Conflict resolution with argumentation dialogues

(Extended Abstract)

Xiuyi Fan, Francesca Toni
Department of Computing, Imperial College London, UK
{x.fan09,ft}@imperial.ac.uk

ABSTRACT

Conflicts exist in multi-agent systems for a number of reasons: agents have different interests and desires; agents hold different beliefs; agents make different assumptions. To resolve conflicts, agents need to better convey information to each other and facilitate fair negotiations yielding jointly agreeable outcomes. We present a two-agent, dialogical conflict resolution scheme developed with the Assumption-Based Argumentation (ABA) framework.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Algorithms

Keywords

Argumentation, Collective Decision Making

1. INTRODUCTION

In this paper, we study conflict resolution in multi-agent systems [3]. We use ABA [2] to represent agents' beliefs and desires. ABA is a general-purpose argumentation framework where arguments are built from *rules* and supported by *assumptions*, and attacks against arguments are directed at the assumptions supporting the arguments, and are provided by arguments for *contraries* of assumptions. Sentences in rules, assumptions and contraries form the underlying *language*. A *claim* is *admissible* iff it is supported by an argument that is in a set of arguments which does not attack itself and counter-attacks all attacks against the set.

In our approach, conflicts are given by different desires, seen as realizations of the same goal. To resolve conflicts between *two* agents is to have dialogues. Through dialogues, agents eliminate misunderstandings by acquiring information from each other. (Sequences of) Successful dialogues allow to identify shared desires and resolve conflicts.

2. MOTIVATING EXAMPLE

Two agents, Jenny (**J**) and Amy (**A**), are planning a film night together. They want to agree on the movie to watch. *Lord of the Rings* (*LoR*) and *Terminator* (*Ter*) are both screening. **J** wants to pick a fun movie. She finds action movies fun. **J** believes *Ter* is an action movie. She does not know much about *LoR*. **J** wants to watch *Ter*. **A** also wants to watch a fun movie. However, **A** thinks fantasy movies are fun. **A** has watched the trailer of *LoR* and believes it is both an action and a fantasy movie. **A** concludes she wants to watch *LoR*. After exchanging information, **J** agrees.

This example can be modelled in terms of two dialogue processes, each consisting of two phases. The first dialogue process is about *Ter*. Here, *Phase I* amounts to the following:

J: Let's see if *Ter* is a good movie to watch.
A: OK.
J: I'll watch *Ter* if it is fun and there is no objection to it.
A: OK.
J: *Ter* would be fun if it is an action movie.
A: OK.
J: Yes, *Ter* is an action movie.
A: OK.
J: I propose we watch *Ter* then.
A: We can watch it unless it has been watched before.
J: OK, it has not.
A: OK.

Since *Ter* satisfies **J**, we move to *Phase II* in which *Ter*'s acceptability with respect to **A** is examined. Now, **A** starts the dialogue and the two agents proceed similarly to the previous dialogue, except this time **A** believes fantasy movies are fun and *Ter* is not a fantasy movie. Hence the dialogue fails.

Since *Ter* is rejected by **A**, the two agents move to the next realization, *LoR*. Using a similar two-phase dialogue process, they find that *LoR* satisfies them both and thus is a conflict resolution.

3. METHODOLOGY

We define agents as equipped with ABA frameworks whose rules are of one of two types: *concession rules* and *non-concession rules*. Non-concession rules ($\mathcal{R}^{NC}$) describe agents' desires, which are strictly firm. Concession rules ($\mathcal{R}^{C}$) describe factual information about the agents' environment, agents' beliefs and agents' desires which can be conceded. Both types of rules may be defeasible or not. However, non-concession rules may be defeasible solely based on an agent's own will. A conflict resolution satisfies all non-concession desires of agents, under the condition that both agents are aware of the other agent's relevant beliefs.

The ABA frameworks of **J** and **A** are in Table 1[1]. The argument in Figure 1 (Left) can be built from the ABA framework of **J**. The claim of this argument is wM(*Ter*). The support of this argument

[1]wM, sM, aM, and fM stand for watchMovie, selectMovie, actionMovie and fantasyMovie, respectively. X and Y are (universally quantified) variables.

Cite as: Conflict resolution with argumentation dialogues (Extended Abstract), Xiuyi Fan, Francesca Toni, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1095-1096.

$\mathcal{R}^{NC}$: wM(X) $\leftarrow$ fun(X), sM(X) (**J, A**)
fun(X) $\leftarrow$ aM(X) (**J**)
fun(X) $\leftarrow$ fM(X) (**A**)
$\mathcal{R}^{C}$: aM(*Ter*) (**J, A**)
fM(*LoR*) (**A**)
Assumptions: sM(X) (**J, A**)
Contraries: $\mathcal{C}$(sM(X)) = {¬ sM(X), sM(Y)| Y $\neq$ X } (**J, A**)

Table 1: ABA frameworks for J and A in the example.

is the assumption sM(*Ter*), and corresponds to selecting the movie *Ter*. Attacks against this argument are arguments for a contrary of this assumption, namely for an element of $\mathcal{C}$(sM(*Ter*)) (no such argument is found in the example). In this example agents have different rules but the same assumptions and contraries. In general, agents may hold different rules, assumptions, and contraries, but will always share the same underlying language $\mathcal{L}$.

We define a *conflict* between two agents a_1 and a_2 (equipped with ABA frameworks AF_1 and AF_2 respectively) with respect to a *goal*, $\mathcal{G}$, as a pair of *realizations* $(\mathcal{G}\delta_1, \mathcal{G}\delta_2)$ such that $\mathcal{G}\delta_1$ and $\mathcal{G}\delta_2$ are admissible claims with respect to AF_1 and AF_2, respectively. In our example, the goal is wM(X), where X is an (implicitly) existentially quantified variable, and the realizations are wM(*Ter*) (for a_1=**J**), and wM(*LoR*) (for a_2=**A**). In general, the goal $\mathcal{G}$ is of the form p(X), where X is a vector of (implicitly) existentially quantified variables, and a realization is of the form $\mathcal{G}\delta$ such that $\delta = \{\mathrm{X}/t\}$, for a vector of terms t, and $\mathcal{G}\delta = \mathrm{p}(t)$ is in $\mathcal{L}$.

We define a *conflict resolution* as a realization, $\mathcal{G}\delta$, such that $\mathcal{G}\delta$ is an admissible claim with respect to AF'_1 and AF'_2, where AF'_x is AF_x with all concession rules from AF_y, for $x, y = 1, 2$ and $x \neq y$. In our example, wM(*LoR*) is a conflict resolution.

We define a *dialogue, $D^{a_j}_{a_i}(s)$, between agents a_i and a_j* (where $i, j = 1, 2$, $i \neq j$) *for a claim s* as a finite sequence of utterances of the form $\langle a_x, a_y, TID, C, ID\rangle$ (where $x, y = 1, 2$, $x \neq y$), in which a_x is the maker and a_y the receiver of the utterance, ID is its identifier, TID is the identifier of the *target* utterance, and C is the content, namely one of (1) a claim, (2) a rule, (3) an assumption, (4) a contrary, (5) π, which represents a *pass*. In $D^{a_j}_{a_i}(s)$, a_i makes the first utterance and $s \in \mathcal{L}$. For two utterances u_k and u_l in a dialogue, if the ID in u_k is the TID in u_l, then u_l is *related to* u_k such that one of two cases holds: (1) the content C_k of u_k is the parent of the content C_l of u_l in an argument; or (2) C_k is an assumption and C_l introduces a contrary of this. A dialogue ends by both agents uttering π consecutively. The informal dialogue in our earlier example can be formalised as in Table 2.

Dialogues are defined in terms of legal-move functions, to determine which utterances agents are allowed to make, and outcome functions, to determine whether dialogues satisfy certain proper-

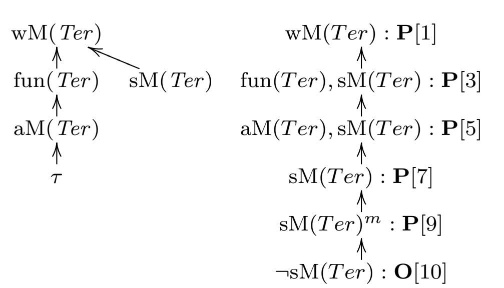

Figure 1: Argument, by J, for watching *Ter* (Left). Dialectical tree for the dialogue in our example (Right).

$\langle J, A, 0, clm(\mathrm{wM}(Ter)), 1\rangle$
$\langle A, J, 0, \pi, 2\rangle$
$\langle J, A, 1, rl(\mathrm{wM}(Ter) \leftarrow \mathrm{fun}(Ter), \mathrm{sM}(Ter)), 3\rangle$
$\langle A, J, 0, \pi, 4\rangle$
$\langle J, A, 5, rl(\mathrm{fun}(Ter) \leftarrow \mathrm{aM}(Ter)), 5\rangle$
$\langle A, J, 0, \pi, 6\rangle$
$\langle J, A, 7, rl(\mathrm{aM}(Ter)), 7\rangle$
$\langle A, J, 0, \pi, 8\rangle$
$\langle J, A, 3, asm(\mathrm{sM}(Ter)), 9\rangle$
$\langle A, J, 5, ctr(\mathrm{sM}(Ter), \neg\mathrm{sM}(Ter)), 10\rangle$
$\langle J, A, 0, \pi, 11\rangle$
$\langle A, J, 0, \pi, 12\rangle$

Table 2: Dialogue in our example.

ties. These functions are defined in such a way that the *dialectical tree* underlying a *successful dialogue* corresponds to a *concrete dispute tree*, as given in [1], with respect to the *ABA framework drawn* from the dialogue. This consists of the rules, assumptions and contraries uttered in the dialogue. The dialogue in Table 2 is successful. The dialectical tree for this dialogue is in Figure 1(Right). The ABA framework drawn from this dialogue consists of

Rules: wM(X) $\leftarrow$ fun(X), sM(X)
fun(X) $\leftarrow$ aM(X)
aM(*Ter*)
Assumptions: sM(X)
Contraries: $\mathcal{C}$(sM(X)) = {¬ sM(X)}

The correspondence between dialectical trees and concrete dispute trees gives, directly from corollary 6.1 in [1], that the claim of a successful dialogue is admissible with respect to the ABA framework drawn from the dialogue.

We define a *conflict resolution dialogue between a_i and a_j for a realization $\mathcal{G}\delta$* as a dialogue $D^{a_i}_{a_j}(\mathcal{G}\delta)$. Here, the agent starting the dialogue, a_i, is the *nominator*, whereas the other agent is the *challenger*. Through the dialogue, the nominator is allowed to utter any rules from its ABA framework, whereas the challenger is only allowed to utter its concession rules.

We define a *successful sequence between a_i and a_j with respect to a goal $\mathcal{G}$* as a sequence $\langle d_1 = D^{a_i}_{a_j}(\mathcal{G}\delta_1), d_2 = D^{a_j}_{a_i}(\mathcal{G}\delta_1), \ldots, d_{2n-1} = D^{a_i}_{a_j}(\mathcal{G}\delta_n), d_{2n} = D^{a_j}_{a_i}(\mathcal{G}\delta_n)\rangle$, for $n \geq 2$, such that both d_{2n-1} and d_{2n} are successful and for all for all $k < n$ either d_{2k-1} or d_{2k} is not successful. Then the following result holds:

THEOREM 3.1. *Given a conflict $(\mathcal{G}\delta_1, \mathcal{G}\delta_2)$ between a_1 and a_2 with respect to some goal $\mathcal{G}$, a conflict resolution $\mathcal{G}\delta$ exists if there is a successful sequence between a_1 and a_2 with respect to $\mathcal{G}$.*

The successful sequence in our example consists of four conflict resolution dialogues, $d_1 = D^J_A(\mathrm{wM}(Ter))$, $d_2 = D^A_J(\mathrm{wM}(Ter))$, $d_3 = D^J_A(\mathrm{wM}(LoR))$, $d_4 = D^A_J(\mathrm{wM}(LoR))$. All except d_2 are successful. d_1 is in Table 2. The other dialogues are omitted for lack of space.

4. REFERENCES

[1] P. Dung, R. Kowalski, and F. Toni. Dialectic proof procedures for assumption-based, admissible argume ntation. *Artificial Intelligence*, 170:114–159, 2006.

[2] P. M. Dung, R. A. Kowalski, and F. Toni. Assumption-based argumentation. In *Argumentation in Artificial Intelligence*, pages 25–44. Springer, 2009.

[3] C. Tessier, L. Chaudron, and H. Müller, editors. *Conflicting agents: conflict management in multi-agent systems*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.