

Defeasible Logic to Model n-person Argumentation Game

Duy Hoang Pham, Subhasis Thakur, Guido Governatori

School of Information Technology and Electrical Engineering
The University of Queensland, Brisbane, Australia
{pham,subhasis,guido}@itee.uq.edu.au}

Abstract

In multi-agent systems, an individual agent can pursue its own goals, which may conflict with those held by other agents. To settle on a common goal for the group of agents, the argumentation/dialogue game provides a robust and flexible tool where an agent can send its explanation for its goal in order to convince other agents. In the setting that the number of agents is greater than two and they are equally trustful, it is not clear how to extend existing argumentation/dialogue frameworks to tackle conflicts from many agents. We propose to use the defeasible logic to model the n-person argumentation game and to use the majority rule as an additional preference mechanism to tackle conflicts between arguments from individual agents.

Introduction

In a group of agents, there are several situations requiring agents to settle on a common goal despite that each agent can pursue its goals which may conflict with other agents. A simple but efficient method to tackle the problem is to give weights over the goals. However, this method is not robust and limits the autonomy of an individual agent. Also, the conflicts among agents are likely to arise from a partial view and incomplete information on working environment of individual agents. To settle conflicts among agents an agent can argue to convince others about its pursued goal and provides evidences to defend its claim. This interaction between agents can be modelled as an argumentation game (Prakken & Sartor 1996; Jennings *et al.* 1998; Parsons & McBurney 2003; Amgoud, Dimopoulos, & Moraitis 2007). In an argumentation game, an agent can propose an explanation for its pursued goal (i.e., an argument), which can be rejected by counter-evidences from other agents. This interaction can be iterated until an agent (the winner) successfully argues its proposal against other agents. The argumentation game approach offers a robust and flexible tool for agents to resolve conflicts by evaluating the status of arguments. Dung's argumentation semantics (Dung 1995) is widely recognised to establish relationships among arguments. The key notion for a set of arguments is whether a set of arguments is self-consistent and provides the base to derive a conclusion. A conclusion is justified, and thus provable, if there is a set of supporting arguments and all counter-arguments are deficient when we

consider the arguments in the set of supporting arguments.

An argumentation game is more complicated when the number of participants is greater than two. It is not clear how to extend existing approaches to cover the argumentation in groups of more than two agents, especially when agents are equally trustful. That is arguments from individual agents have the same weight. In this case, the problem amounts to how to decide which argument has precedence over competitive arguments. In other words, the problem is to determine the global collective preference of a group of agents.

The main idea behind our approach is that if individual preferences of agents are not sufficient to solve a conflict (for example, we have several arguments without any relative preference over them), the group of agents uses the majority rule (Lin 1996) over initial proposals to determine the "most common" claim known as the "topic" of a dialogue. That is the topic preferred by the majority of the group. An agent either supports the topic or defends its own claim against the topic. Our majority mechanism simplifies the complexity of the n-person argumentation and provides a strategy for an agent to select an argument for defending its proposal. That is an argument causing more "supporters" to reconsider "their attitude" will be preferred by defending agents.

Each of our agents has three types of knowledge: its private knowledge, background knowledge, and knowledge obtained from other agents. The background knowledge presents the expected behaviour of a member of the group that is commonly shared by the group. The knowledge about other agents growing during the interactions enables an agent to efficiently convince others about its own goal. Essentially, the background knowledge is preferred over other sources because it represents common expectations and constraints of the group. Any argument violating the background knowledge is not supported by the group.

Defeasible logic is chosen as our underlying logic for the argumentation game due to its efficiency, simplicity in representing incomplete and conflicting information and relationships with logic programming (Antoniou *et al.* 2006). Furthermore, the logic has a powerful and flexible reasoning mechanism (Antoniou *et al.* 2000; Maher *et al.* 2001) which enables our agents to capture Dung's argumentation semantics by using two features of defeasible reasoning, namely the ambiguity propagating (the preference over conflicts is

unknown) and ambiguity blocking (the preference is given).

Our paper is structured as follows. In the second section, we briefly introduce essential notions of defeasible logics, the construction of arguments using defeasible reasoning with respect to (w.r.t) ambiguous information, and the majority rule. In the third section, we introduce n-person argumentation 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 different individual knowledge sources in order to propose and to defend its goal against the other agents. The fourth section provides an overview of research works related to our approach. The final section concludes the paper.

Background

Defeasible Logic

Following the presentation in (Billington 1993), the basic components of defeasible logic (DL) are: *facts*, *strict rules*, *defeasible rules*, *defeaters*, and a *superiority relation*.

Facts are undeniable statements, which are always true. *Strict rules*, similar to rules in classical logics, are rules whose conclusions are unquestionable. *Defeasible rules* are different from strict rules in the way that their conclusions can be overridden by contrary evidences. *Defeaters* are rules that cannot be used to draw any conclusion but to prevent some conclusions from some defeasible rules by producing evidence to the contrary. The *superiority relation* defines priorities among rules. That is, one rule may override the conclusion of another rule when we have to solve a conflict between rules with opposite conclusions.

A defeasible theory D is a triple $(F, R, >)$ where F is a finite set of facts, R a finite set of rules, and $>$ a superiority relation on R .

The language of DL consists of a finite set of literals. Given a literal l , we use $\sim l$ to denote the propositional literal complementary to l , that is if $l = p$, then $\sim l = \neg p$, and if $l = \neg p$, then $\sim l = p$.

A rule r in R is composed of an antecedent or body $A(r)$ and a consequent or head $C(r)$. $A(r)$ consists of a finite set of literals while $C(r)$ contains a single literal. $A(r)$ can be omitted from the rule if it is empty. The set of rules R can include all three types of rules, namely R_s (strict rules), R_d (defeasible rules), and R_{df} (defeaters). We will use R_{sd} for the set of strict and defeasible rules, and $R[q]$ for the set of rules whose head is q .

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 .

Provability is based on the concept of a derivation (or proof) in $D = (F, R, >)$. Informally, definite conclusions can

derive from strict rules by forward chaining, while defeasible conclusions can obtain 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 for definitely and defeasibly provable conclusions¹:

- $+\Delta$: 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 finite², and it can be computed in linear time (Maher 2001). In addition, several efficient implementations have been proposed (see (Maher *et al.* 2001)).

Example 1 Consider the defeasible theory D has a set of defeasible rules:

$$R_d = \{r_1 := a; r_2 := \sim a; r_3 := b; r_4 : a \Rightarrow \sim b\}$$

and a superiority relation

$$> = \{r_2 > r_1\}$$

r_1 and r_2 have empty therefore they are applicable to derive $+\partial a$ and $+\partial \sim a$ respectively. These conclusions are clearly ambiguous. Thanks to the superiority relation, the conclusion of a is overridden. That means $-\partial a$ is in the conclusions from theory D . As a result, $+\partial b$ is added to the conclusion set without any ambiguity. r_4 is no longer applicable due to $-\partial a$.

Defeasible logic can be extended by an ambiguity propagating variant (See (Governatori *et al.* 2004; Antoniou *et al.* 2000)). The superiority relation is not considered in the inference process. The inference with the ambiguity propagation introduces a new tag Σ . A literal $p (+\Sigma p)$ 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.

¹ Refer to (Antoniou *et al.* 2001) for proof conditions for all tagged conclusions.

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

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

We can achieve ambiguity propagation behaviour by making a minor change to the inference condition for $+\partial_{AP}$ ³

$$\begin{array}{l}
+\partial_{AP}: \text{ If } P(i+1) = +\partial q \text{ then either} \\
(1) +\Delta q \in P(1..i) \text{ or} \\
(2.1) \exists r \in R_{sd}[q] \forall a \in A(r) : +\partial_{AP} a \in P(1..i) \text{ and} \\
(2.2) -\Delta \sim q \in P(1..i) \text{ and} \\
(2.3) \forall s \in R_{sd}[\sim q] \\
\quad \exists a \in A(s) : -\partial_{AP} a \in P(1..i) \text{ or}
\end{array}$$

Example 2 We modify the defeasible theory D in example 1 by removing the superiority relation:

$$R_d = \{r_1 : \Rightarrow a; r_2 : \Rightarrow \sim a; r_3 : \Rightarrow b; r_4 : a \Rightarrow \sim b\}$$

Without the superiority relationship, there is no means to decide between a and $\sim a$ due to both of r_1 and r_2 are applicable. In a setting where the ambiguity is blocked, b is not ambiguous because r_3 for b is applicable whilst r_4 is not since its antecedent is not provable. If the ambiguity is propagated, we have evidence supporting all of four literals since all of the rules is applicable. $+\Sigma a, +\Sigma \sim a, +\Sigma b$ and $+\Sigma \sim b$ are included in the conclusion set. Moreover we can derive $-\partial a, -\partial \sim a, -\partial b$ and $-\partial \sim b$ showing that the resulting logic exhibits an ambiguity propagating behaviour. In the second setting b is ambiguous, and its ambiguity depends on that of a .

Argumentation by Defeasible Logic

In what follows, we briefly introduce the basic notions of an argumentation system using defeasible logic as underlying logical language. Moreover, we present the acceptance of an argument w.r.t Dung's semantics.

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, \dots, b_n label the children of h then there is a rule in R with body b_1, \dots, 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.

In general, arguments are defined to be proof trees (or monotonic derivations). Defeasible logic requires a more general notion of proof tree that admits infinite trees, therefore the distinction is kept between an unrefuted, but infinite, 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.

³The proof for $-\partial_{AP}$ is derived from that of $+\partial_{AP}$ using the strong negation principle (Maher et al. 2001).

- 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 those of literals which are constituted in these arguments. An argument A attacks a defeasible argument B if a conclusion of A is the complement of a conclusion of B , and that conclusion of B is not part of a strict sub-argument of B . A set of arguments \mathcal{S} attacks a defeasible argument B if there is an argument A in \mathcal{S} that attacks B .

A defeasible argument A is undercut by a set of arguments \mathcal{S} if \mathcal{S} supports an argument B attacking a proper non-strict sub-argument of A . An argument A is undercut by \mathcal{S} means we can show that some premises of A cannot be proved if we accept the arguments in \mathcal{S} .

It is noticed that the concepts of the attack and undercut concern only defeasible arguments and sub-arguments. For strict arguments we stipulate that they cannot be undercut or attacked.

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 \mathcal{S} . The concepts of provability depend on the methods used by the reasoning mechanism to tackle ambiguous information. According to the features of the defeasible reasoning, we have two definitions of acceptable arguments (definition 2 and 3).

Definition 2 In case of the reasoning with the ambiguity propagation, an argument A for p is acceptable w.r.t a set of arguments \mathcal{S} if A is finite, and

1. A is strict, or
2. every argument attacking A is attacked by \mathcal{S} .

Definition 3 If the reasoning with the ambiguity blocking is used, an argument A for p is acceptable w.r.t a set of arguments \mathcal{S} if A is finite, and

1. A is strict, or
2. every argument attacking A is undercut by \mathcal{S} .

Due to the concept of acceptance, we can determine the status of an argument. If an argument can resist a reasonable refutation, this argument is justified (definition 4). If an argument can not overcome attacks from other arguments, this argument is rejected (definition 5).

Definition 4 Let D be a defeasible theory. We define J_i^D as follows.

- $J_0^D = \emptyset$
- $J_{i+1}^D = \{a \in \text{Args}_D \mid a \text{ is acceptable w.r.t } J_i^D\}$

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

Definition 5 Let D be a defeasible theory and \mathcal{T} be a set of arguments. We define $R_i^D(\mathcal{T})$ as follows.

- $R_0^D(\mathcal{T}) = \emptyset$
- $R_{i+1}^D(\mathcal{T}) = \{a \in \text{Args}_D \mid a \text{ is rejected by } R_i^D(\mathcal{T}) \text{ and } \mathcal{T}\}$.

The set of rejected arguments in a defeasible theory D w.r.t \mathcal{T} is $R\text{Args}^D(\mathcal{T}) = \bigcup_{i=1}^{\infty} R_i^D(\mathcal{T})$.

Majority Rule

The *majority rule* from (Lin 1996) retrieves a maximal amount of consistent knowledge from a set of agents' knowledge. Conflicts between agents can be tackled by considering not only the number of agents supporting that information but also the importance (reliability) of the agents. The approach provides a useful and efficient method to discover information largely held by agents. The majority knowledge can be used either to reinforce the current knowledge of an agent or to introduce new information into the agent's knowledge.

Due to possible conflicting information within a source, the merging operator by majority cannot directly apply to our framework. Instead, the majority rule pools potential joint conclusions derived by the defeasible reasoning, which resolves possible conflicts.

Considering the knowledge sources $\{T_1, \dots, T_n\}$, C_i denotes the set of tagged conclusions that can be derived by the defeasible reasoning from the corresponding theory T_i . The level that the theory T_i supports a literal l corresponds to its weight represented w_{T_i} as follows:

$$\text{support}(l, T_i) = \begin{cases} w_{T_i} & l \in C_i \\ 0 & \text{otherwise} \end{cases}$$

The majority knowledge from the others, T_{maj} , whose elements are inferred from $\{C_1, \dots, C_n\}$ by the majority rule, is determined by the formula:

$$T_{maj} = \left\{ c : \sum_{T_i} \text{support}(c, T_i) > \frac{\sum w_{T_i}}{2} \right\}$$

n-Person Argumentation Framework

In this section, we develop our framework by using the argument construction from the defeasible reasoning. In particular, we define an external model which describes interactions between agents in order to achieve a goal supported by the majority. Also, we present an internal model which illustrates the reasoning method on knowledge from other agents exposed during interactions.

Model Agents' Interaction

This section describes the basic scenario where an individual agent exchanges arguments to promote its own goal and to reach an agreement by the majority. Consider a set of agents \mathcal{A} sharing a set of goals \mathcal{G} and external constraints represented as a defeasible theory T_{bg} . These external constraints are also known as background knowledge which provides common expectations and restrictions among agents in \mathcal{A} . An individual agent in \mathcal{A} can have its own view on the working environment, therefore can pursue its own goals. In this work, we model the interactions among these agents in order to establish a goal accepted by the majority of the group. Due to the partial view and incomplete information of an agent, we believe the argumentation game is a useful method to tackle this problem.

Determine common goal. In order to pursue a goal, an agent generates an argument for its goal. This goal is considered as its main claim. The process to determine the goal supported by the group involves multiple steps in 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. The group of agents determines the dialogue topic by using the majority rule over the set of claims (i.e. the goals from players)⁴. The claim supported by more than a half of the group is selected. If the group can not settle a topic, the previous step is repeated. The dialogue terminates early if agents fail to achieve a majority goal and they do not have any new goal to propose.
3. An agent can rest if its claim is supported by the majority. Otherwise, the agent can provide a new argument to defend its claim against the common one. At this step, the group creates a set of majority arguments $Args_i^{maj}$ and a set of majority premises P_i^{maj} where i indicates the iteration. An agent utilises these sets to select its new arguments for subsequence steps. Also, new arguments are required to be justified by the background knowledge.
4. A dialogue terminates when all agents pass for an iteration (i.e., do not propose a new argument). Now, the group can settle on the common goal and the explanation accepted by the majority of the group.

Example 3 Suppose that there is three agents A_1, A_2 , and A_3 . A_1 and A_2 respectively propose $Args^{A_1}$ and $Args^{A_2}$

$$Args^{A_1} = \{ \Rightarrow e \Rightarrow b \Rightarrow a \}$$

$$Args^{A_2} = \{ \Rightarrow e \Rightarrow c \Rightarrow a \}$$

whilst A_3 claims

$$Args^{A_3} = \{ \Rightarrow d \Rightarrow \sim a \}.$$

The topic of the dialogue accepted by the majority is a .

Identify majority arguments. Once the group successfully identifies the common claim, the group is divided into two sub-groups namely "pros-group" and "cons-group". Agents in the pros-group support the common claim whilst the cons-group does not. By using the majority rule the agents in the cons-group determine their defensive arguments by attacking the "most common" premise among the arguments from the pros-group. That will force the pros-group to reconsider their claim.

At iteration i , G_i^{maj} is the claim by the majority of active agents (the agents broadcast their arguments). $Args_i^{maj}$ represents the set of majority arguments which are played by the agents to support G_i^{maj} .

$$Args_i^{maj} = \bigcup_{j=0}^{|\mathcal{A}|} Args^{A_j} | Args^{A_j} \vdash G^{maj}$$

⁴Note that agents in \mathcal{A} have the same weight, therefore the majority rule is applied knowledge sources such that each source has the weight of 1.

where $Args^{A_j}$ is the argument played by agent A_j . The set of majority premises at iteration i is

$$P_i^{maj} = \{p \mid p \in Args_i^{maj}\}$$

We define the preference over P_i^{maj} as given $p_1, p_2 \in P_i^{maj}$, $p_2 \succeq p_1$ if the frequency of p_1 in $Args_i^{maj}$ is less than that of p_2 .

Let $i = 0$ be the first iteration that agents in the group reach a common claim. The topic of the dialogue is set to G_0^{maj} . Given two consecutive iterations: i and $i + 1$, G_i^{maj} and G_{i+1}^{maj} are incompatible claims. That is the pros-group for G_i^{maj} at iteration i is attacked by the cons-group which gives G_{i+1}^{maj} in the next iteration as the counter-evidence. In the case that the cons-group does not have an argument which directly attacks G_i^{maj} , the cons-group uses the order of premises in P_i^{maj} as a preference mechanism to select a counter-argument. The idea is that P_i^{maj} eventually contains premises which are sub-claims of G_i^{maj} . The higher order a premises p in P_i^{maj} has, the more agents in the pros-group support p . Consequently, if p is rebutted, the pros-group should revise its attitude towards the claim.

Example 4 Reconsidering example 3, we have

$$\begin{aligned} G_0^{maj} &= a \\ Args_0^{maj} &= \{e \Rightarrow b \Rightarrow a; e \Rightarrow c \Rightarrow a\} \\ P_0^{maj} &= \{a^2, b^1, c^1, e^2\} \end{aligned}$$

The superscript of a premise in P_0^{maj} represents its frequency in $Args_0^{maj}$. Since the main claim of A_3 does not pass the majority selection, A_3 can defend its proposal by attacking either b or c or e in the next step. An argument against e is likely to be a better selection compared with those against b or c . Another alternative is that A_3 proposes a new argument for $\sim a$ stronger than any of the arguments played by A_1 and A_2

Model Agent's Internal Knowledge

The motivation, which drives an agent to participate in the dialogue, is to promote its own goal. However, its argument for the goal will be accepted if the argument is shared by the majority of the group. To gain the acceptance of the majority, the agent should consider common constraints and expectations of the group, governed by the background knowledge, as well as the attitude of other agents when proposing a claim. The majority rule over the knowledge obtained from other agents enables an agent to probe a common attitude among agents.

At the beginning of the dialogue, the majority rule determines the main claim. In the follow iteration, this rule identifies sub-claims to help an agent to effectively defend its original claim. The idea is that an agent should launch an argument which is likely to alter the majority opinion. The majority rule provides a preference that is among the supportive arguments for the main claim, identifying the most common premise if an agent refutes this premise.

Knowledge representation. An 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:

$$\mathcal{T}_{other} = \{T_j : 1 \leq j \leq |\mathcal{A}| \& j \neq me\}$$

T_j is obtained from agent $Ag_j \in \mathcal{A}$ during iterations (proposing arguments for individual goals). All of these knowledge is represented in defeasible logic. $T_j \in \mathcal{T}_{other}$ is constructed from an argument Ar proposed by the agent Ag_j . At iteration i , the theory obtained from Ag_j is accumulated from previous steps $T_i^j = \bigcup_{k=0}^i T_k^j$.

In our framework, agents can have conflicting knowledge due to partial view and incomplete information sources. We assume that the defeasible theories contain only defeasible rules and defeasible facts (rules with empty body). Therefore, the knowledge of an agent can be rebutted by that from other agents.

Knowledge Integration. To generate an argument, an agent should ponder knowledge from multiple sources. In this section, we present two simple methods to integrate knowledge sources based on ambiguity blocking and ambiguity propagation: given two sources of knowledge, if the preference between these two sources is known we can perform the ambiguity blocking integration; Otherwise, we can select ambiguity propagation integration.

Ambiguity blocking integration. This integration extends the standard defeasible reasoning by creating a new superiority relation from that of the knowledge sources i.e. given two knowledge 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 denotes as $T_{INT} = T_{sp} \oplus T_{in}$. Now the standard defeasible reasoning can be applied for T_{INT} to produce a set of arguments $Args_{AB}^{T_{sp} \oplus T_{in}}$.

Example 5 Given two defeasible theories

$$\begin{aligned} T_{bg} &= \{R_d = \{r_1 : e \Rightarrow c; r_2 : g, f \Rightarrow \sim c, r_3 : \Rightarrow e\}; \\ &> = \{r_2 > r_1\}\} \end{aligned}$$

and

$$T_{me} = \{R_d = \{r_1 : \Rightarrow d; r_2 : d \Rightarrow \sim a; r_3 : \Rightarrow g\}\}$$

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

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

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

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 the underlying logics, the combination can produce arguments being beyond those from individual knowledge. That is the combination can produce arguments which are totally new to the two sources. 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} . In other words, T_{bg} governs essential behaviours (expectations) of the group. Any attack to T_{bg} is not supported by members of \mathcal{A} .

Agent 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 of T_{me} .
2. Use the ambiguity blocking feature to construct the set of arguments $Args_{AB}^{T_{bg}}$ from T_{bg} and the set of arguments $Args_{AB}^{T_{me}'}$ from T_{me}' .
3. Remove any argument in $Args^{T_{me}'}$ attacked by those in $Args_{AB}^{T_{bg}}$, obtaining the justified arguments by the background knowledge

$$JArgs^{T_{me}'} = \{a \in Args^{T_{me}'} \text{ and } a \text{ is not attacked by } Args_{AB}^{T_{bg}}\}$$

Example 6 Given two defeasible theories, T_{bg} and T_{me} , in example 5. We have

$$\begin{aligned} Args^{T_{bg}} &= \{ \Rightarrow e; \\ &\quad \Rightarrow e \Rightarrow c \} \\ Args^{T_{bg} \ni T_{me}} &= \{ \Rightarrow e; \\ &\quad \Rightarrow e \Rightarrow c; \\ &\quad \Rightarrow d; \\ &\quad \Rightarrow g; \\ &\quad \Rightarrow d \Rightarrow \sim a \} \end{aligned}$$

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

$$JArgs^{T_{bg} \ni T_{me}} = Args^{T_{bg} \ni T_{me}} e;$$

Pondering knowledge from the others. During the dialogue, an agent can exploit the knowledge that other agents exposed in order to defend its main claims. Due to possible conflicts in proposals from other agents, an agent can use the sceptical semantics of the ambiguity propagation reasoning in order to retrieve the consistent knowledge. That is given competing arguments, the agent does not have any preference over them and they will be rejected. The consistent knowledge from the others allows an agent to discover ‘‘collective wisdom’’ distributed among agents. From those arguments, agent A_{me} should justify arguments against the set

of majority premises P_i^{maj} at iteration i of the dialogue. The judgement is done by using the arguments from the background knowledge $Args^{T_{bg}}$. The procedure runs as follows:

1. Create a new defeasible theory $T_{me}'' = T_{bg} \ni T_{me} + \mathcal{T}_{other}$.
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 $JArgs^{T_{me}''} = \{a | a \in Args_{AP}^{T_{me}''} \text{ and } a \text{ is accepted by } Args^{T_{bg}}\}$.

At iteration i of the dialogue, the group determines the set P_i^{maj} containing premises support by the majority. In order to refute the majority claim, A_{me} can select an argument from $JArgs_{AB}^{T_{me}'} \cup JArgs_{AP}^{T_{me}''}$ that attack a premise $p \in P_i^{maj}$. The preference of an argument against p is determined by the weight of p . Since the weight of p is proportional to the number of agents supporting p . If p is attacked, the majority can change in favour to A_{me} .

Example 7 Suppose that

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

and the private knowledge of A_{me} has

$$T_{me} = \{R_d = \{r_1 : \Rightarrow d; r_2 : d \Rightarrow \sim a; r_3 : \Rightarrow g\}\}$$

Agent A_{me} currently plays $\Rightarrow d \Rightarrow \sim a$ and knows about other agents

$$\mathcal{T}_{other} = \{T_1, T_2\}$$

where

$$T_1 = \{\Rightarrow e \Rightarrow f \Rightarrow b \Rightarrow a\}$$

and

$$T_2 = \{\Rightarrow e \Rightarrow c \Rightarrow a\}$$

and at this step the majority premises

$$P_i^{maj} = \{a^2, e^2, f^1, b^1, c^1\}.$$

The superscript of an element of $P_{m_j}^i$ represents the frequency (weight) of this element.

The defeasible reasoning with ambiguity propagation for the combination $T_{bg} + T_{me} + \mathcal{T}_{other}$ generates a set of arguments $\Rightarrow g, \Rightarrow e, \Rightarrow e \Rightarrow f \Rightarrow b, \Rightarrow g, f \Rightarrow \sim c, \Rightarrow g, f \Rightarrow \sim c$ is due to the superiority relation in T_{bg} . Given the current knowledge of A_{me} this is only argument that A_{me} can play.

Related Work

Substantial work have been done on argumentation games in the artificial intelligence and Law-field. (Prakken & Sartor 1996) 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. (Lial 1998) presents an early specification and implementation of an argumentation game based on the Toulmin argument-schema without a specified underlying logic. (Lodder 2000) presented The

Pleadings Game as a normative formalization and fully implemented computational model, using conditional entailment. The goal of the model was to identify issues in the argumentation rather than as in our case elaborating on the status of the main claim.

Using the defeasible logic to capture concepts of the argumentation game is supported by (Letia & Vartic 2006; Nilsson, Eriksson Lundström, & Hamfelt 2005) and recently (Thakur *et al.* 2007; Eriksson Lundström *et al.* 2008). (Letia & Vartic 2006) 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). (Nilsson, Eriksson Lundström, & Hamfelt 2005) provides an extension of defeasible logic to include the step of the adversarial dialogue by defining a metaprogram for an alternative computational algorithm for ambiguity propagating defeasible logic 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 (Thakur *et al.* 2007; Eriksson Lundström *et al.* 2008). In our work, we define the notion of majority acceptance and a method to weight arguments. In (Thakur *et al.* 2007), the strength of unchallenged rules is upgraded over iterations. That is the conclusions supported by these rules are not rebutted by the current iteration, these conclusions are unarguable in follow iterations. The upgrade is applied to all participants during iterations of the argumentation game. (Eriksson Lundström *et al.* 2008) distinguishes participants of the argumentation game. That is one participant must provide a strong argument (i.e. a definite proof) in order to defeat arguments from other participants. Both of the works do not directly handle the challenge coming from multiple participants.

We extend the protocol of a argumentation game to settle on a common goal. The termination condition of our framework is either there is no more argument to rebut or an agent can pass its proposal at one iteration. 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 (Amgoud, Dimopoulos, & Moraitis 2007) provides a unified and general formal framework for the argumentation-based negotiation dialogue between two agents for a set of offers. The work provides a formal connection between the status of a argument including accepted, rejected, and undecided with possible actions of an agent (accept, reject, and negotiate respectively). One important feature of the framework is that this representation is independent with logical languages modelling knowledge of an agent. Moreover, an agent's knowledge is evolved by accumulating arguments during interactions.

We have advantages of using the defeasible logic since it provides us an elegant tool to naturally capture the above statuses of arguments. Accepted, rejected, undecided conditions can be simulated by the proof conditions of defeasible reasoning w.r.t ambiguity of premises. If the preference over

knowledge sources is known, the accepted and rejected arguments is corresponding to $(+\partial, -\partial)$ using the feature of ambiguity blocking. Otherwise, three conditions of arguments are derived from $(+\partial, -\partial)$ and $+\Sigma$). These notions correspond to the existence of a positive proof, a negative proof, and a positive support of a premise. In addition, defeasible logic provides a compact representation to accommodate new information from other agents.

From the perspective of coordination among agents, (Parsons & McBurney 2003) presents an argumentation based communication, 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. Also, (Rueda, Garcia, & Simari 2002) proposes a communication mechanism based on argumentation for collaborative BDI agents, in which agents exchange their proposals and counter-proposals in order to reach a mutual agreement. During the course of conversations, an agent can retrieve missing literals (regarded as sub-goals) or fulfil its goals by requesting collaboration from other agents. However, these works did not clearly show how an agent can tackle conflicts from multiple agents, especially when the preference over exchanged arguments is unknown.

The main difference in our framework is the external model where more than two agents can argue to settle on a common goal. Since there is no preference over the proposal of individual agents, the majority rule enables the group to identify the majority preference over individual claims. On one hand, we present the notion of the acceptance by the majority of agents. On the other hand, this notion relaxes the complexity of n-persons argumentation game by partitioning agents into two sub-groups: one supports the major claim; the other opposes it. Moreover, the majority rule allows an agent to probe the attitudes of the group in order to dynamically create a preference over its defensive arguments if its main claim is not accepted by the majority of agents. The strategy to defend against the topic of the dialogue is to attack the most common premise among the arguments supporting the topic.

In our framework, an individual agent efficiently tackle with conflicts from multiple sources of knowledge owing to the use of the defeasible logic as the underlying logic. The construction of arguments requires an individual agent to integrate the background knowledge commonly shared among agents, knowledge from other agents, and its private knowledge. The background knowledge has the priority over the other sources, therefore when integrating any conflict with this knowledge is blocked. Since all agents are equally trustful, the knowledge from other agents has the same weight. To achieve a consensus from knowledge of other agents and to discover "collective wisdom", the ambiguity propagation is applied over all knowledge sources of an individual agent.

Conclusions

This paper has presented an n-person argumentation framework based on the defeasible logic. In the framework, we propose an external model based on the argumentation/dialogue game which enables agents in a group settle

on a common goal. An agent proposes its goal including the explanation and argues with other agents about the goal. At the termination, the group identifies a common goal accepted by the majority of the group and the supportive argument for the goal.

We also propose an internal model of an agent where an individual agent can efficiently construct arguments from multiple sources of knowledge including the background knowledge presenting the common constraints and expectations of the group, knowledge from others which is evolved during iterations, and its private knowledge. The background knowledge is preferred over the other sources of knowledge. Due to the flexibility of defeasible logic in tackling the ambiguous information, these types of knowledges can be efficiently integrated with the private knowledge of an agent (with or without a preference over the knowledge sources) to generate and justify its arguments. The majority rule relaxes the complexity of n-persons argumentation dialogue game. This rule is used to identify the topic of the dialogue among the claims of agents. That is the majority acceptance of an argument. Also, an agent can use the majority rule as a method to select an argument which challenges the major number of agents in order to better defend its goal.

References

- [Amgoud, Dimopoulos, & Moraitis 2007] Amgoud, L.; Dimopoulos, Y.; and Moraitis, P. 2007. A unified and general framework for argumentation-based negotiation. In *AA-MAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, 1–8. New York, NY, USA: ACM.
- [Antoniou *et al.* 2000] Antoniou, G.; Billington, D.; Governatori, G.; and Maher, M. J. 2000. A flexible framework for defeasible logics. In *Proc. American National Conference on Artificial Intelligence (AAAI-2000)*, 401–405.
- [Antoniou *et al.* 2001] Antoniou, G.; Billington, D.; Governatori, G.; and Maher, M. J. 2001. Representation results for defeasible logic. *ACM Transactions on Computational Logic* 2(2):255–287.
- [Antoniou *et al.* 2006] Antoniou, G.; Billington, D.; Governatori, G.; and Maher, M. J. 2006. Embedding defeasible logic into logic programming. *Theory and Practice of Logic Programming* 6(6):703–735.
- [Billington 1993] Billington, D. 1993. Defeasible logic is stable. *Journal of Logic and Computation* 3:370–400.
- [Dung 1995] Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77(2):321–358.
- [Eriksson Lundström *et al.* 2008] Eriksson Lundström, J.; Governatori, G.; Thakur, S.; and Padmanabhan, V. 2008. An asymmetric protocol for argumentation games in defeasible logic. In *10 Pacific Rim International Workshop on Multi-Agents*, volume 5044. Springer.
- [Governatori *et al.* 2004] Governatori, G.; Maher, M. J.; Antoniou, G.; and Billington, D. 2004. Argumentation Semantics for Defeasible Logic. *J Logic Computation* 14(5):675–702.
- [Jennings *et al.* 1998] 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*.
- [Letia & Vartic 2006] Letia, I. A., and Vartic, R. 2006. 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*, 359–362. Washington, DC, USA: IEEE Computer Society.
- [Lial 1998] Lial, T. B.-C. 1998. 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: Jurix: The Eleventh Conference*, 5–20. Nijmegen: Gerard Noodt Instituut.
- [Lin 1996] Lin, J. 1996. Integration of weighted knowledge bases. *Artificial Intelligence* 83:363–378.
- [Lodder 2000] Lodder, A. R. 2000. Thomas F. Gordon, the pleadings game - an artificial intelligence model of procedural justice. *Artif. Intell. Law* 8(2/3):255–264.
- [Maher *et al.* 2001] Maher, M. J.; Rock, A.; Antoniou, G.; Billington, D.; and Miller, T. 2001. Efficient defeasible reasoning systems. *International Journal of Artificial Intelligence Tools* 10(4):483–501.
- [Maher 2001] Maher, M. J. 2001. Propositional defeasible logic has linear complexity. *Theory and Practice of Logic Programming* 1(6):691–711.
- [Nilsson, Eriksson Lundström, & Hamfelt 2005] Nilsson, J. F.; Eriksson Lundström, J.; and Hamfelt, A. 2005. A metalogic formalization of legal argumentation as game trees with defeasible reasoning. In *Proceedings of ICAIL'05, International Conference on AI and Law*, Proceedings of ICAIL. ACM.
- [Parsons & McBurney 2003] Parsons, S., and McBurney, P. 2003. Argumentation-based dialogues for agent coordination, group decision and negotiation. *Group Decision and Negotiation* (12):415–439.
- [Prakken & Sartor 1996] Prakken, H., and Sartor, G. 1996. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law* 4:331–368.
- [Rueda, Garcia, & Simari 2002] Rueda, S. V.; Garcia, A. J.; and Simari, G. R. 2002. Argument-based negotiation among bdi agents. *Journal of Computer Science and Technology* 2(7).
- [Thakur *et al.* 2007] Thakur, S.; Governatori, G.; Padmanabhan, V.; and Eriksson Lundström, J. 2007. Dialogue games in defeasible logic. In *20th Australian Joint Conference on Artificial Intelligence, AI 2007*, volume 4830, 497–506. Springer.