CIMI Research project

Ranking semantics for evaluating arguments

Coordinator: Jonathan Ben-Naim
Duration: 36 months
Partners: IRIT, CRIL, University of Balgrade

1 Summary of the project

Argumentation is a reasoning model based on the production and evaluation of interacting arguments. It is used for reasoning about incomplete, uncertain, inconsistent information, for making decisions under uncertainty, and for modelling different types of dialogues, namely negotiation. Whatever the application is, an argumentation framework consists of an *argumentation graph*, that is *arguments* and *attacks* between them, and a *semantics* for evaluating the arguments, i.e., for identifying the acceptable ones.

The dominant family of semantics in argumentation literature looks for sets of arguments, called *extensions*, each of which represents a coherent point of view. These semantics are founded on the strong hypothesis according to which the effect of an attack is binary, that is an attack may either kill its target or has no effect. Consequently, existing extension semantics satisfy axioms that are not suitable in certain applications of argumentation theory, namely decision making and negotiation.

In a recent paper [1], we investigated an alternative hypothesis which says that an attack cannot kill its target but *weakens* it. We proposed axioms that a semantics should satisfy, some of them are mandatory while others are optional, that is their choice depends on the application at hand. We also proposed two semantics that satisfy some of the axioms. In [2], we used the new semantics for defining a novel family of paraconsistent logics. It turns out that the latter are more discriminating than existing logics that are founded on extension semantics. This shows that the new semantics are very promising and overcome the limitations of the existing ones.

The main goal of the project is to further develop ranking semantics by investigating more axioms, more semantics, and by applying the new semantics to various theoretical applications, namely paraconsistent reasoning and decision making.

2 Background and present state of the art in the field

Argumentation is a form of common sense reasoning for drawing conclusions or making decisions by means of arguments. It is widely studied in Artificial Intelligence, namely for reasoning about inconsistent information (e.g., [3, 4, 5]), making decisions (e.g., [6, 7, 8]), and modelling agents interactions (e.g., [9, 10, 11]). Whatever the application is, an argumentation framework consists of an *argumentation graph*, that is *arguments* and *attacks* between them, and a *semantics* for evaluating the arguments. A semantics specifies which arguments are acceptable.

In the argumentation literature, arguments are mainly evaluated using *extension*-based semantics (or extension semantics for short) as introduced by Dung in his seminal paper [12]. Extension semantics are functions transforming any argumentation graph into one or several subsets of arguments, called *extensions*, each of which representing a coherent point of view. Using the extensions, the set of arguments is partitioned into three disjoint categories: i) the arguments which are in all the extensions, ii) the arguments that are in some but not all the extensions, and ii) the arguments which do not belong to any extension. Examples of extension semantics are the well-known stable and preferred semantics proposed by Dung in [12], as well as their refinements like the recursive semantics [13], semi-stable semantics [14], and ideal semantics [15]. These semantics evaluate arguments solely on the basis of the attack relation and do not take into account the internal structure of arguments. Their input is a plain directed graph whose nodes and arrows represent abstract arguments and attacks. Moreover, they are based on the following principles:

- The impact of an attack from an argument b to an argument a is binary. Indeed, if b is accepted the attack *kills* a, otherwise it has no effect on a.
- The quality of attacks prevail over their quantity. Indeed, one successful attack is sufficient to kill an argument, whereas a large number of weak attacks may fail to kill the argument.
- One successful attack against an argument a has the same effect on a as any number of successful attacks. Indeed, one such attack is sufficient to kill a, several attacks cannot kill a to a greater extent.
- The arguments that have the same status (i.e., arguments of the same category) are considered as equally acceptable.

These principles are certainly rational in applications like paraconsistent reasoning. As a matter of fact, in his paper [12], Dung has shown that his semantics allow to capture various well-established non-monotonic reasoning formalisms. However, the same principles may be debatable in other applications of argumentation theory. For example, the killing principle may be problematic in decision-making and in dialogues, because an attack does not necessarily kill its target but just weakens it. Think about a committee which recruits young researchers. Once an argument against a candidate is given, even if this argument is attacked, the initial argument is still considered by the members of the committee but with a lower strength. Similarly, the number of attackers may play a key role in decision making, especially when preferences are not available. To say it differently, when the quality is not sufficient to make decisions, one considers the quantity [16]. The last principle is also debatable since it gives the same importance to attacked and non-attacked arguments.

The above discussion reveals that there is no single way of evaluating arguments. The principles underlying semantics may be different from one application to another.

3 Objectives of the project - Workplan

Since early works on argumentation in Artificial Intelligence, arguments are evaluated by constructing extensions and by adopting the killing principle. In this project, we explore the *first* alternative approach for defining semantics. Our approach differs in two main aspects from extension semantics:

- Semantics are based on the weakening principle instead of killing.
- Arguments are not necessarily evaluated via extensions.
- A semantics takes as input an argumentation graph and returns a ranking on the set of arguments. Thus, it rank-orders the arguments from the most acceptable to the least acceptable ones.

The technical tasks are organised along three main lines of work, which reflect the overall organization of the project.

3.1 Ranking semantics (Year 1)

A first goal of the present proposal is to construct and axiomatically analyze semantics where the output takes the form of a *ranking* showing the relative acceptability of each argument (from the most acceptable to the least acceptable ones). More precisely, the four main points are the following:

- *Axioms.* A first important task consists in establishing axioms for ranking semantics. By axiom, we mean a principle that can be seen as desirable, be it general or specific, mandatory or optional, etc. The axioms represent criteria useful to better understand, judge, and compare different semantics.
- *Semantics.* Another important task is to construct semantics. For example, with attack graphs as input, a simple semantics consists in ranking the arguments solely on the basis of the number of attackers, i.e., the less an argument is attacked, the higher it is ranked.
- *Satisfaction.* A third task consists in showing that a semantics $\mathbf{S}$ satisfies a set of axioms. The more the axioms are numerous and desirable, the more $\mathbf{S}$ is justified from a theoretical point of view.
- *Characterization.* A more challenging task is to characterize a semantics $\mathbf{S}$ by a set of axioms $\mathbf{A}$, that is, prove that $\mathbf{S}$ is the only semantics that satisfies all the elements of $\mathbf{A}$. Such a characterization allows us to understand, judge, and compare $\mathbf{S}$ with other semantics with maximal precision. Indeed, the properties of $\mathbf{S}$ are exactly the consequences of the axioms of $\mathbf{A}$.

The four previous tasks will be carried out for four categories of argumentation frameworks:

- *Plain graphs*, that is argumentation graphs where all arguments are equally important, and all attacks are equally relevant.
- *Weighted graphs*, that is graphs where arguments are equipped with weights representing their intrinsic importance. Similarly, attacks are weighted where the weight of an attack represents its relevance.
- *Bipolar graphs*, that is argumentation graphs that contain two kinds of relations between arguments: *attacks* and *supports*.
- *Logic-based graphs*, that is argumentation graphs whose arguments have a logical structure.

3.2 Validating semantics with conflict measures (Year 2)

Another objective consists of evaluating the amount of disagreements in an argumentation graph. To put it differently, we propose *conflict measures* that calculate at what extent an argumentation graph is conflicting. Such measures are relevant for analysing the behaviour of a semantics. Indeed, they allow to check whether a given semantics treats equally graphs that have the same amount of conflicts. They also allow to compare the outputs of a semantics for respectively low and highly conflicting graphs.

There are three tasks that need to be carried out:

- To define the notion of conflict measure and propose a set of axioms that such a measure should satisfy.
- To propose intuitive candidate measures that satisfy the axioms or even are characterized by them.
- To analyse our ranking semantics with the proposed measures.

3.3 Paraconsistent logics based on argumentation (Year 3)

Argumentation is an alternative approach for handling inconsistency. Starting from a knowledge base encoded in a particular logical language, it builds arguments and attack relations between them using a consequence operator associated with the language, then it evaluates the arguments using an acceptability semantics. Finally, it draws the conclusions that are supported by the "winning" arguments.

Recently, it was shown in [1] that argumentation systems that are based on Tarskian logics [17] and Dung's acceptability semantics [12] follow the same line of research as well-known syntactic approaches for handling inconsistency, namely the one that computes the maximal (for set inclusion) consistent subbases (MCSs) of a knowledge base, and draws the conclusions that follow from all of them [18]. Indeed, it was shown that there are full correspondences between stable (respectively preferred, naive) extensions and MCSs. These formalisms are intuitive, but have a *sceptical behaviour* towards inconsistency. They only draw conclusions that follow from the formulae of the knowledge base that are not involved in inconsistency. They thus leave inconsistency *unsolved*. Let us consider the following illustrative example.

Example 1 *Let $\Phi = \{p, \neg p, q, p \to \neg q\}$ be a propositional knowledge base. This base has three maximal (for set inclusion) consistent subbases:*

- $\Phi_1 = \{p, q\}$,
- $\Phi_2 = \{p, p \to \neg q\}$,
- $\Phi_3 = \{\neg p, q, p \to \neg q\}$.

The common consequences of the three subbases are the tautologies. The same output is reached by any argumentation system that uses grounded, stable, semi-stable, preferred, and naive semantics.

Note that none of the two conflicts $\{p, \neg p\}$ and $\{p, q, p \to \neg q\}$ is solved. Such output may seem unsatisfactory in general and in multi agent systems where one needs an efficient way for solving conflicts between agents. Let us now have a closer look at the knowledge base Φ. The four formulae in Φ do not have the same responsibility for inconsistency. For instance, the degree of blame of p is higher than the one of q since it is involved in more conflicts. Moreover, p is frontally opposed while q is opposed in an indirect way. Similarly, $\neg q$ is more to blame than q since it follows from the controversial formula p.

In a recent paper [2], we proposed a novel family of argumentation-based logics which take advantage of such information in order to handle inconsistency. The new logics are built on top of Tarskian logics. Their main novelty lies in the semantics that are used for evaluating the arguments. Indeed, they use our ranking semantics proposed in [3]. We have shown that such logics are more discriminating than those based on Dung's semantics. Moreover, they satisfy several desirable properties, which shows that the approach is promising.

Our goal is to further develop such ranking logics.

- To define new ranking logics using new ranking semantics.
- To investigate the links between the axioms satisfied by a ranking semantics and those satisfied the resulting ranking logic.
- To compare argumentation-based ranking logics with model-based paraconsistent logics (e.g., multi-valued logics) and proof-based paraconsistent logics (e.g., Gentzen calculus).

4 Description of the consortium

Three partners are involved in this project: IRIT (Jonathan Ben-Naim), CRIL (Srdjan Vesic) and University of Belgrade (Dragan Doder).

Jonathan Ben-Naim (*Project coordinator*) is a CNRS researcher at IRIT. He received his Ph.D. in computer science from the university of Marseille in 2006. His main research domain consists of relation-based evaluation systems, i.e., systems for evaluating entities (agents, arguments, web pages, ...) on the basis of relations between them (vote, attacks, hypertext links,...).

Srdjan Vesic has a PhD is computer science. He is a CNRS researcher at CRIL - Lens. His research interests focus on reasoning under uncertainty and/or inconsistency, argumentation theory, computational social choice, decision making theory, logic-based knowledge representation and argumentation-based negotiation.

Dragan Doder has a PhD in mathematics. He is currently assistant professor at the university of Belgrade. His research interest focuses on applications of mathematical logic in artificial intelligence and computer science, and in particular reasoning about uncertainty, probability and inconsistency.

5 Added value of the proposed collaboration

In order to accomplish the project, several domains of expertise are needed: argumentation theory (Srdjan Vesic), evaluation systems (Jonathan Ben-Naim), and notions in mathematics and graph theory (Dragan Doder). We believe the only to way to reach the objectives of the project is to bring these particular expertise together.

6 Financial plan

The project duration is 3 years. We request 22000 euros for funding for the following points:

- 2 masters students: 6000 euros [2 * (500 euros * 6 months)]
- Conferences: 12000 euros (2 international conferences per year; 3 * 2 * 2000 euros)
- Research visits to IRIT for Dragan Doder and Srdjan Vesic: 2 * 2000 euros

References

[1] L. Amgoud, J. Ben-Naim, Ranking-based semantics for argumentation frameworks, in: Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM'13), 2013, pp. 134–147.

[2] L. Amgoud, J. Ben-Naim, Argumentation-based ranking logics, in: Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'15), 2015, pp. 1511–1519.

[3] L. Amgoud, P. Besnard, Logical limits of abstract argumentation frameworks, Journal of Applied Non-Classical Logics 23 (3) (2013) 229–267.

[4] A. García, G. Simari, Defeasible logic programming: an argumentative approach, Theory and Practice of Logic Programming 4 (1-2) (2004) 95–138.

[5] G. Governatori, M. Maher, G. Antoniou, D. Billington, Argumentation semantics for defeasible logic, Journal of Logic and Computation 14 (5) (2004) 675–702.

[6] L. Amgoud, H. Prade, Using arguments for making and explaining decisions, Artificial Intelligence Journal 173 (2009) 413–436.

[7] B. Bonet, H. Geffner, Arguing for decisions: A qualitative model of decision making, in: Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence (UAI'96), Morgan Kaufmann, 1996, pp. 98–105.

[8] C. Labreuche, Argumentation of the decision made by several aggregation operators based on weights, in: Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'06), 2006, pp. 683–690.

[9] E. Black, A. Hunter, A generative inquiry dialogue system, in: Proceedings of the 6th International Joint Conference on Autonomous Agents and Multi-Agents systems (AAMAS'07), 2007.

[10] N. Hadidi, Y. Dimopoulos, P. Moraitis, Argumentative alternating offers, in: Proceedings of the international conference on Autonomous Agents and Multi-Agent Systems (AAMAS'10), 2010, pp. 441–448.

[11] C. Reed, Dialogue frames in agent communication, in: Proceedings of the 3rd International Conference on Multi Agent Systems (ICMAS'98), 1998, pp. 246–253.

[12] P. M. Dung, On the Acceptability of Arguments and its Fundamental Role in Non-Monotonic Reasoning, Logic Programming and n-Person Games, AIJ 77 (1995) 321–357.

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

[14] M. Caminada, Semi-stable semantics, in: Proceedings of the 1st International Conference on Computational Models of Argument (COMMA'06), IOS Press, 2006, pp. 121–130.

[15] P. Dung, P. Mancarella, F. Toni, Computing ideal skeptical argumentation, Artificial Intelligence Journal 171 (2007) 642–674.

[16] D. Dubois, H. Fargier, J. Bonnefon, On the qualitative comparison of decisions having positive and negative features, Journal of Artificial Intelligence Research 32 (2008) 385–417.

[17] A. Tarski, Logic, Semantics, Metamathematics (E. H. Woodger, editor)), Oxford Uni. Press, 1956, Ch. On Some Fundamental Concepts of Metamathematics.

[18] N. Rescher, R. Manor, On inference from inconsistent premises, Journal of Theory and decision 1 (1970) 179–219.

[19] L. Amgoud, J. Ben-Naim, D. Doder, S. Vesic, Ranking arguments with parameterized semantics, in: Submitted, 2015.