which agents communicate via arguments, maintaining a simple notion of credibility between agents.

The argument structure of SBDS is domain specific, permitting two categories of arguments. As it deviates from the argumentation structure representation in this paper (with distinction between facts and assumptions). Below is a loose description of the argument structure below in SBDS.

Definition 15 (SBDS Argument) An SBDS-argument is a pair $\langle Prem, Con\rangle$, and belongs to one of the following two categories:

1. isgoods (variable-value assignment proposals)
   - *Prem* - an ordered sequence of variable-value assignments
   - *Con* - the variable-value assignment for the agent stating the argument
2. nogoods (variable-value assignment rejections)
   - *Prem* - a set of variable-value assignments which are not permitted
   - *Con* - exact copy of the premise

As the argument structure of SBDS differ from the monotonic logic example given in this paper, a domain-specific conflict and attacks relation was defined. The spirit of these relation remains the same as those provided in our work. This further strengthens our claim that our proposal complements and provides assistance to solving problems in a whole range of different domains.

6 CONCLUSION

In most argumentation systems, the source of the argument plays a minimal role. In this paper, we have introduced the notion of sources and credibility. We have shown that by simply tagging arguments with sources, the resulting defeat relation is more fluent and the dynamics of argumentation systems changes. We have shown that ignoring this important attribute of human argumentation reduces the capabilities of current argumentation systems. This paper focuses on the importance of information source in argumentation and extending this to the notion of credibility. The notion of credibility plays an important role in the agent decision making process during argumentation. We have also shown that the system is capable of emulating feature of human argumentation that was not captured in existing systems. Finally, we have shown that this system is not purely theoretical but can also be applied into the practical domain such as argumentation-based negotiation and distributed constraint satisfaction.

6.1 Future Works

Throughout this paper, we have indicated room for improvements. We point to improvements that can be made on mapping sources to credibility by augmenting the function to consider context. We also like to point out that the current mapping may not be satisfactory to some audience. Suggestion of generalising the mapping function via the use of semi-ring structure, combining two orthogonal metric (one measuring strength of the argument, the other measuring the credibility) are currently under investigation. By using this approach, we are in the position to provide a notion of graded defeat. This modification would provide a method to infer the global credibility and defeat relation from a given set of agents.

REFERENCES

Bondarenko, A., Dung, P. M., Kowalski, R. A., and Toni, F. (1997). An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101.

Brewka, G. (2001). Dynamic argument systems: A formal model of argumentation processes based on situation calculus. *Journal of Logic and Computation*, 11.

Carbogim, D. V. (2000). *Dynamics in Formal Argumentation*. PhD thesis, University of Edinburgh. College of Science and Engineering. School of Informatics.

Dung, P. M. (1995). An argumentation-theoretic foundations for logic programming. *Journal of Logic Programming*, 22(2):151–171.

Dung, P. M. and Son, T. C. (2001). An argument-based approach to reasoning with specificity. *Artificial Intelligence*, 133(1-2):35–85.

Harvey, P., Chang, C. F., and Ghose, A. (2005a). Practical application of support based distributed search. In *Proceedings of the 17th IEEE International Conference on Tools with AI*.

Harvey, P., Chang, C. F., and Ghose, A. (2005b). Support-based distributed search. In Meisels, A., editor, *Proceedings of the 6th International Workshop on Distributed Constraint Reasoning*. http://www.dsl.uow.edu.au/people/harvey/dcr05.pdf.

Harvey, P., Chang, C. F., and Ghose, A. (2006). Support-based distributed search. In *Proceedings of the 5th International Joint Conference on Autonomous Agents & Multi Agent Systems*.

Jennings, N. R., Parsons, S., Noriega, P., and Sierra, C. (1998). On argumentation-based negotiation. In *Proceedings of the International Workshop on Multi-Agent Systems*, Boston, USA.

Kraus, S., Sycara, K. P., and Evenchik, A. (1998). Reaching agreements through argumentation: A logical model and implementation. *Artificial Intelligence*, 104(1-2):1–69.

Parsons, S., Sierra, C., and Jennings, N. R. (1998). Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292.

Pollock, J. L. (1970). The structure of epistemic justification. *American Philosophical Quarterly*, monograph series 4:62–78.

Pollock, J. L. (1991). Self-defeating arguments. *Minds and Machines*, 1(4):367–392.

Prakken, H. (1993). An argumentation framework in default logic. *Annals of Mathematics and Artificial Intelligence*, 9(1-2):93–132. Bibsource = DBLP,http://dblp.uni-trier.de.

Prakken, H. (2001). Relating protocols for dynamic dispute with logics for defeasible argumentation. *Synthese*, 127(1):187–219.

Prakken, H. and Vreeswijk, G. (2002). *Logics for Defeasible Argumentation*, volume 4 of *Handbook of Philosophical Logic*, pages 218–319. Kluwer Academic Publishers, second edition edition.

Sierra, C., Jennings, N. R., Noriega, P., and Parsons, S. (1997). A framework for argumentation-based negotiation. In Singh, M. P., Rao, A. S., and Wooldridge, M., editors, *Proceedings of the 4th International Workshop on Agent Theories, Architectures, and Languages*, pages 177–192.

Verheij, B. H. (1994). Reason based logic and legal knowledge representation. In Carr, I. and Narayanan, A., editors, *Proceedings of the 4th National Conference on Law, Computers and Artificial Intelligence*, pages 154–165. University of Exeter.

Vreeswijk, G. (1997). Abstract argumentation systems. *Artificial Intelligence*, 90(1-2):225–279.