Paretian Argumentation Frameworks for Pareto Optimal Arguments[☆]

Hiroyuki Kido[a,*], Yukio Ohsawa[a], Katsumi Nitta[b]

[a] *The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
{kido,ohsawa}@sys.t.u-tokyo.ac.jp, +81-3-5841-8449
[b] *Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503, Japan*
nitta@dis.titech.ac.jp, +81-45-924-5214

Abstract

Argument-based reasoning offers promising interaction and computation mechanisms for multi-agent negotiation and deliberation. Arguments in this context are typically statements of beliefs or actions related to agents' subjective values, preferences, and so on. Consequences of such arguments can and should be evaluated using various criteria, and therefore, it is desirable that semantics supports these criteria as principles for accepting arguments. This paper gives an instance of Dung's abstract argumentation framework to deal with Pareto optimality, i.e., a fundamental criterion for social welfare. We show that the instance allows Dung's acceptability semantics to interpret Pareto optimal arguments, without loss of generality. We discuss the prospects of justified Pareto optimal arguments and Pareto optimal extensions.

Keywords: argumentation, acceptability semantics, abstract argumentation frameworks, Pareto optimality

1. Introduction

Dung's acceptability semantics [4], often called Dungean semantics, gives unified principles to reformulate consequence notions of various kinds of non-

[☆]This paper is a substantial extension and revision of the conference papers [1, 2] and Japanese journal paper [3].

[*]Corresponding author

monotonic reasoning. In the last fifteen years, a great number of papers in the field of argumentation in artificial intelligence have addressed a variety of problems by means of abstract argumentation frameworks instantiated with suitable logical languages and inference rules, and demand solutions from his semantics. Bench-Capon [5] describes that 2000 to 2007 witnessed an intensive study developing Dung's theory in various directions. While at the same time, a number of research have addressed limitations of Dung's theory. On the one hand, Dung's abstract argumentation frameworks have been extended to treat variety of notions associated with argumentation [6, 7, 8, 9, 10], and on the other hand, Dungean semantics has been refined to interpret detailed consequence notions of argumentation [11, 12, 13, 14, 15]. Moreover, some researcher argues that limitations of Dung's theory come from studying argumentation in an abstract level [16]. However, Dung's theory is still well-used semantics due to its generality. It is true that abstractness and generality are two sides of the same coin. In fact, bridging multidisciplinary fields such as nonmonotonic reasoning, n-person games and stable marriage problem is impossible without abstract studies of argumentation.

In this paper, we argue that semantics for argumentation should support welfare criteria for acceptable arguments. Arguments in the context of multi-agent negotiation or deliberation involve beliefs and actions associated with values and preferences that differ from an agent to another. In this context, agents do not necessarily pursue the truth, but can sometimes pursue a common ground or compromise taking into account individual agents' values or preferences. This observation leads us to the critical opinion that Dungean semantics and their modifications have no mention of welfare criteria, although they definitely give certain types of criteria for acceptable arguments. In fact, Prakken [17] sees Dung's theory as principles to judge whether propositions are acceptable as true. Moreover, Rahwan and Larson [18] argue for the importance of welfare semantics for argumentation.

This observation motivates us to figure out the relationship between Dungean semantics and Pareto optimality. Pareto optimality or efficiency is known as a fundamental criterion for social welfare and an outcome is said to be Pareto optimal if no agents can be made better off without making someone else worse off. For example, outcomes of practical reasoning, i.e., reasoning about what to do, are courses of action that an agent or a group of agents takes, and the decisions depend on desires, aims, or values individual agents have. So, they are certain to avoid Pareto improvable decisions because otherwise there exists another decision that makes some agents better off and no one worse off.

This paper gives a particular kind of abstract argumentation frameworks, called Paretian argumentation frameworks. An abstract argumentation framework is a tuple of a set of arguments and a binary relation on the set. It is abstract in the sense that no internal structures of arguments nor a basis of attack relations is assumed there. A Paretian argumentation framework is its specialization in the sense that it is an abstract argumentation framework whose attack relation is substituted by the union of the inverse complements of agents' preference relations on the set of arguments. Namely, attacking is invoked when an argument is not worse than another. In this paper, we use the word "defeat," instead of "attack," with the intention of expressing not only the original meaning of attacks, i.e., contradiction, implicitly assumed in abstract argumentation frameworks, but also preference relations between arguments. However, in principle, we do not substitute the defeat relation into a combination of the preference and attack relations in order to avoid getting involved with the issue of inconsistency [16], i.e., the problem of combining attack and preference relations causing conflict in extensions. We show that preferred extensions of Paretian argumentation frameworks coincide with Pareto optimal solutions.

The contributions of this paper are summarized as follows. First, we show the relationships between Dung's notion of acceptability, emerging from studies of nonmonotonic logic, and Pareto's notion of Pareto optimality, emerging from studies of economics. In particular, we show that Pareto optimality can be identified by a fixed point of the characteristic function given by Dungean semantics. This implies the fact that, in a particular case, Dungean semantics gives a preference interpretation in term of Pareto optimality. Second, we give a foundation for developing multi-agent negotiation and deliberation dialogues searching for Pareto optimal solutions and their expansions, i.e., justified Pareto optimal arguments and Pareto optimal extensions.

This paper is organized as follows. Section 2 gives a motivating example for Pareto optimality in argumentation. Section 3 defines a Paretian argumentation framework and Section 4 evaluates the framework by relating it to Pareto optimality. In Section 5, we show prospects of the framework and discuss related work. In Section 6, we describe conclusions and future work.

2. Motivating Example

Let us consider simple argument by which agents i and j try to decide which apartment they buy. They have common concerns about safety, access to transportation and quietness, but their preferences are different. Agent i prefers safety rather than quietness and quietness rather than transportation. Agent j prefers transportation rather than safety and safety rather than quietness. The following arguments are assumed to be put forward by agents i and j at some point.

Argument a **:** We ought to buy apartment A because living in a safe area is desirable and the apartment is located in a safe area.

Argument b **:** We ought to buy apartment B because good access to transportation is desirable and the apartment has good access to transportation.

Argument c**:** We ought to buy apartment C because living in a quite area is desirable and the apartment is located in a quite area.

Which apartment should rational agents buy, or more generally, what is the consequence of this argument? We think that rational agents are at least certain to decide to buy a socially good apartment. Pareto optimality is a fundamental criterion for goodness, and a solution is Pareto optimal if no agents can be made better off without making someone else worse off. In the above example, buying apartment A or B, i.e., accepting argument a or b, is Pareto optimal because i prefers safety to quietness, and quietness to transportation, while j prefers transportation to safety, and safety to quietness.

Our idea is to characterize arguments a and b supporting Pareto optimal solutions as arguments in an extension, i.e., as arguments successfully defended from defeating arguments. We will show that this is achieved by substituting the inverse complement of agents' preference relations on the set of arguments, into the defeat relation of an abstract argumentation framework. Namely, the defeat relation is substantiated with the idea that an argument is not worse than another. For example, a defeats c, c defeats b and a defeats b for agent i because a is not worse than c, c is not worse than b and a is not worse than b for i, in terms of i's preference on arguments. On the other hand, b defeats a, a defeats c and b defeats c for agent j in terms of j's preference. Figure 1 shows the abstract argumentation framework where

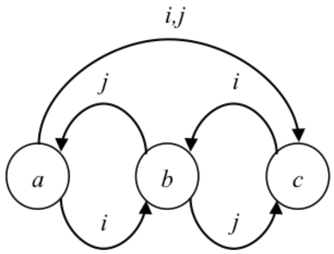

Figure 1: Interpretation of Pareto optimality based on an abstract argumentation

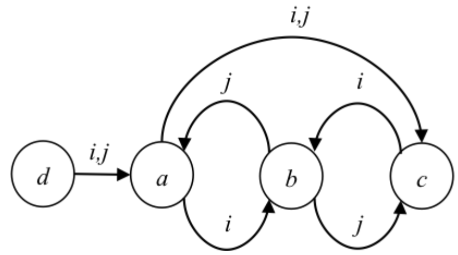

Figure 2: Consideration on justification of arguments

the defeat relation is substantiated by the union of their defeat relations. For readability, each arrow has an extra information specifying whose defeat relation activates. It is the case that both a and b are successfully defended from defeating arguments, but c is not.

Now, we further the idea that consequences of argument about what to do depend on consequences of argument about what to believe. For example, let us consider the following argument put forward by agent j.

Argument d **:** Apartment A is not located in a safe area because a murder occurred and the murderer is still at large.

Now, which apartment should rational agents buy? Argument a fails to justify buying apartment A in terms of Dungean semantics. In fact, a is not in any extension because it is defeated by argument d, but no argument defeats d. Figure 2 shows this situation where the defeat relation from d to a represents contradiction, but not preference. In terms of both Pareto optimality and Dungean semantics, it is obvious that buying apartment B or C, i.e., accepting argument b or c, becomes Pareto optimal because although j prefers b to c, agent i prefers c to b. What is interesting here is that argument c becomes Pareto optimal only by considering theoretical argument d. The new semantics evaluates b and c as acceptable on the ground that they are Pareto optimal within justified arguments, i.e., justified Pareto optimal arguments.

In this paper, we assume that arguments in our framework can be divided into theoretical arguments and practical arguments. The latter ones represent options/choices (e.g. buying apartment A, B, or C) over which agents have preferences, and these preferences are translated into defeat relations according to these agents. On the other hand, agents do not have preferences over theoretical arguments. Those arguments exist to defeat/reinstate other practical arguments (e.g. argument d in this example). Theoretical arguments can defeat theoretical and practical arguments while practical arguments can only defeat practical arguments (by translating preference relations). All agents agree on the (non-)existence of defeats from theoretical arguments, unlike the defeats between practical arguments.

3. Paretian Argumentation Frameworks

This section aims to introduce a Paretian argumentation framework. It is defined as a particular kind of an abstract argumentation framework. Dung's theory of abstract argumentation [4] reformulates consequence notions of various sorts of nonmonotonic reasoning. It gives four kinds of argumentation-theoretic notions of extensions called preferred, stable, grounded and complete extensions, defined as follows.

Definition 1 (Abstract argumentation). [4] An abstract argumentation framework is a pair $AF = \langle AR, defeats \rangle$ where AR is a set of arguments, and $defeats$ is a binary relation on AR, i.e. $defeat \subseteq AR \times AR$.

- A set $S \subseteq AR$ of arguments is said to be *conflict-free* if there are no arguments $a, b \in S$ such that a defeats b, i.e., $(a, b) \notin defeat$, for all $a, b \in S$.
- An argument $a \in AR$ is *acceptable* with respect to a set $S \subseteq AR$ of arguments iff for each argument $b \in AR$: if b defeats a then b is defeated by an argument in S.
- A conflict-free set of arguments S is *admissible* iff each argument in S is acceptable with respect to S.
- A *preferred extension* of AF is a maximal (with respect to set inclusion) admissible set of AF.
- A conflict-free set of arguments S is called *stable extension* of AF iff S defeats each argument which does not belong to S.
- The *characteristic function*, $F_{AF} : Pow(AR) \rightarrow Pow(AR)$, of AF is defined by $F_{AF}(S) = \{a \mid a \textit{ is acceptable with respect to } S\}$.

- The *grounded extension* of AF is the least fixed point of F_{AF}.

- An admissible set S of arguments is called a *complete extension* of AF iff each argument, which is acceptable with respect to S, belongs to S.

Each preferred, stable, grounded and complete extension is said to be an extension given by preferred, stable grounded and complete semantics, respectively, and we collectively call them Dungean semantics.

We assume an arbitrary but fixed abstract argumentation framework $AF = \langle AR, defeat\rangle$. In addition, we assume a set of agents and agent's preference relation defined as a preorder, i.e., a binary relation satisfying reflexivity and transitivity[1]. AG denotes the set of agents and $\succsim_i \subseteq AR \times AR$ denotes the preference relation, for all $i \in AG$. As usual, the complement of $\succsim_i$ is denoted and defined as $\not\succsim_i = \{(x, y) \in AR \times AR | (x, y) \notin \succsim_i\}$ and the inverse of $\succsim_i$ is $\precsim_i = \{(x, y) \in AR \times AR | (y, x) \in \succsim_i\}$. So, the inverse complement of $\succsim_i$ is denoted and defined as $\not\precsim_i$.

Now, we define a particular kind of abstract argumentation frameworks, called Paretian argumentation frameworks.

Definition 2 (Paretian argumentation framework). A *Paretian argumentation framework* is a pair $AF = \langle AR, defeat\rangle$ where AR is a set of arguments and $defeat = \bigcup_{i \in AG} \not\precsim_i$.

A Paretian argumentation framework is characterized by the defeat relation defined as the union of the inverse complements of agents' preference relations.

Example 1. Consider the set of arguments, agents' preference relations defined as follows.

$$
\begin{aligned}
AR &= \{a, b, c, d\} \\
\succsim_i &= \{(a, a), (b, b), (c, c), (d, d), (b, a), (b, d), (c, a), (c, b), (c, d)\} \\
\succsim_j &= \{(a, a), (b, b), (c, c), (d, d), (a, c), (a, d), (b, a), (b, c), (c, d), (b, d), (d, c)\}
\end{aligned}
$$

Note that the preference relations are preorder, but not partial order. In fact, $c \succsim_j d$ and $d \succsim_j c$ are the case, but $c \neq d$. Thus, $AF = \langle AR, \not\precsim_i \cup \not\precsim_j\rangle$ is a

[1]Normally, a preference relation is assumed to be a partial order, i.e., a binary relation satisfying reflexivity, antisymmetry and transitivity. We, however, relax it to develop our idea as general as possible.

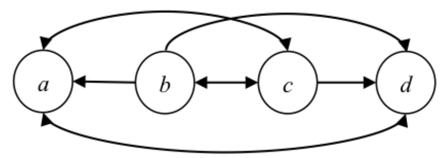

Figure 3: Paretian argumentation framework $\langle AR, \not\gtrsim_i \cup \not\gtrsim_j \rangle$

Paretian argumentation framework where $\not\gtrsim_i$ and $\not\gtrsim_j$ are defined as follows.

$$\begin{aligned}\not\gtrsim_i &= \{(a,d),(b,a),(b,d),(c,a),(c,b),(c,d),(d,a)\}\\ \not\gtrsim_j &= \{(a,c),(a,d),(b,a),(b,c),(b,d)\}\end{aligned}$$

Figure 3 shows AF represented by a directed graph where each node represents an argument and each edge from x to y represents either $x \not\gtrsim_i y$ or $x \not\gtrsim_j y$.

4. Correctness of Paretian Argumentation Frameworks

Welfare economics is a branch of economics that is concerned with the evaluation of alternative economic situations (states, configurations) from the point of view of the society's well being [19]. One of fundamental criteria for evaluating society's well being is Pareto optimality, or Pareto efficiency. Let O be a set of outcomes, I be a set of agents and $\geq_i$ be agent i's *preference* represented by a partial order, i.e., a binary relation satisfying reflexivity, antisymmetry and transitivity, on O. The notion of Pareto optimality is defined on the notion of Pareto dominance.

Definition 3 (Pareto dominance). An outcome $o_1 \in O$ Pareto dominates outcome $o_2 \in O$ iff, for all $i \in I$, $o_1 \geq_i o_2$ and there exists $j \in I$ such that $o_1 >_j o_2$.

An outcome is said to be Pareto optimal if no outcome dominates it. In Pareto optimal outcomes, no agents can be made better off without making someone else worse off.

Definition 4 (Pareto optimality). An outcome $o_1 \in O$ is Pareto optimal (or Pareto efficient) iff o_1 is not Pareto dominated by any ouctome in O.

In order to develop our idea as general as possible, we relax the preference relation assumed in the definition of Pareto optimality to a preorder.

Now, we start analyzing relationships between Pareto optimality and Dung's semantics.

Lemma 1. *Let $AF = \langle AR, defeat \rangle$ be a Paretian argumentation framework. An argument $a \in AR$ is in a preferred extension of AF iff $\{a\}$ is an admissible set.*

PROOF. ($\Leftarrow$) From Definition 1, a preferred extension is a maximal admissible set. So, if $\{a\}$ is admissible then there exists a preferred extension S such that $\{a\} \subseteq S$. ($\Rightarrow$) We show that the contradiction is derived from the assumptions that a is in a preferred extension S and $\{a\}$ is not admissible. Under the assumptions, there exists an argument $b \in AR$ defeating a, that is not defeated by a and is defeated by a third argument $c \in S \setminus \{a\}$. Formally, the following formula holds.

$$\exists b \in AR \exists c \in S.(b,a) \in defeat \wedge (a,b) \notin defeat \wedge (c,b) \in defeat \quad (1)$$

Since the complement of $defeat$ is transitive, the following formulae hold.

$$\begin{aligned} & \forall a,b,c \in AR.(a,b) \notin defeat \wedge (c,a) \notin defeat \rightarrow (c,b) \notin defeat \\ \Leftrightarrow \quad & \forall a,b,c \in AR.(a,b) \notin defeat \wedge (c,b) \in defeat \rightarrow (c,a) \in defeat \quad (2) \end{aligned}$$

The following formula is derived from (1) and (2).

$$\exists b \in AR.(b,a) \in defeat \wedge \exists c \in S.(c,a) \in defeat \quad (3)$$

(3) implies $(c,a) \in defeat$ where $a, c \in S$. This contradicts the assumption that S is conflict-free. $\square$

The following lemma shows the relationship between admissible sets of arguments and Pareto optimality.

Lemma 2. *Let O be a set of outcomes, I be a set of agents, and $\succsim_i$ be i's preference represented by a preorder on O, for all $i \in I$. $o \in O$ is Pareto optimal iff $\{o\}$ is an admissible set in the Paretian argumentation framework $\langle AR, defeat \rangle$ where $AR = O$ and $defeat = \bigcup_{i \in I} \not\precsim_i$.*

PROOF. Based on the assumption that $\succsim_i$ is a preorder, we show that the outcome o in an admissible set turns out to be Pareto optimal.

$$
\begin{aligned}
& \nexists i \in I(o \not\succsim_i o) \wedge \forall o_1 \in O(\exists i \in I(o_1 \not\succsim_i o) \rightarrow \exists j \in I(o \not\succsim_j o_1)) \\
\Leftrightarrow\; & \forall o_1 \in O(\exists i \in I(o_1 \not\succsim_i o) \rightarrow \exists j \in I(o \not\succsim_j o_1)) \\
\Leftrightarrow\; & \forall o_1 \in O(\exists i \in I(o_1 \succ_i o \vee o \not\succsim_i o_1 \wedge o_1 \not\succsim_i o) \rightarrow \exists j \in I(o \not\succsim_j o_1)) \\
\Leftrightarrow\; & \forall o_1 \in O(\exists i \in I(o_1 \succ_i o) \vee \\
& \exists k \in I(o \not\succsim_k o_1 \wedge o_1 \not\succsim_k o) \rightarrow \exists j \in I(o \not\succsim_j o_1)) \\
\Leftrightarrow\; & \forall o_1 \in O((\exists i \in I(o_1 \succ_i o) \rightarrow \exists j \in I(o \not\succsim_j o_1)) \wedge \\
& (\exists k \in I(o \not\succsim_k o_1 \wedge o_1 \not\succsim_k o) \rightarrow \exists l \in I(o \not\succsim_l o_1))) \\
\Leftrightarrow\; & \forall o_1 \in O(\exists i \in I(o_1 \succ_i o) \rightarrow \exists j \in I(o \not\succsim_j o_1)) \\
\Leftrightarrow\; & \nexists o_1 \in O(\exists i \in I(o_1 \succ_i o) \wedge \forall j \in I(o \precsim_j o_1)) \qquad (4)
\end{aligned}
$$

(4) is equivalent to the definition of Pareto optimality. □

Lemma 1 and Lemma 2 lead to the following theorem.

Theorem 1. *Let O be a set of outcomes, I be a set of agents, and $\succsim_i$ be i's preference relation represented by a preorder on O, for all $i \in I$. $o \in O$ is Pareto optimal iff o is in a preferred extension of the Paretian argumentation framework $AF = \langle AR, defeat \rangle$ where $AR = O$ and $defeat = \bigcup_{i \in I} \not\succsim_i$.*

Theorem 1 shows that evaluation of Pareto optimal solutions can be translated to evaluation of preferred extensions of a particular Paretian argumentation framework. In other words, preferred semantics credulously justifies Pareto optimal solutions.

5. Discussions on Paretian Argumentation Frameworks

5.1. Prospect on Justified Pareto optimal arguments

Paretian argumentation frameworks have the ability to make Pareto optimality to more realistic. Basically, Pareto optimality implicitly assumes mutually incompatible solutions and no reason nor justification supporting the solutions is considered. By contrast, handling Pareto optimality by Paretian argumentation frameworks allows us to evaluate Pareto optimal solutions supported by a justified reasons. This idea is developed on a combination of Paretian argumentation framework and Dung's abstract argumentation framework where Dungean semantics interprets Pareto optimal arguments in the former framework and justified arguments in the latter.

We assume that a practical argument is informally distinguished from a theoretical argument where a practical argument is intuitively explained as an argument whose conclusion is a statement about a proposal related to an action, while a theoretical argument is an argument whose conclusion is a statement about proposition related to truth. A formal distinction needs internal structures of arguments, but they are beyond the scope of this paper. We only show examples of practical and theoretical arguments and assume disjoint sets $Targ \subseteq AR$ of theoretical arguments and $Parg \subseteq AR$ of practical arguments taken as given.

Example 2. The following statement is a practical argument.

> *Living in a safe area is desirable. An apartment A is located in a safe area and living at A is sufficient for us to live in a safe area. Therefore, we ought to live at the apartment A.*

The following statement is a theoretical argument.

> *The apartment A is not located in a safe area because a murder occurred in the area and the murderer is still at large.*

It is natural to think that practical arguments cannot defeat theoretical arguments although theoretical arguments can defeat both theoretical and practical arguments. Formally, we assume an attack relation $attack \subseteq Targ \times (Targ \cup Parg)$ to represent incompatibility between arguments. It is also natural to assume that agents have preferences on practical arguments. Formally, we assume a preorder $\succsim_i \subseteq Parg \times Parg$ to represent agent i's preference, for all agents $i \in AG$. We now consider the abstract argumentation framework AF^* and Paretian argumentation framework F^* defined as follows.

$$\begin{aligned} AF^* &= \langle Targ \cup Parg, attack \rangle \\ F^* &= \langle Parg, \bigcup_{i \in AG} \not\precsim_i \rangle \end{aligned}$$

Theorem 1 implies the fact that preferred extensions of F^* coincide with Pareto optimal arguments. We combine a Paretian argumentation framework and an abstract argumentation framework where the set of arguments in

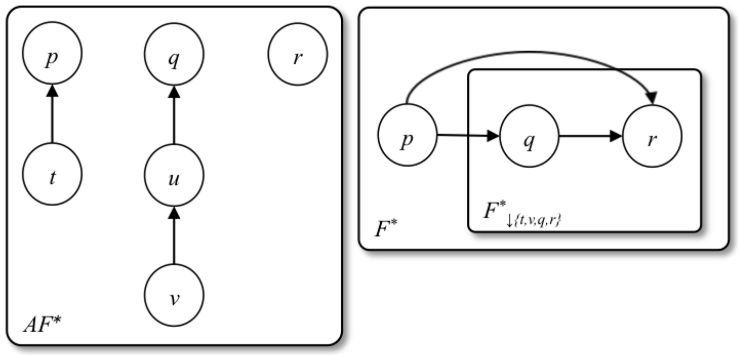

Figure 4: $F^*_{\downarrow\{t,v,q,r\}}$ obtained by combining AF^* and F^*

the Paretian argumentation framework is restricted to justified arguments in the abstract argumentation framework. A notion of restriction is defined as follows.

Definition 5 (Restriction). Let $AF = \langle AR, defeat \rangle$ be an abstract argumentation framework and $S \subseteq AR$ be a set of arguments. The *restriction* of F to S, denoted by $F_{\downarrow S}$, is defined as the tuple $\langle AR \cap S, defeat^* \rangle$ where $defeat^* = defeat \cap (S \times S)$.

We define a justified Pareto optimal argument as a Pareto optimal arguments restricted to justified arguments.

Definition 6 (Justified Pareto optimal argument). $a \in AR$ is a *justified Pareto optimal argument* iff a is in a preferred extension of $F^*_{\downarrow S}$ where S is the grounded extension of AF^*.

Example 3. Figure 4 shows examples of $AF^* = \langle Targ \cup Parg, attack \rangle$ and $F^* = \langle Parg, AG, \bigcup_{i \in AG} \not\precsim_i \rangle$ where $Targ = \{t, u, v\}$, $Parg = \{p, q, r\}$, $attack = \{(t,p), (u,q), (v,u)\}$, $\succsim_i = \{(p,p), (q,q), (r,r), (p,q), (p,r), (q,r), (r,q)\}$ and $\succsim_j = \{(p,p), (q,q), (r,r), (p,q), (p,r), (q,r)\}$. We can see that $\{t, v, q, r\}$ is the grounded extension of AF^*. Figure 4 also shows the restriction of F^* to $\{t, v, q, r\}$, i.e., $F^*_{\downarrow\{t,v,q,r\}} = \langle \{q, r\}, \{(q, r)\} \rangle$. We can see that $\{q\}$ is the preferred extension of $F^*_{\downarrow\{t,v,q,r\}}$. Therefore, q is the justified Pareto optimal argument. Note that no consideration on justification yields p as a Pareto optimal argument since $\{p\}$ is the preferred extension of F^*.

5.2. Prospect on Pareto optimal extensions

In general, different agents can have conflicting views of acceptable arguments even though they see the same argumentation. So if the argumentation is taken place in the context of consensus building or agreement making, then a certain kind of criterion is required to reconcile conflicts. Pareto optimality is an effective criterion for the purpose and we show that combining Pareto optimality with Dung's extensions allows agents to choose Pareto optimal extensions in terms of society's well-being. We assume that each agent has a preference on sets of arguments, and it is defined as a preorder $\succsim_i \subseteq Pow(AR) \times Pow(AR)$, for all $i \in AG$. We now consider the abstract argumentation framework AF^+ and Paretian argumentation framework F^+ defined as follows.

$$\begin{aligned} AF^+ &= \langle AR, defeat \rangle \\ F^+ &= \langle Pow(AR), \bigcup_{i \in AG} \not\precsim_i \rangle \end{aligned}$$

Theorem 1 implies the fact that preferred extensions of F^+ coincide with sets of Pareto optimal sets of arguments. We combine a Paretian argumentation framework and an abstract argumentation framework where the power set of the set of arguments in the Paretian argumentation framework is restricted to extensions of the abstract argumentation framework. A notion of a Pareto optimal extension is defined as Pareto optimal sets of arguments where the sets of arguments are restricted to extensions.

Definition 7 (Pareto optimal extension). $S \in Pow(AR)$ is a Pareto optimal extension iff S is a preferred extension of $F^+_{\downarrow T}$ where T is a set of extensions of AF^+.

Example 4. Figure 5 shows examples of $AF^+ = \langle AR, defeat \rangle$ and $F^+ = \langle Pow(AR), \bigcup_{i \in AG} \not\precsim_i \rangle$ where $AR = \{a, b\}$, $defeat = \{(a, b), (b, a)\}$, $\succsim_i = \{(\emptyset, \emptyset), (\{a\}, \{a\}), (\{b\}, \{b\}), (\{a, b\}, \{a, b\}), (\{a, b\}, \{a\}), (\{a, b\}, \{b\}), (\{a, b\}, \emptyset), (\{a\}, \{b\}), (\{a\}, \emptyset), (\{b\}, \{a\}), (\{b\}, \emptyset)\}$ and $\succsim_j = \{(\emptyset, \emptyset), (\{a\}, \{a\}), (\{b\}, \{b\}), (\{a, b\}, \{a, b\}), (\{a, b\}, \{a\}), (\{a, b\}, \{b\}), (\{a, b\}, \emptyset), (\{a\}, \{b\}), (\{a\}, \emptyset), (\{b\}, \emptyset)\}$. We can see that $\emptyset$, $\{a\}$ and $\{b\}$ are all extensions of AF^+. Figure 5 also shows the restriction of F^+ to $\{\emptyset, \{a\}, \{b\}\}$, i.e., $F^+_{\downarrow\{\emptyset,\{a\},\{b\}\}} = \langle \{\emptyset, \{a\}, \{b\}\}, \{(\{a\}, \{b\}), (\{a\}, \emptyset), (\{b\}, \emptyset)\} \rangle$. We can see that $\{a\}$ is the preferred extension of $F^+_{\downarrow\{\emptyset,\{a\},\{b\}\}}$. Therefore, $\{a\}$ is the Pareto

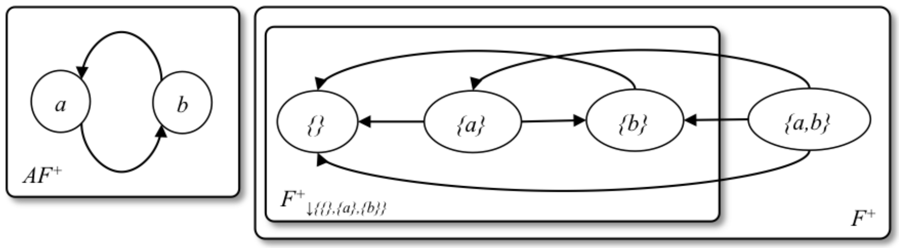

Figure 5: $F^+_{\downarrow\{\emptyset,\{a\},\{b\}\}}$ obtained by combining AF^+ and F^+

optimal extension. Note that no consideration on extension yields $\{a, b\}$ as a Pareto optimal set of arguments since $\{a, b\}$ is the preferred extension of F^+.

5.3. Related work

Some researchers working on argumentation-theoretic semantics introduce nonclassical semantics such as stage semantics [11], semi-stable semantics [12], ideal semantics [13], CF2 semantics [14], prudent semantics [15], and so on. These semantics are defined on Dung's abstract AF (argumentation frameworks) and aim to overcome or improve some limitations or drawbacks of Dungean semantics. Others extend Dung's abstract AF to value-based AF [6], preference-based AF [7, 8], bipolar AF [9], extended AF [10], and so on. These frameworks aim to enhance expressive power of argumentation frameworks taking into account various attributes pertaining to argumentation such as values, preferences, supports, attacks to attacks, respectively. Our Paretian AF is motivated by our novel idea relating Dung's theory and Pareto optimality. In particular, we agree with the standpoints on value-based and preference-based AF [6, 7] that consideration on values and preferences is necessary for handling persuasiveness. However, this paper does not stand on the position that values and preferences can cancel out inconsistency. This is consistent with the claim [8] that handling preferences in argumentation frameworks can cause a problem of inconsistency. The main thesis in this paper is that extensions of Paretian argumentation frameworks coincide with Pareto optimal arguments when defeat relations are appropriately instantiated by preference relations. The problems of how to instantiate arguments to avoid inconsistency are outside the scope of this paper.

Rahwan and Larson [18] argue for the need for welfare semantics for argumentation and explore the relationships between extensions defined by Dungean semantics and Pareto optimality. Dung and Thang [20] also introduce the notion of Pareto optimality in a process of argument-based negotiation. However, identifying Pareto optimal solutions themselves is not a subject for computation in their research.

Paretian AF is applicable to argument-based deliberation. Deliberation is a type of dialogue in which a group of agents or a single agent tries, through looking at a set of alternatives, to make a decision about which course of action among the possible alternatives to take [21]. Many argument-based approaches for deliberation or practical reasoning apply Dungean semantics as a fundamental principle for evaluating acceptable arguments. For instance, a problem of deciding single agent's course of action is formalized as instantiations of abstract argumentation frameworks where the agent has more than one desires. In fact, Bench-Capon and Prakken [22] propose two kinds of practical reasoning, positive and negative practical syllogisms, deriving desirable and undesirable actions, respectively. Dungean semantics is used for evaluating arguments, and consequently decides what the best action is. Prakken [23] gives a combined formalization for skeptical epistemic reasoning interleaved with credulous practical reasoning. The paper distinguishes practical arguments from theoretical arguments by informally dividing logical formulas into epistemic and practical ones. Epistemic and practical arguments are evaluated by skeptical, i.e., grounded, semantics and credulous, i.e., preferred, semantics defined by Dungean semantics, respectively. Although these research handle practical reasoning in argumentation, they do not discuss the relationship to efficiency. We think that a decision of a course of action and the notion of efficiency are inseparable even when argumentation by single agent.

The welfare acceptability semantics is also applicable to argument-based negotiation requiring concession or compromise. Sawamura et al. [24] introduce seven dialectical inference rules on dialectical logic DL and weaker dialectical logic DM [25] to realize concession and compromise from inconsistent theory. Their argument-based negotiation mechanism equipped with the inference rules helps agents to reach agreement from conflict situations. Kido et al. [26] propose reasoning for compromise on a lattice, and illustrate that compromise arguments incorporating the reasoning realize compromise-based justification. Amgoud et al. [27] propose an abstract framework for argument-based negotiation, and introduce the notion of concession as an essential element of negotiation. Fan and Toni [28] develop assumption-based frameworks and propose a dialogue-based mechanism for conflict resolution. Their mechanism allows agents to make concession by incorporating others' knowledge, i.e., concession rules. However, none of them discuss the relationship between Pareto optimality with the notions of concession and compromise. We think that concessions and compromises should be chosen from Pareto optimal solutions.

6. Conclusions and Future Work

We introduced a Paretian argumentation framework as an important instance of Dung's abstract argumentation framework. We showed that preferred extensions of Paretian argumentation frameworks coincide with Pareto optimal arguments. We embodied the idea of justified Pareto optimal arguments and Pareto optimal extensions by combining abstract argumentation frameworks and Paretian argumentation frameworks.

One interesting future work will be applications of Paretian argumentation frameworks to dialogue games and formal dialogue systems. The attempt will allow agents to dialectically search for and argue about justified Pareto optimal solutions and Pareto optimal extensions.

Acknowledgements

This work was partially supported by Japan Science and Technology Agency, Core Research for Evolutionary Science and Technology.

References

[1] H. Kido, K. Nitta, Practical argumentation semantics for socially efficient defeasible consequence, in: Proc. of the 10th International Conference on Autonomous Agents and Multiagent Systems, 2011, pp. 267–274.

[2] H. Kido, Practical argumentation semantics for pareto optimality and its relationships with values, in: Proc. of the 8th International Workshop on Argumentation in Multiagent Systems, 2011, pp. 28–45.

[3] H. Kido, K. Nitta, Practical argumentation semantics for credulously justifying pareto optimal defeasible consequences, Transactions of the

Japanese Society for Artificial Intelligence (in Japanese) 27 (2) (2012) 52–60.

[4] P. M. Dung, On the acceptability of arguments and its funedamental role in nonmonotonic reasoning, logic programming, and n-person games, Artificial Intelligence 77 (1995) 321–357.

[5] T. J. M. Bench-Capon, P. E. Dunne, Argumentation in artificial intelligence, Artificial Intelligence 171 (10–15) (2007) 619–641.

[6] T. J. M. Bench-Capon, Persuasion in practical argument using value-based argumentation frameworks, Journal of Logic and Computation 13 (3) (2003) 429–448.

[7] L. Amgoud, C. Cayrol, Inferring from inconsistency in preference-based argumentation frameworks, Journal of Automated Reasoning 29 (2) (2002) 125–169.

[8] L. Amgoud, Repairing preference-based argumentation frameworks, in: Proc. of the 21st International Joint Conference on Artificial Intelligence, 2009, pp. 665–670.

[9] C. Cayrol, M. C. Lagasquie-Schiex, On the acceptability of arguments in bipolar argumentation frameworks, in: Proc. of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 2005, pp. 378–389.

[10] S. Modgil, M. Luck, Argumentation based resolution of conflicts between desires and normative goals, in: Proc. of the 5th International Workshop on Argumentation in Multi-Agent Systems, 2009, pp. 19–36.

[11] B. Verheij, Two approaches to dialectical argumentation: admissible sets and argumentation stages, in: Proc. of the 8th Dutch Conference on Artificial Intelligence, 1996, pp. 357–368.

[12] M. Caminada, Semi-stable semantics, in: Proc. of the 1st International Conference on Computational Models of Argument, 2006, pp. 121–130.

[13] P. M. Dung, P. Mancarella, F. Toni, A dialectic procedure for sceptical, assumption-based argumentation, in: Proc. of the 1st International Conference on Computational Models of Argument, 2006, pp. 145–156.

[14] P. Baroni, M. Giacomin, G. Guida, Scc-recursiveness: a general schema for argumentation semantics, Artificial Intelligence 168 (1-2) (2005) 162–210.

[15] S. Coste-Marquis, C. Devred, P. Marquis, Prudent semantics for argumentation frameworks, in: Proc. of the 17th International Conference on Tools with Artificial Intelligence, 2005, pp. 568–572.

[16] M. Caminada, Y. Wu, On the limitations of abstract argumentation, in: Proc. of the 23rd Benelux Conference on Artificial Intelligence, 2011, pp. 59–66.

[17] H. Prakken, G. Vreeswijk, Logics for Defeasible Argumentation, handbook of philosophical logic, 2nd Edition, Vol. 4, Springer, 2001, pp. 219–318.

[18] I. Rahwan, K. Larson, Pareto optimality in abstract argumentation, in: Proc. of the 23rd National Conference on Artificial Intelligence, 2008, pp. 150–155.

[19] A. Koutsoyiannis, Modern Microeconomics, Palgrave Macmillan, 1979.

[20] P. M. Dung, P. M. Thang, F. Toni, Towards argumentation-based contract negotiation, in: Proc. of the 2nd International Conference on Computational Models of Argument, 2008, pp. 134–146.

[21] D. Walton, Argumentation methods for artificial intelligence in law, Springer, 2005.

[22] T. J. M. Bench-Capon, H. Prakken, Justifying actions by accruing arguments, in: Proc. of the 1st International Conference on Computational Models of Argument, 2006, pp. 247–258.

[23] H. Prakken, Combining sceptical epistemic reasoning with credulous practical reasoning, in: Proc. of the 1st International Conference on Computational Models of Argument, 2006, pp. 311–322.

[24] H. Sawamura, M. Yamashita, Y. Umeda, Applying dialectic agents to argumentation in e-commerce, Electronic Commerce Research 3 (3-4) (2003) 297–313.

[25] R. Routley, R. K. Meyer, Dialectical logic, classical logic, and the consistency of the world, Studies in East European Thought 16 (1-2) (1976) 1–25.

[26] H. Kido, M. Kurihara, D. Katagami, K. Nitta, Formalizing dialectical reasoning for compromise-based justification, in: Proc. of the 3rd International Conference on Agents and Artificial Intelligence, 2011, pp. 355–363.

[27] L. Amgoud, Y. Dimopoulos, P. Moraitis, A general framework for argumentation-based negotiation, in: Proc. of the 4th International Workshop on Argumentation in Multi-Agent Systems, 2008, pp. 1–17.

[28] X. Fan, F. Toni, Argumentation dialogues for two-agent conflict resolution, in: Proc. of the 4th International Conference on Computational Models of Argument, 2012, pp. 249–260.