# Coalition Formability Semantics with Conflict-Eliminable Sets of Arguments

# (Extended Abstract)

Ryuta Arisaka
National Institute of Informatics, Tokyo, Japan
ryutaarisaka@gmail.com

Ken Satoh
National Institute of Informatics, Tokyo, Japan
ksatoh@nii.ac.jp

## ABSTRACT

Taking a model in political alliance among political parties, we will contemplate abstract-argumentation-theoretic characterisation of profitability, and then formability, of coalitions. We will deliver theoretical results around the semantics.

**1 Introduction** It is reasonable that we regard a coalition as a set of arguments its members express. Exploring abstract argumentation for an apt characterisation of coalition formability is worthwhile, thus. While a few papers [Amgoud 2005, Boella et al. 2008, Riley et al. 2012] in this domain apply Dung's acceptability semantics, we like to look into another direction for the sort of coalition formation we have in mind, which is one that may be found in a political-alliance-like coalition. It is unique by: being *more organisational than individual* (i.e. not possible to freely move agents across multiple political parties during the parliament term); involving *partial internal conflicts* (i.e. factions within a political party may argue against one another over details of certain political agendas); and exhibiting *asymmetry in attacks to and from a coalition* (i.e. while a political alliance must argue only by conflict-free portion of the arguments of the party's participants in order to retain credibility, the other political parties are unhindered by the personal circumstance of the alliance). As a single political party or alliance that dominates the parliament has total freedom in policy-making, a larger than smaller coalition is, all else equal, better for this type of coalitions. The LDP and its factions (the Japanese politics) highlight these points; see [Arisaka and Satoh 2016b].

In this work, we consider the following questions: (1) suppose a set of arguments that may contain partial internal conflicts (as an abstract representation of a political party) and suppose also rational criteria of profitability, with which other (disjoint) similar set(s) of arguments can it profit from forming a coalition?; and (2) suppose the profitability relation, some rational principles to judge the goodness of a coalition and such a set of arguments, which other (disjoint) similar sets of arguments can it actually form a coalition? We contemplate abstract-argumentation-theoretic characterisation of coalition profitability and formability semantics, and show theoretical results.

**2 Technical Details** Partial internal conflicts in a coalition necessitate a weaker condition, *conflict-eliminability*, than conflict-freeness, which we characterise with *argument capacity*, a positive integer representing the amount of information in an argument, given to each argument, and an *attack strength*, again a positive integer, given to each attack. An argument has an identity and a capacity by our design. A positive capacity means positive content, the greater the more. Attacks on an argument reduce its content. The idea is then: a set of arguments is conflict-eliminable just when no members of the set attack an argument of the same set with a greater attack strength than the argument's capacity. An attack from a set of arguments on an argument is defeating only when its strength is greater than the target's capacity. We define several properties for the attack which we define to be a partial function from a pair of a set of arguments (attacking arguments) and an argument (a targetted argument) to a numerical value (attack strength). Let us just assume a *coherent set of arguments* from this point on, which satisfies: that the set is finite; and that no two arguments share the same identity. For the first condition of the partial function, when there is an attack, there must be argument(s) that attack an argument **[Coherence]**. Secondly, there is an attack from a set of arguments on an argument just because each member of the set is attacking the argument **[Quasi-closure by subset relation]**. This can be contrasted with the group attacks in Nielsen-Parsons' formulation [Nielsen and Parsons 2006] which does not postulate a similar condition. It must be noted, however, that that there is an attack on an argument does not mean that the attack is defeating in our framework. Nielsen-Parsons' group attacks can be completely characterised in our framework, provided the number of arguments is finite(ary), by adjusting the numerical attack strengths. Thirdly, if the partial function is defined for two sets of arguments on the same argument, it is also defined for the set union of the two sets on that argument **[Closure by set union]**. This is the reverse of the previous condition. Fourthly, attack strength is positive **[Attack with positive strength]**. Fifthly, say an attack occurs from some set $S$ of arguments on an argument; now, increase the argument capacity of just one argument in $S$, keeping all else equal; then the attack which occured before the capacity increase should still occur **[Attack monotonicity 1]**. Sixthly, if an attack occurs from $S$ on $s$ with some strength, any superset of $S$ does not decrease the attack strength on $s$ **[Attack monotonicity 2]**. Seventhly, let us say that $S$ is attacking $s$. This intuitively means that $S$ intends to suppress $s$. Now, if the capacity of $s$ increases, $S$ still intends to suppress $s$ just as strongly or even more strongly, but not less strongly, for there are more materials in $s$ that $S$ could attack **[Attack monotonicity 3]**. Eighthly, $S$ may not attack any $s \in S$ **[No self attacks]**. Our argumentation framework comprises a coherent set of arguments $S$ and the partial function. $S_1 \subseteq S$ is conflict-eliminable just when no subset of $S_1$ defeats any $s \in S_1$.

The numerical values allow us also to extract *intrinsic arguments* of a conflict-eliminable set, which are the arguments that would re-

**Appears in:** *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

main if any partial internal conflicts within the set were resolved away. They are obtained by taking capacity of an argument minus the maximum attack strength incurred on $s$ by a subset of $S_1$, simultaneously for every $s \in S_1$. Now, replace $S_1$ with its intrinsic arguments, keeping all the other arguments in $S$ intact, and also remove from the domain of the partial function every $(S_2, s)$ satisfying $S_2 \subseteq S_1$ and $s \in S_1$. We call the pair of the updated arguments and the updated partial function $S_1$'s view of $S$. Clearly, there is no attack occurring in $S_1$'s intrinsic arguments in $S_1$'s view (of $S$). The notion of view captures the perspective of a coalition in a given argumentation framework. They characterise a coalition's - rather than its members' - arguments (see Introduction).

The asymmetry in attacks to and from a coalition can be rather easily formulated through these notions. As a coalition may attack external arguments only by its intrinsic arguments, we define that $S_1 \subseteq S$ c-attacks $s \in S$ iff $S_1$ is conflict-eliminable and there exists a subset of $S_1$'s intrinsic arguments that attack $s$ in $S_1$'s view. We define that $S_1 \subseteq S$ c-defeats $s \in S$ iff there is a subset of $S_1$'s intrinsic arguments that attack $s$ in $S_1$'s view with a greater attack strength than $s$'s capacity. We can now define c-admissibile and c-preferred sets, analogous to admissible and preferred sets in abstract argumentation but for conflict-eliminable sets. We say that $S_1 \subseteq S$ is c-admissible iff $S_1$ is conflict-eliminable, and there exists some subset of $S_1$'s intrinsic arguments for every $S_x$ *in $S_1$'s view that attacks* $s \in S_1$ such that it c-defeats some $s_x \in S_x$. In this definition it is $S_x$ *in $S_1$'s view* and not $S_x \subseteq S$; and $s \in S_1$ and not $s$ in $S_1$'s intrinsic arguments. This is because we are presuming conflict-eliminability (instead of conflict-freeness) of $S_1$. $S_x \subseteq S$ would include any partial conflict in $S_1$ as an attack; however, c-admissibility, which is the admissibility of a conflict-eliminable set in the view of the set, should not be defined to defeat it. On the other hand, $s \in S_1$ is due to the asymmetry in attacks to and from a conflict-eliminable set.

**3 Coalition Semantics and Results** Our semantics are relativised to conflict-eliminable sets. We can talk about whether a coalition is profitable *for* a conflict-eliminable set. We can talk about coalition formation of conflict-eliminable sets based on coalition profitability of both sets. A few notations shall be assumed for profitability semantics. We say that a conflict-eliminable set $S_1$ is one-directionally attacked iff there is a subset $S_x$ in $S_1$'s view such that $S_x$ attacks $s \in S_1$ and $S_1$ does not c-attack any $s_x \in S_x$. We say that $S_1$ is in a better state than, or as good a state as, $S_2$ just when (1) both $S_1$ and $S_2$ are conflict-eliminable, and (2) either: $S_2$ is c-admissible; $S_1$ is one-directionally attacked; or neither $S_1$ nor $S_2$ is c-admissible or one-directionally attacked. Informally, if $S_x \subseteq S$ is c-admissible, then it is fully defended from external attacks and is good. If $S_x$ is one-directionally attacked, then $S_x$ does not have any answer to external attacks, which is bad. Any conflict-eliminable set that does not belong to either of them is better than the second but worse than the first. Our definition of the state respects the quality. Finally, we say that coalition is permitted between $S_1$ and $S_2$ just when they do not overlap in $S$ and $S_1 \cup S_2$ is conflict-eliminable.

**Proposition** If coalition is permitted between $S_1$ and $S_2$, both $S_1$ and $S_2$ are necessarily conflict-eliminable.

We set forth 3 conditions for the relation that $S_1$ profits from forming a coalition into $S_2$, to be written $S_1 \trianglelefteq S_2$. Firstly, $S_1 \subseteq S_2$ (a larger set is better). Secondly, $S_2$ is not in a worse state than $S_1$ (a better state is better). Thirdly, let $\mathsf{Attacker}(S_1)$ be the arguments that attack any $s \in S_1$, then $S_2$ either c-defeats or contains at least as many number of arguments in $\mathsf{Attacker}(S_1)$ as $S_1$ either c-defeats or contains them (a set attacked by a smaller number of arguments is better). We will refer to the last condition by (fewer attackers) later. It is easy to see that there is some argumentation framework and some two subsets $S_1 \neq S_2$ of its arguments such that $S_1 \trianglelefteq S_2$. We state other results.

**Theorem (Existence)** If, for any $S_1$, there is some $S_2$ such that coalition is permitted between $S_1$ and $S_2$, $S_1 \trianglelefteq S_1 \cup S_2$, and $S_1 \cup S_2$ is c-admissible, then there is some $S_3$ such that coalition is permitted between $S_1$ and $S_3$, $S_1 \trianglelefteq S_1 \cup S_3$, $S_1 \cup S_3$ is c-preferred, and $S_2 \subseteq S_3$.

**Theorem (Mutual Maximality)** For any $S_1 \subseteq S_x$ such that $S_1$ is conflict-eliminable and that $S_x$ is c-preferred, both $S_1 \trianglelefteq S_x$ and $S_x \backslash S_1 \trianglelefteq S_x$.

In general, though, the mutual profitability is not guaranteed.

**Theorem (Asymmetry)** There is some argumentation framework where $S_1 \trianglelefteq S_1 \cup S_2$ but not $S_2 \trianglelefteq S_1 \cup S_2$.

Now, let $\mathsf{Max}(S_1)$ be the set of all $S_x$ such that $S_1 \trianglelefteq S_x$ and that there is no $S_x \subset S_y$ such that $S_1 \trianglelefteq S_y$, and let us say that profitability of a conflict-eliminable set $S_1$ is weakly continuous iff there is some $S_z \in \mathsf{Max}(S_1)$ such that, for any $S_w \subseteq S_z$, if coalition is permitted between $S_1$ and $S_w \backslash S_1$, then $S_1 \trianglelefteq S_w$. Let us also say that profitability of $S_1$ is continuous iff it is weakly continous for $S_1$ for any $S_z \in \mathsf{Max}(S_1)$. At first, we may expect that the continuation property holds good. However:

**Theorem (Profitability Discontinuation)** There are conflict eliminable sets $S_1, S_2, S_x$ such that $S_x \in \mathsf{Max}(S_1)$ and $S_2 \subseteq S_x$, but not $S_1 \trianglelefteq S_2$.

Profitability continuation holds in certain special cases, however.

**Theorem (Profitability Continuation)** Let $S_x$ be a c-preferred set. Then profitability is weakly continous for any $S_1 \subset S_x$ iff any disjoint pair $S_y, S_z$ of subsets of $S_x$ satisfy $S_y \trianglelefteq S_y \cup S_z$.

We use the profitability relation and the following rational utility postulates to express coalition formability semantics. (I) Coalition is good when it is profitable at least to one party. (II) Coalition is good when it is profitable to both parties. (III) Coalition is good when maximal future profits are expected from it. Of these, the first two can be understood immediately with the profitability relation. Our interpretation for the last postulate is as follows. Suppose a party, some conflict-eliminable set $S_1 \subseteq S$ in our context, considers coalition formation with another conflict-eliminable set $S_2$. We know that $S_2$ is some subset of $S_1 \subseteq S \backslash S_1$. Before $S_1$ forms a coalition with $S_2$, we have $\mathsf{Max}(S_1)$ as the set of maximal coalitions possible for $S_1$. Once the coalition is formed, we have $\mathsf{Max}(S_1 \cup S_2)$ as the set of maximal coalitions possible for the coalition. Here, clearly $\mathsf{Max}(S_1 \cup S_2) \subseteq \mathsf{Max}(S_1)$. What this means is that a particular choice of $S_2$ blocks any possibilities in $\mathsf{Max}(S_1) \backslash \mathsf{Max}(S_1 \cup S_2)$: they become unrealisable from $S_1 \cup S_2$. Hence $S_1$ has an incentive not to form a coalition with such a $S_2$ if all the members of $\mathsf{Max}(S_1 \cup S_2)$ are strictly and comparatively less profitable than some member of $\mathsf{Max}(S_1)$. We reflect this intuition on $\trianglelefteq$. We write $S_1 \trianglelefteq_{\mathsf{m}} S_2$ just when $S_1 \trianglelefteq S_2$, and some $S_x \in \mathsf{Max}(S_2)$ is such that, for all $S_y \in \mathsf{Max}(S_1)$, if $S_y$ is better (larger) than $S_x$ by either the set size, the state, or (fewer attackers), then $S_x$ is better than $S_y$ in one of the three criteria. We define four formability semantics: $\mathsf{W}$, $\mathsf{M}$, $\mathsf{WS}$ and $\mathsf{S}$ respecting (I), (II), (I + III) and (II + III).

$$\mathsf{W}(S_1) = \{S_2 \subseteq S \mid S_1 \trianglelefteq S_1 \cup S_2 \text{ or } S_2 \trianglelefteq S_1 \cup S_2\}.$$
$$\mathsf{M}(S_1) = \{S_2 \subseteq S \mid S_1 \trianglelefteq S_1 \cup S_2 \text{ and } S_2 \trianglelefteq S_1 \cup S_2\}.$$
$$\mathsf{WS}(S_1) = \{S_2 \subseteq S \mid S_1 \trianglelefteq_{\mathsf{m}} S_1 \cup S_2 \text{ or } S_2 \trianglelefteq_{\mathsf{m}} S_1 \cup S_2\}.$$
$$\mathsf{S}(S_1) = \{S_2 \subseteq S \mid S_1 \trianglelefteq_{\mathsf{m}} S_1 \cup S_2 \text{ and } S_2 \trianglelefteq_{\mathsf{m}} S_1 \cup S_2\}.$$

Intuitively, if $\rho(S_1)$ ($\rho \in \{\mathsf{W}, \mathsf{M}, \mathsf{WS}, \mathsf{S}\}$), $S_1$ is comfortable with forming a coalition with $S_2 \in \rho(S_1)$ under the given criteria.

**Theorem (Relation)** The following all hold good for a conflict-eliminable set $S_1$. (1 )$\mathsf{M}(S_1) \subseteq \mathsf{W}(S_1)$. (2) $\mathsf{WS}(S_1) \subseteq \mathsf{W}(S_1)$. (3) $\mathsf{S}(S_1) \subseteq \mathsf{M}(S_1)$. (4) $\mathsf{S}(S_1) \subseteq \mathsf{WS}(S_1)$. Meanwhile, neither $\mathsf{WS}(S_1) \subseteq \mathsf{M}(S_1)$ nor $\mathsf{M}(S_1) \subseteq \mathsf{WS}(S_1)$ is necessary.

## REFERENCES


Leila Amgoud. 2005. An argumentation-based model for reasoning about coalition structures. In *Argumentation in Multi-Agent Systems*. 217–228.

Leila Amgoud and Jonathan Ben-Naim. 2016. Axiomatic Foundations of Acceptability Semantics. In *KR*. 2–11.

Leila Amgoud and Philippe Besnard. 2009. Bridging the Gap between Abstract Argumentation Systems and Logic. In *SUM*. 12–27.

Ofer Arieli. 2012. Conflict-Tolerant Semantics for Argumentation Frameworks. In *JELIA*. 28–40.

Ryuta Arisaka and Ken Satoh. 2016a. Balancing Rationality and Utility in Logic-Based Argumentation with Classical Logic Sentences and Belief Contraction. In *PRIMA*. 168–180.

Ryuta Arisaka and Ken Satoh. 2016b. Coalition Formability Semantics with Conflict-Eliminable Sets of Arguments. *CoRR abs/1605.00495* (2016).

Ringo Baumann and Gerhard Brewka. 2015. AGM Meets Abstract Argumentation: Expansion and Revision for Dung Frameworks. In *IJCAI*. 2734–2740.

Floris Bex, Henry Prakken, Chris Reed, and Douglas Walton. 2003. Towards a Formal Account of Reasoning about Evidence: Argumentation Schemes and Generalisations. *Artificial Intelligence and Law* 11, 2 (2003), 125–165.

Guido Boella, Souhila Kaci, and Leendert van der Torre. 2009a. Dynamics in Argumentation with Single Extensions: Abstraction Principles and the Grounded Extension. In *ECSQARU*. 107–118.

Guido Boella, Souhila Kaci, and Leendert van der Torre. 2009b. Dynamics in Argumentation with Single Extensions: Attack Refinement and the Grounded Extension. In *AAMAS*. 1213–1214.

Guido Boella, Leendert van der Torre, and Serena Villata. 2008. Social Viewpoints for Arguing about Coalitions. In *PRIMA*. 66–77.

Nils Bulling, Jürgen Dix, and Carlos I. Chesñevar. 2008. Modelling Coalitions: ATL + Argumentation. In *AAMAS*. 681–688.

Martin Caminada and Leila Amgoud. 2005. An Axiomatic Account of Formal Argumentation. In *AAAI*. 608–613.

Martin Caminada and Leila Amgoud. 2007. On the evaluation of argumentation formalisms. *Artificial Intelligence* 171, 5-6 (2007), 286–310.

Claudette Cayrol, Florence Dupin de Saint-Cyr, and Marie-Christine Lagasquie-Schiex. 2010. Change in Abstract Argumentation Frameworks: Adding an Argument. *Journal of Artificial Intelligence Research* 38 (2010), 49–84.

Sylvie Coste-Marquis and Sébastien Konieczny. 2014. A Translation-Based Approach for Revision of Argumentation Frameworks. In *JELIA*. 397–411.

Sylvie Coste-Marquis, Sébastien Konieczny, Jean-Guy Mailly, and Pierre Marquis. 2014. On the Revision of Argumentation Systems: Minimal Change of arguments Statuses. In *KR*.

Phan M. Dung. 1995. On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming, and n-Person Games. *Artificial Intelligence* 77, 2 (1995), 321–357.

Phan M. Dung. 2009. Assumption-based argumentation. In *Argumentation in AI*. Springer, 25–44.

Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. 2011. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence* 175, 2 (2011), 457–486.

Dov M. Gabbay and Odinaldo Rodrigues. 2014. An equational approach to the merging of argumentation networks. *Journal of Logic and Computation* 24, 6 (2014), 1253–1277.

Nikos Gorogiannis and Anthony Hunter. 2011. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artificial Intelligence* 175, 9-10 (2011), 1479–1497.

João Leite and João Martins. 2011. Social Abstract Argumentation. In *IJCAI*. 2287–2292.

Hengfei Li, Nir Oren, and Timothy J. Norman. 2012. Probabilistic Argumentation Frameworks. In *TAFA*. 1–16.

Søren Holbech Nielsen and Simon Parsons. 2006. A generalization of Dung's Abstract Framework for Argumentation. In *ArgMAS*. 54–73.

Luís Moniz Pereira and Alexandre Miguel Pinto. 2007. Approved Models for Normal Logic Programs. In *LPAR*. 454–468.

Henry Prakken. 2004. Analysing reasoning about evidence with formal models of argumentation. *Law, Probability & Risk* 3, 1 (2004), 33–50.

Henry Prakken. 2010. An abstract framework for argumentation with structured arguments. *Argument and Computation* 1 (2010), 93–124.

Luke Riley, Katie Atkinson, and Terry Payne. 2012. A Dialogue Game for Coalition Structure Generation with Self-Interested Agents. In *COMMA*. 229–236.

Nicolás Rotstein, Martín O. Moguillansky, Alejandro J. García, and Guillermo R. Simari. 2008. An abstract Argumentation Framework for Handling Dynamics. In *Proceedings of the Argument, Dialogue and Decision Workshop in NMR 2008*. 131–139.