Settling on the Group's Goals: An n-Person Argumentation Game Approach

Duy Hoang Pham[1,2], Subhasis Thakur[1], Guido Governatori[2]

[1] School of Information Technology and Electrical Engineering
The University of Queensland, Brisbane, Australia
{pham,subhasis}@itee.uq.edu.au
[2] National ICT Australia, Queensland Research Laboratory
guido.governatori@nicta.com.au

Abstract. Argumentation games have been proved to be a robust and flexible tool to resolve conflicts among agents. An agent can propose its explanation and its goal known as a claim, which can be refuted by other agents. The situation is more complicated when there are more than two agents playing the game.
We propose a weighting mechanism for competing premises to tackle with conflicts from multiple agents in an n-person game. An agent can defend its proposal by giving a counter-argument to change the "opinion" of the majority of opposing agents. During the game, an agent can exploit the knowledge that other agents expose in order to promote and defend its main claim.

1 Introduction

In multi-agent systems, there are several situations requiring a group of agents to settle on common goals despite each agent's pursuit of individual goals which may conflict with other agents. To resolve the conflicts, an agent can argue to convince others about its pursued goal and provides evidence to defend its claim. This interaction can be modelled as an argumentation game [1,2,3]. In an argumentation game, an agent can propose an explanation for its goal (i.e., an argument), which can be rejected by counter-evidence from other agents. This action can be iterated until an agent either successfully argues its proposal against other agents or drops its initial claim.

The argumentation game approach offers a robust and flexible tool to resolve conflicts by evaluating the status of arguments from agents. Dung's argumentation semantics [4] is widely recognised to establish relationships (undercut, defeated, and accepted) among arguments. The key notion for a set of arguments is whether a set of arguments is self-consistent and provides the basis to derive a conclusion.

An argumentation game is more complicated when the group has more than two agents. It is not clear how to extend existing approaches to resolve conflicts from multiple agents, especially when agents have equal weight. In this case, the problem amounts to deciding which argument has precedence over competing arguments. The main idea behind our approach is the global collective preference over individual proposals, which enables an agent to identify the key arguments and premises from opposing agents in order to generate counter-arguments. These arguments cause a majority of opposing

11th Pacific Rim International Conference on Multi-Agents (PRIMA 2008)
Hanoi, 15–16 December 2008, 2008.
© Springer 2008.
The original publication is available at www.springerlink.com.

agents to reconsider their claims, therefore, an agent has an opportunity to change "attitudes" of others.

Each of our agents is equipped with its private knowledge, background knowledge, and knowledge obtained from other agents. The background knowledge, commonly shared by the group, presents the expected behaviours of a member of the group. Any argument violating the background knowledge is not supported by the group. The knowledge about other agents, growing during the game, enables an agent to efficiently convince others about its own goal. Defeasible logic is chosen as our underlying logic for the argumentation game due to its efficiency and simplicity in representing incomplete and conflicting information. Furthermore, the logic has a powerful and flexible reasoning mechanism [5,6] which enables our agents to flawlessly capture Dung's argumentation semantics by using two features of defeasible reasoning, namely the ambiguity propagating and ambiguity blocking. Our paper is structured as follows. In section 2, we briefly introduce notions of defeasible logic and the construction of the argumentation semantics. Section 3 introduces our n-person argumentation game framework using defeasible logic. We present firstly the external model of agents' interaction, which describes a basic procedure for an interaction between agents. Secondly, we define the internal model, which shows how an agent can deal with individual knowledge sources to propose and defend its goal against other agents. Finally, we show the justification of arguments generated by an agent during the game w.r.t. the background knowledge of the group. Section 4 provides an overview of research works related to our approach. Section 5 concludes the paper.

2 Background

In this section, we briefly present essential notions of defeasible logic (DL) and the construction of Dung's argumentation semantics by using two features of defeasible reasoning including ambiguity blocking and propagating.

2.1 Defeasible Logic

Following the presentation in [7], a defeasible theory D is a triple $(F,R,>)$ where F is a finite set of facts, R is a finite set of rules, and $>$ is a superiority relation on R. The language of defeasible logic consists of a finite set of literals, l, and their complement $\sim l$.

A rule r in R is composed of an antecedent (body) $A(r)$ and a consequent (head) $C(r)$, where $A(r)$ consists of a finite set of literals and $C(r)$ contains a single literal. $A(r)$ can be omitted from the rule if it is empty. There are three types of rules in R, namely R_s (strict rules), R_d (defeasible rules), and R_{dft} (defeaters).

A conclusion derived from the theory D is a tagged literal and is categorised according to how the conclusion can be proved:

- $+\Delta q$: q is definitely provable in D.
- $-\Delta q$: q is definitely unprovable in D.
- $+\partial q$: q is defeasibly provable in D.

– $-\partial q$: q is defeasibly unprovable in D.

The provability is based on the concept of a derivation (or proof) in $D = (F, R, >)$. Informally, definite conclusions can be derived from strict rules by forward chaining, while defeasible conclusions can be obtained from defeasible rules iff all possible "attacks" are rebutted due to the superiority relation or defeater rules. A derivation is a finite sequence $P = (P(1), \dots, P(n))$ of tagged literals satisfying proof conditions (which correspond to inference rules for each of the four kinds of conclusions). $P(1..i)$ denotes the initial part of the sequence P of length i. In the follows, we present the proof conditions for definitely and defeasibly provable conclusions[1]:

$+\Delta$: If $P(i+1) = +\Delta q$ then
  (1) $q \in F$ or
  (2) $\exists r \in R_s[q]\ \forall a \in A(r) : +\Delta a \in P(1..i)$

$+\partial$: If $P(i+1) = +\partial q$ then either
(1) $+\Delta q \in P(1..i)$ or
  (2.1) $\exists r \in R_{sd}[q]\ \forall a \in A(r) : +\partial a \in P(1..i)$ and
  (2.2) $-\Delta{\sim}q \in P(1..i)$ and
  (2.3) $\forall s \in R_{sd}[{\sim}q]$ either
    (2.3.1) $\exists a \in A(s) : -\partial a \in P(1..i)$ or
    (2.3.2) $\exists t \in R_{sd}[q]$ such that $t > s$ and
      $\forall a \in A(t) : +\partial a \in P(1..i)$

The set of conclusions of a defeasible theory is finite[2], and it can be computed in linear time [10].

DL can be extended by an ambiguity propagating variant [11,5]. The superiority relation is not considered in the inference process. Inference with the ambiguity propagation introduces a new tag Σ, a positive support for a literal $+\Sigma q$ is defined as:

$$+\Sigma\text{: If } P(i+1) = +\Sigma q \text{ then } \exists r \in R_{sd}[q]\text{: } \forall a \in A(r) : +\Sigma a \in P(1..i)$$

$+\Sigma p$ means p is supported by the defeasible theory and there is a monotonic chain of reasoning that would lead us to conclude p in the absence of conflicts. A literal that is defeasibly provable ($+\partial$) is supported, but a literal may be supported even though it is not defeasibly provable. Thus support is a weaker notion than defeasible provability.

2.2 Argumentation Semantics

In what follows, we briefly introduce the basic notions of an argumentation system using defeasible reasoning. We also present the acceptance of an argument w.r.t. Dung's semantics.

---

[1] For a full presentation and proof conditions of DL and its properties refer to [8,9].

[2] It is the Herbrand base that can be built from the literals occurring in the rules and the facts of the theory

Definition 1. *An argument A for a literal p based on a set of rules R is a (possibly infinite) tree with nodes labelled by literals such that the root is labelled by p and for every node with label h:*

1. *If $b_1,\ldots,b_n$ label the children of h then there is a rule in R with body $b_1,\ldots,b_n$ and head h.*
2. *If this rule is a defeater then h is the root of the argument.*
3. *The arcs in a proof tree are labelled by the rules used to obtain them.*

DL requires a more general notion of proof tree that admits infinite trees, so that the distinction is kept between an infinite un-refuted chain of reasoning and a refuted chain. Depending on the rules used, there are different types of arguments.

- A supportive argument is a finite argument in which no defeater is used.
- A strict argument is an argument in which only strict rules are used.
- An argument that is not strict, is called defeasible

Relationships between two arguments A and B are determined by literals constituting these arguments. An argument A *attacks* a defeasible argument B if a literal of A is the complement of a literal of B, and that literal of B is not part of a strict sub-argument of B. A set of arguments $\mathscr{S}$ attacks a defeasible argument B if there is an argument A in $\mathscr{S}$ that attacks B.

A defeasible argument A is *undercut* by a set of arguments $\mathscr{S}$ if $\mathscr{S}$ supports an argument B attacking a proper non-strict sub-argument of A. An argument A is undercut by $\mathscr{S}$ means some literals of A cannot be proven if we accept the arguments in $\mathscr{S}$.

The concepts of the attack and the undercut concern only defeasible arguments and sub-arguments. A defeasible argument is assessed as valid if we can show that the premises of all arguments attacking it cannot be proved from the valid arguments in $\mathscr{S}$. The concepts of provability depend on the method used by the reasoning mechanism to tackle ambiguous information. According to the features of the defeasible reasoning, we have the definition of acceptable arguments (definition 2).

Definition 2. *An argument A for p is acceptable w.r.t. a set of arguments $\mathscr{S}$ if A is finite, and*

1. *If reasoning with the ambiguity propagation is used: (a) A is strict, or (b) every argument attacking A is attacked by $\mathscr{S}$.*
2. *If reasoning with the ambiguity blocking is used: (a) A is strict, or (b) every argument attacking A is undercut by $\mathscr{S}$.*

The status of an argument is determined by the concept of acceptance. If an argument can resist a reasonable refutation, this argument is justified. If an argument cannot overcome attacks from other arguments, this argument is rejected. We define the sets of justified arguments as follows:

Definition 3. *Let D be a defeasible theory. We define J_i^D as follows.*

- $J_0^D = \emptyset$
- $J_{i+1}^D = \{a \in Args_D |$ *a is acceptable w.r.t.* $J_i^D\}$

The set of justified arguments in a defeasible theory D is $JArgs^D = \bigcup_{i=1}^{\infty} J_i^D$.

3 n-Person Argumentation Game

In this section, we utilise the argumentation semantics presented in section 2.2 to model agents' interactions in an n-person argumentation game. Also, we propose a knowledge structure which enables an agent to construct its arguments w.r.t. knowledge from other agents as well as to select a defensive argument.

3.1 Agents' Interactions

In an argumentation game, a group of agents $\mathscr{A}$ shares a set of goals $\mathscr{G}$ and a set of external constraints T_{bg} represented as a defeasible theory, known as a background knowledge. This knowledge provides common expectations and restrictions in $\mathscr{A}$. An agent has its own view on the working environment, therefore, can autonomously pursue its own goal. In this work, we model interactions between agents to settle on goals commonly accepted by the group. Also, at each step of the game, we show how an agent can identify a goal and sub-goals for its counter arguments. This information is critical for those agents whose main claims are refuted either directly by arguments from other agents or indirectly by the combination of these arguments.

Settling on common goals. An agent can pursue a goal in the set of common goals $\mathscr{G}$ by proposing an explanation for its goal. The group justifies proposals from individual agents in order to identify commonly-accepted goals using a dialogue as follows:

1. Each agent broadcasts an argument for its goal. The system can be viewed as an argumentation game with n-players corresponding to the number of agents.
2. An agent checks the status of its argument against those from the other agents. There are three possibilities: (a) *directly refuted* if its argument conflicts with those from others; (b) *collectively refuted* if its argument does not conflict with individual arguments but violates the combination of individual arguments (See section 3.2); (c) *collectively accepted* if its argument is justified by the combination (See section 3.3).
3. According to the status of its main claim, an agent can: (a) defend its claim; (b) attack a claim from other agents; (c) rest. An agent repeats the previous step until the termination conditions of the game are reached.
4. The dialogue among agents is terminated if all agents can pass their claims. For a dispute, agents stop arguing if they do not have any more argument to propose.

Weighting opposite premises. In a dialogue, at each iteration, an agent is required to identify goals and sub-goals which are largely shared by other agents. This information is highly critical for agents, whose main claims are refuted either directly by other agents or collectively by the combination of arguments from others in order to effectively convince other agents.

To achieve that an agent, A_{me}, identifies a sub-group of agents, namely "opp-group", which directly or collectively attacks its main claim. A_{me} creates $Args^{opp}$ as the set of opposing arguments from the opp-group and P^{opp} as the set of premises in $Args^{opp}$.

Essentially, $Args^{opp}$ contains arguments attacking A_{me}'s claim. Each element of P^{opp} is weighted by its frequency in $Args^{opp}$. We define the preference over P^{opp} as given $p_1, p_2 \in P^{opp}$, $p_2 \succeq p_1$ if the frequency of p_2 in $Args^{opp}$ is greater than that of p_1. Basically, the more frequent an element $q \in P^{opp}$ is the more agents use this premise in their arguments. Therefore the refutation of q challenges other agents better than the premises having lower frequency since this refutation causes a larger number of agents to reconsider their claims.

Defending the main claim. At iteration i, $Args_i^{opp}$ represents the set of arguments played by the opp-group:

$$Args_i^{opp} = \bigcup_{j=0}^{|\mathscr{A}|} Args_i^{A_j} \text{ s.t. } Args_i^{A_j} \text{ directly attacks } Args_i^{A_{me}}$$

where $Args^{A_j}$ is the argument played by agent A_j. If A_j rests at iteration i, its last argument (at iteration k) is used $Args_i^{A_j} = Args_k^{A_j}$. The set of opposite premises at iteration i is:

$$P_i^{opp} = \{p | p \in Args_i^{opp} \text{ and } p \notin Args_i^{A_{me}}\}$$

The preference over elements of P^{opp} provides a mechanism for A_{me} to select arguments for defending its main claim.

Example 1. Suppose that agent A_1 and A_2 respectively propose $Args^{A_1} = \{\Rightarrow e \Rightarrow b \Rightarrow a\}$ and $Args^{A_2} = \{\Rightarrow e \Rightarrow c \Rightarrow a\}$ whilst agent A_3 claims $Args^{A_3} = \{\Rightarrow d \Rightarrow \sim a\}$. From A_3's view, its claim directly conflicts with those of A_1 and A_2. The arguments and premises of the opp-group are:

$$Args_i^{opp} = \{\Rightarrow e \Rightarrow b \Rightarrow a; \Rightarrow e \Rightarrow c \Rightarrow a\} \text{ and } P_i^{opp} = \{a^2, b^1, c^1, e^2\}$$

The superscript of elements in P_i^{opp} represents the frequency of a premise in $Args_i^{opp}$. A_3 can defend its claim by providing a counter-argument that refute $\sim a$ – the major claim of the opp-group. Alternatively, A_3 can attack either b or c or e in the next step. An argument against e is the better selection compared with those against b or c since A_3's refutation of e causes both A_1 and A_2 to reconsider their claims.

Attacking an argument. In this situation, individual arguments of other agents do not conflict with that of A_{me} but the integration of these arguments does. Agent A_{me} should argue against one of these arguments in order to convince others about its claim.

At iteration i, let the integration of arguments be $T_i^{INT} = T_{bg} \bigcup_{j=0}^{|\mathscr{A}|} T_j^i$, where T_j^i is the knowledge from agent j supporting agent j's claim, and $JArgs^{INT}$ be the set of justified arguments from integrated knowledge of other agents (See section 3.3). The set of opposite arguments is defined as:

$$Args_i^{opp} = a | a \in JArgs^{INT} \text{ and } a \text{ is attacked by } Args_i^{A_{me}}$$

and the set of opposite premises is:

$$P_i^{opp} = \{p | p \in Args_i^{opp} \text{ and } (p \notin Args_i^{A_{me}} \text{ or } p \text{ is not attacked by } Args_i^{A_{me}})\}$$

The second condition is to guarantee that A_{me} is self-consistent and does not play any argument against itself. In order to convince other agents about its claim, A_{me} is required to provide arguments against any premise in P^{opp}. In fact, the order of elements in P^{opp} offers a guideline for A_{me} on selecting its attacking arguments.

3.2 Agent's Knowledge Structure

In this section, we present a knowledge structure which allows an agent to incorporate background knowledge and knowledge exposed by individual agents during the game. Also, we propose two simple methods to integrate knowledge sources w.r.t. ambiguity information.

Knowledge representation. Agent A_{me} has three types of knowledge including the background knowledge T_{bg}, its own knowledge about working environment T_{me}, and the knowledge about others:

$$\mathscr{T}_{other} = \{T_j : 1 \leq j \leq |\mathscr{A}| \& \text{ excluding } T_{me}\}$$

where T_j is obtained from agent Ag_j during iterations and T_j is represented in DL. At iteration i, the knowledge obtained from Ag_j is accumulated from previous steps:

$$T_j^i = \bigcup_{k=0}^{i-1} T_j^k + Args_i^{Ag_j}$$

In our framework, the knowledge of an agent can be rebutted by other agents. It is reasonable to assume that defeasible theories contain only defeasible rules and defeasible facts (defeasible rules with empty body).

Knowledge integration. To generate arguments, an agent integrates knowledge from different sources. Given ambiguous information between two sources, there are two possible methods to combine them: ambiguity blocking is selected if the preference between these sources is known; otherwise, ambiguity propagation is applied.

Ambiguity blocking integration. This method extends the standard defeasible reasoning by creating a new superiority relation from that of the knowledge sources i.e. given two sources as T_{sp} – the superior theory, and T_{in} – the inferior theory, we generate a new superiority relation $R_d^{sp} > R_d^{in}$ based on rules from two sources. The integration of the two sources is denoted as $T_{INT} = T_{sp} \ni T_{in}$. Now, the standard defeasible reasoning can be applied for T_{INT} to produce a set of arguments $Args_{AB}^{T_{sp} \ni T_{in}}$.

Example 2. Given two defeasible theories

$$T_{bg} = \{R_d = \{r_1 : e \Rightarrow c; r_2 : g, f \Rightarrow \sim c, r_3 :\Rightarrow e\}; >= \{r_2 > r_1\}\}$$
$$T_{me} = \{R_d = \{r_1 :\Rightarrow d; r_2 : d \Rightarrow \sim a; r_3 :\Rightarrow g\}\}$$

The integration produces $T_{bg} \ni T_{me} =$

$$\{R_d = \{r_1^{T_{bg}} : e \Rightarrow c; r_2^{T_{bg}} : g, f \Rightarrow \sim c, r_3^{T_{bg}} :\Rightarrow e; r_1^{T_{me}} :\Rightarrow d; r_2^{T_{me}} : d \Rightarrow a; r_3^{T_{me}} :\Rightarrow g\};$$
$$>= \{r_2^{T_{bg}} > r_1^{T_{bg}}\}\}$$

The integrated theory inherits the superiority relation from T_{bg}. That means the new theory reuses the blocking mechanism from T_{bg}.

Ambiguity propagation integration. Given two knowledge sources T_1 and T_2, the reasoning mechanism with ambiguity propagation can directly apply to the combination of theories denoted as $T_{INT} = T_1 + T_2$. The preference between two sources is unknown, therefore, there is no method to solve conflicts between them. The supportive and opposing arguments for any premise are removed from the final set of arguments. The set of arguments obtained by this integration is denoted by $Args_{AP}^{T_1+T_2}$.

3.3 Argument Justification

The motivation of an agent to participate in the game is to promote its own goal. However, its claim can be refuted by different agents. To gain the acceptance of the group, at the first iteration, an agent should justify its arguments by common constraints and expectations of the group governed by the background knowledge T_{bg}. The set of arguments justified by T_{bg} determines arguments that an agent can play to defend its claim. In subsequent iterations, even if the proposal does not conflict with other agents, an agent should ponder knowledge from others to determine the validity of its claim. That is an agent is required a justification by collecting individual arguments from others.

Justification by background knowledge. Agent A_{me} generates the set of arguments for its goals by combining its private knowledge T_{me} and the background knowledge T_{bg}. The combination is denoted as $T'_{me} = T_{bg} \ni T_{me}$ and the set of arguments is $Args^{T'_{me}}$. Due to the non-monotonic nature of DL, the combination can produce arguments beyond individual knowledges. From A_{me}'s view, this can bring more opportunities to fulfil its goals. However, A_{me}'s arguments must be justified by the background knowledge T_{bg} since T_{bg} governs essential behaviours (expectations) of the group. Any attack to T_{bg} is not supported by members of $\mathscr{A}$. A_{me} maintains the consistency with the background knowledge T_{bg} by following procedure:

1. Create $T'_{me} = T_{bg} \ni T_{me}$. The new defeasible theory is obtained by replicating all rules from common constraints T_{bg} into the internal knowledge T_{me} while maintaining the superiority of rules in T_{bg} over that in T_{me}.
2. Use the ambiguity blocking feature to construct the set of arguments $Args^{T_{bg}}$ from T_{bg} and the set of arguments $Args_{AB}^{T'_{me}}$ from T'_{me}.

3. Remove any argument in $Args_{AB}^{T'_{me}}$ attacked by those in $Args^{T_{bg}}$, obtaining the justified arguments by the background knowledge $JArgs^{T'_{me}} = \{a \in Args_{AB}^{T'_{me}}$ and a is accepted by $Args^{T_{bg}}\}$

Example 3. Consider two defeasible theories:

$$T_{bg} = \{R_d = \{r_1 : e \Rightarrow c; r_2 : g, f \Rightarrow \sim c, r_3 :\Rightarrow e\}; >= \{r_2 > r_1\}\}$$
$$T_{me} = \{R_d = \{r_1 :\Rightarrow d; r_2 : d \Rightarrow \sim a; r_3 :\Rightarrow g\}\}$$

We have sets of arguments from the background theory and the integrated theory:

$$Args^{T_{bg}} = \{\Rightarrow e; \Rightarrow e \Rightarrow c\}$$
$$Args^{T_{bg} \ni T_{me}} = \{\Rightarrow e; \Rightarrow e \Rightarrow c; \Rightarrow d; \Rightarrow g; \Rightarrow d \Rightarrow \sim a\}$$

In this example, there is not any attack between arguments in $Args^{T_{bg}}$ and $Args_{AB}^{T_{bg} \ni T_{me}}$. In other words, arguments from $Args^{T_{bg} \ni T_{me}}$ are acceptable by those from $Args^{T_{bg}}$. The set of justified arguments w.r.t. $Args^{T_{bg}}$ $JArgs^{T'_{me}} = Args_{AB}^{T_{bg} \ni T_{me}}$.

Collective justification. During the game, A_{me} can exploit the knowledge exposed by other agents in order to defend its main claims. Due to possible conflicts in individual proposals, an agent uses the sceptical semantics of the ambiguity propagation reasoning to retrieve the consistent knowledge. Essentially, given competing arguments an agent does not have any preference over them, therefore, these arguments will be rejected. The consistent knowledge from the others allows an agent to discover "collective wisdom" distributed among agents in order to justify its claim.

The justification of collective arguments, which are generated by integrating all knowledge sources, is done by the arguments from the background knowledge $Args^{T_{bg}}$. The procedure runs as follows:

1. Create a new defeasible theory $T''_{me} = T_{bg} + T_{me} + \mathscr{T}_{other}$.
2. Generate the set of arguments $Args_{AP}^{T''_{me}}$ from T''_{me} using the feature of ambiguity propagation.
3. Justify the new set of arguments $JArgs^{T''_{me}} = \{a | a \in Args_{AP}^{T''_{me}}$ and a is accepted by $Args^{T_{bg}}$.

$JArgs^{T''_{me}}$ allows A_{me} to verify the status of its arguments for its claim $JArgs^{T'_{me}}$. If arguments in $JArgs^{T''_{me}}$ and $JArgs^{T'_{me}}$ do not attack one another, A_{me}'s claims are accepted by other agents. Any conflict between two sets shows that accepting arguments in $JArgs^{T''_{me}}$ stops A_{me} to achieve its claims in next steps. The set of arguments $Args^{opp}$ against A_{me} is identified as any argument in $JArgs^{T''_{me}}$ attacking A_{me}'s arguments. A_{me} also establishes P^{opp} to select its counter-argument. It is noticed that A_{me} is self-consistent.

Example 4. Suppose the background knowledge T_{bg} and the private knowledge T_{me} of A_{me} are:

$$T_{bg} = \{R_d = \{r_1 : e \Rightarrow c; r_2 : g, f \Rightarrow \sim c\}; >= \{r_2 > r_1\}\}$$
$$T_{me} = \{R_d = \{r_1 :\Rightarrow e; r_2 : c \Rightarrow d; r_3 :\Rightarrow g\}\}$$

Agent A_{me} currently plays $\{\Rightarrow e \Rightarrow c \Rightarrow d\}$ and knows about other agents:

$$\mathscr{T}_{other} = \{T_1, T_2\} \text{ where } T_1 = \{\Rightarrow h \Rightarrow f \Rightarrow b \Rightarrow a\} \text{ and } T_2 = \{\Rightarrow e \Rightarrow c \Rightarrow a\}$$

The claim of A_3 is acceptable w.r.t. arguments played by the other agents. However, the combination $T_{bg} + T_{me} + \mathscr{T}_{other}$ shows the difference. This combination generates $\{\Rightarrow g, \Rightarrow e, \Rightarrow e \Rightarrow f \Rightarrow b, \Rightarrow g, f \Rightarrow \sim c\}$. $\{\Rightarrow g, f \Rightarrow \sim c\}$ is due to the superiority relation in T_{bg} which rebuts the claim of A_3. Therefore, the set of opposing arguments $Args^{opp} = \{\Rightarrow g, f \Rightarrow \sim c\}$ and $P^{opp} = \{f^1\}$. Given this information, A_3 should provide a counter-evidence to f in order to pursue c. Moreover, A_3 should not expose g to the other agents. Otherwise, A_3 has to drop its initial claim d.

4 Related Work

Substantial works have been done on argumentation games in the artificial intelligence and law-field. [1] introduces a dialectical model of legal argument, in the sense that arguments can be attacked with appropriate counterarguments. In the model, the factual premises are not arguable; they are treated as strict rules. [12] presents an early specification and implementation of an argumentation game based on the Toulmin argument-schema without a specified underlying logic. [13] presented the pleadings game as a normative formalization and fully implemented computational model, using conditional entailment.

Settling on a common goal among agents can be seen as a negotiation process where agents exchange information to resolve conflicts or to obtain missing information. The work in [14] provides a unified and general formal framework for the argumentation-based negotiation dialogue between two agents. The work establishes a formal connection between the status of a argument (accepted, rejected, and undecided) with an agent's actions (accept, reject, and negotiate respectively). Moreover, an agent's knowledge is evolved by accumulating arguments during interactions.

[3] presents an argumentation-based coordination, where agents can exchange arguments for their goals and plans to achieve the goals. The acceptance of an argument of an agent depends on the attitudes of this agent namely credulous, cautious, and sceptical. In [15], agents collaborate with one another by exchanging their proposals and counter-proposals in order to reach a mutual agreement. During conversations, an agent can retrieve missing literals (regarded as sub-goals) or fulfil its goals by requesting collaboration from other agents.

We have advantages of using DL since it flawlessly captures the statuses of arguments, such as accepted, rejected, and undecided by the proof conditions of DL. The statuses are derived from the notions of $+\partial, -\partial$ and $+\Sigma$ corresponding to a positive proof, a negative proof, and a positive support of a premise. Consequently, an agent can take a suitable action either to provide more evidence or to accept an argument from others. In addition, DL provides a compact representation to accommodate new information.

Using DL to capture concepts of the argumentation game is supported by [16,17] and recently [18,19]. [16] focuses on persuasive dialogues for cooperative interactions among agents. It includes in the process cognitive states of agents such as knowledge and beliefs, and presents some protocols for some types of dialogues (e.g. information seeking, explanation, persuasion). [17] provides an extension of DL to include the step of the adversarial dialogue by defining a meta-program for an alternative computational algorithm for ambiguity propagating DL while the logic presented here is ambiguity blocking.

We tackle the problem of evolving knowledge of an agent during iterations, where the argument construction is an extension of [18,19], which define the strengthening of an argument after each step in an argumentation game. [19] differs from [18] in distinguishing participants that one participant must provide a strong argument (i.e. a definite proof) in order to defeat the other. In our work, we define the notion of collective acceptance for an argument and a method to weight arguments defending against opposing arguments by using both features of ambiguity blocking and propagating.

The works in literature did not clearly show how an agent can tackle with conflicts from multiple agents, especially when the preference over arguments is unknown. The main difference in our framework is the external model where more than two agents can argue to settle on goals commonly accepted by the group. Our weighting mechanism enables an agent to build up a preference over premises constituting opposing arguments from other agents. As a result, an agent can effectively select an argument among those justified by the group's background knowledge to challenge other agents.

We also propose the notion of collective justification to tackle the side-effect of accepting claims from individual agents. Individual arguments for these claims may not conflict with one another, but the integration of these arguments can result in conflicting with an agent's claim. This notion is efficiently deployed in our work due to the efficiency of defeasible logic in handling ambiguous information.

5 Conclusions

We presented an n-person argumentation game based on defeasible logic, which enables a group of more than two agents to settle on goals commonly accepted by the group. During an argumentation game, each agent can use knowledge from multiple sources including the group's constraints and expectations, other agents' knowledge, and its own knowledge in order to argue to convince other agents about its goals. The knowledge about the group's constraints and expectations plays a critical role in our framework since this knowledge provides a basis to justify new arguments non-monotonically inferred from the integration of different sources.

In this work, we propose a simple weighting mechanism, which is based on the frequency of premises in arguments attacking an agent's claim, in order to tackle the problem of conflicts from multiple agents. In the future work, we will extend this mechanism to incorporate the notion of trustful arguments from trusted agents to better select a rebuttal argument and resolve conflicts among agents.

References

1. Prakken, H., Sartor, G.: A dialectical model of assessing conflicting arguments in legal reasoning. Artificial Intellifgence and Law **4** (1996) 331–368

2. Jennings, N.R., Parsons, S., Noriega, P., Sierra, C.: On argumentation-based negotiation. In: Proceedings of the International Workshop on Multi-Agent Systems. (1998)
3. Parsons, S., McBurney, P.: Argumentation-based dialogues for agent coordination. group decision and negotiation. Group Decision and Negotiation (12) (2003) 415–439
4. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2) (1995) 321–358
5. Antoniou, G., Billington, D., Governatori, G., Maher, M.J.: A flexible framework for defeasible logics. In: Proc. American National Conference on Artificial Intelligence (AAAI-2000). (2000) 401–405
6. Maher, M.J., Rock, A., Antoniou, G., Billignton, D., Miller, T.: Efficient defeasible reasoning systems. International Journal of Artificial Intelligence Tools **10**(4) (2001) 483–501
7. Billington, D.: Defeasible logic is stable. Journal of Logic and Computation **3** (1993) 370–400
8. Antoniou, G., Billington, D., Governatori, G., Maher, M.J.: Representation results for defeasible logic. ACM Transactions on Computational Logic **2**(2) (2001) 255–287
9. Antoniou, G., Billington, D., Governatori, G., Maher, M.J.: Embedding defeasible logic into logic programming. Theory and Practice of Logic Programming **6**(6) (2006) 703–735
10. Maher, M.J.: Propositional defeasible logic has linear complexity. Theory and Practice of Logic Programming **1**(6) (2001) 691–711
11. Governatori, G., Maher, M.J., Antoniou, G., Billington, D.: Argumentation Semantics for Defeasible Logic. J Logic Computation **14**(5) (2004) 675–702
12. Bench-Capon, T.J.: Specification and implementation of Toulmin dialogue game. In J.C. Hage, T.J.M. Bench-Capon, A.K.C.d.V.M.C.G., ed.: Jurix 1998. (1998) 5–20
13. Lodder, A.R.: Thomas F. Gordon, The Pleadings Game – an artificial intelligence model of procedural justice. Artif. Intell. Law **8**(2/3) (2000) 255–264
14. Amgoud, L., Dimopoulos, Y., Moraitis, P.: A unified and general framework for argumentation-based negotiation. In: Proceedings of the 6th international joint conference on AAMAS. (2007) 1–8
15. Rueda, S.V., Garcia, A.J., Simari, G.R.: Argument-based negotiation among bdi agents. Journal of Computer Science and Technology **2**(7) (2002)
16. Letia, I.A., Vartic, R.: Defeasible protocols in persuasion dialogues. In: WI-IATW '06: Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology. (2006) 359–362
17. Nilsson, J.F., Eriksson Lundstrøm, J., Hamfelt, A.: A metalogic formalization of legal argumentation as game trees with defeasible reasoning. In: Proceedings of ICAIL'05, International Conference on AI and Law. (jun 2005)
18. Thakur, S., Governatori, G., Padmanabhan, V., Eriksson Lundström, J.: Dialogue games in defeasible logic. In: 20th Australian Joint Conference on Artificial Intelligence, AI 2007. Volume 4830 of LNAI., Springer (2007) 497–506
19. Eriksson Lundström, J., Governatori, G., Thakur, S., Padmanabhan, V.: An asymmetric protocol for argumentation games in defeasible logic. In: 10 Pacific Rim International Workshop on Multi-Agents. Volume 5044 of LNCS., Springer (2008)