Workshop on Argument Strength

Program and Abstracts

November 30 – December 2, 2016

Research Group
for Non-Monotonic Logic and
Formal Argumentation

Institute for Philosophy II

Ruhr University Bochum

RUB

Research Group for
Non-Monotonic Logic
and Formal Argumentation

Organization and Sponsors

The workshop is organized by the *Research Group for Non-Monotonic Logic and Formal Argumentation* at the *Ruhr University* of Bochum. The current members of the research group are:

- ☺ AnneMarie Borg
- ☺ Christian Straßer
- ☺ Dunja Šešelja
- ☺ Jesse Heyninck
- ☺ Mathieu Beirlaen
- ☺ Pere Pardo

Further information on the workshop can be found in the following link:

`http://homepage.ruhr-uni-bochum.de/defeasible-reasoning/Argument-Strength-2016.html`

For more information about the group and its members, we refer to our home page:

`http://homepages.ruhr-uni-bochum.de/defeasible-reasoning/index.html`

The research group is supported by A *Sofja Kovalevskaja* award, granted to Christian Straßer in 2014 by the *Alexander von Humboldt Foundation*, funded by the *German Ministry for Education and Research*.

Additional funding for this workshop was rewarded by the Ruhr University Research School PLUS (funded by Germanys Excellence Initiative [DFG GSC 98/3]).

Workshop Venue

The venue of the workshop is **Beckmanns Hof**, located on the university campus. (Warning: there is also a hotel in Bochum called Beckmannshof which is different from our Beckmanns Hof!)

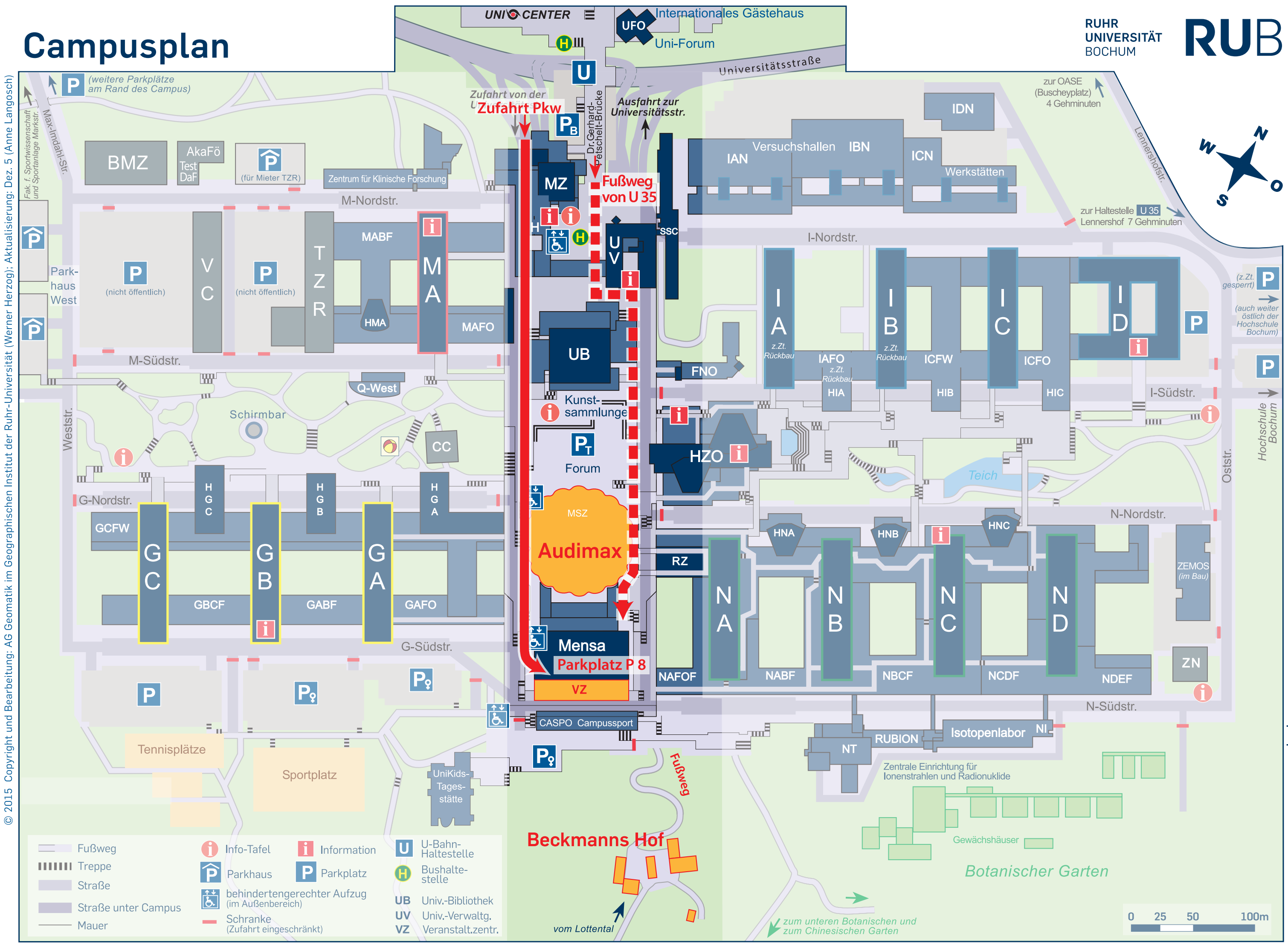
On Wednesday and Thursday, the workshop will be held in room ‘Shanghai’ on the ground floor. On Friday, we move to room ‘Tokio’ on the upper floor.

- ☺ **Getting from Bochum Hauptbahnhof to the Ruhr University of Bochum:** From Bochum Hauptbahnhof (central station) take the U35 towards Bochum Querenburg (Hustadt) and get out at stop “Ruhr Universität”. (Ticket needed: Preisstufe A, €2,60). On weekdays the subway U35 leaves every 5 minutes and reaches the university within 9 minutes.
- ☺ **Getting from the U-Bahn stop “Ruhr Universität” to Beckmanns Hof:** From the train station of the U35 (“Ruhr Universität”) go up the pedestrian bridge, turn right from the exit and walk towards the university. Your route takes you directly to the building of the university library. Keep walking till you pass the University library on your right, and continue until you reach the University “Mensa” building. Pass by the “Mensa” by walking down the stairs, cross a campus street and a parking lot, and further walk down to the Beckmanns Hof. (Around 10 min. walk in total.)

The next pages contain maps indicating how to walk from the RUB U-Bahn station (the blue square with the white ‘U’ at the top of the map) to Beckmanns Hof. The shortest route follows the dashed line.

Campusplan

RUHR UNIVERSITÄT BOCHUM RUB



www.rub.de


UNI CENTER
Internationales Gästehaus
UFO
Uni-Forum
Universitätsstraße
Zufahrt Pkw
Dr.Gerhard-Petschelt-Brücke
Ausfahrt zur Universitätsstr.
Versuchshallen
IAN
Klinische Forschung
MZ
Fußweg von U 35
SSC
I-Nordstr.
MA
IA
MAFO
UV
UB
FNO
IAFO
HIA
West
Kunstsammlungen
CC
Forum
HZO
HGA
MSZ
Audimax
HNA
GA
RZ
NA
NB
GAFO
G-Südstr.
Mensa
Parkplatz P 8
VZ
NAFOF
NABF
CASPO Campussport
NT
Fußweg
UniKids-Tagesstätte
Beckmanns Hof
U-Bahn-Haltestelle
Bushaltestelle
UB Univ.-Bibliothek
UV Univ.-Verwaltg.
VZ Veranstalt.zentr.
vom Lottental

Program

	Wednesday, November 30th 2016 – room 'Shanghai'
13:30 – 14:00	Registration and welcome
14:00 – 15:00	Henry Prakken *Argument strength and probability*
15:00 – 15:30	Adam Purtee and Lenhart Schubert *Simple rules for probabilistic commonsense reasoning*
15:30 – 16:00	Coffee Break
16:00 – 16:30	Elise Bonzon, Jérôme Delobelle, Sébastien Konieczny and Nicolas Maudet *Argumentation ranking semantics based on propagation*
16:30 – 17:00	Leila Amgoud, Jonathan Ben-Naim, Dragan Doder and Srdjan Vesic *On the notion of compensation between number and strength of attackers in ranking-based semantics*
17:00 – 18:00	Gabriele Kern-Isberner *A plausibility-based semantics for defeasible argumentation*

	Thursday, December 1st 2016 – room ‘Shanghai’
9:00 – 10:00	Leon van der Torre *On the acceptability of weighted arguments*
10:00 – 10:30	Sjur Dyrkolbotn, Truls Pedersen and Jan Broersen *On the possibility of rational indifference in abstract argumentation with structured arguments*
10:30 – 11:00	Coffee break
11:00 – 11:30	Chenwei Shi and Sonja Smets *Modeling the strength of arguments in dynamic epistemic probabilistic logic*
11:30 – 12:00	Paweł Łupkowski and Mariusz Urbański *The influence of an agent's credibility on the questioning agenda in a multi-agent environment*
12:00 – 12:30	Carlo Proietti *Group polarization and argument strength*
12:30 – 14:00	Lunch Break
14:00 – 14:30	Roberto Confalionieri, Oliver Kutz, Tarek R. Besold, Rafael Peñaloza and Daniele Porello *Argumentation, coherence, and blending*
14:30 – 15:00	Thomas F. Gordon *Defining argument weighing functions*
15:00 – 15:30	Robin E. Bloomfield and Kate Netkachova *Argument strength – an engineering perspective*
15:30 – 16:00	Coffee Break
16:00 – 16:30	James Freeman *Comparative argument strength: a formal inquiry*
16:30 – 17:30	Gerhard Brewka *Weighted abstract dialectical frameworks*
19:00 – . . .	Workshop dinner at Restaurant Karawane Location: Große Beckstraße 27, 44787 Bochum Directions: `http://restaurant-karawane.de/bochum/anfahrt.html`

	Friday, December 2nd 2016 – room 'Tokio'
9:00 – 10:00	Dov Gabbay *Distortions in argument strength generated in the reasoning of sex offenders*
10:00 – 10:30	Corina Andone, Florin Coman-Kund *The importance of argument strength in policy-making: the case of European Union directives*
10:30 – 11:00	Coffee break
11:00 – 11:30	Leila Amgoud and Jonathan Ben-Naim *Axiomatic foundations of argument evaluation*
11:30 – 12:00	Marcin Selinger *Argument evaluation based on proportionality*
12:00 – 12:30	Vlasta Sikimić *How specific arguments defeat general dogmas: lack of parsimony in molecular biology*
12:30 – 14:30	Lunch break
14:30 – 15:00	David Godden and Frank Zenker *A probabilistic analysis of argument cogency*
15:00 – 15:30	William Peden *The Ellsberg paradox and the weight of arguments*
15:30 – 16:00	Gabriella Pigozzi and Leon van der Torre *Arguing about constitutive and regulative norms*
16:00 – 16:30	Coffee break
16:30 – 17:30	Beishui Liao *Probabilistic argumentation and its computation – a survey*

Abstracts of Invited Talks

Weighted Abstract Dialectical Frameworks

1 Dec
16:30

Gerhard Brewka
University of Leipzig

ADFs (Abstract Dialectical Frameworks) are a generalization of Dung's argumentation frameworks which – rather than focusing entirely on attack – are able to express arbitrary relations among the nodes in an argument graph, like support, combined attack and so on. This is achieved by using explicit acceptance conditions, represented as propositional formulae, for the nodes. We briefly review ADFs and their semantics which are defined in terms of (pre)fixed points of a characteristic operator working on three-valued interpetations.

In an ADF, an argument is either accepted (t), rejected (f), or undecided (u). We discuss how the ADF approach can be generalized to allow for more fine-grained distinctions. We consider acceptance degrees taken form an arbitrary domain of values possessing an adequate truth ordering and an information ordering. We show how to accommodate such values using an adequate characteristic operator. We illustrate the approach using degrees in the unit interval.

Distortions in Argument Strength Generated in the Reasoning of Sex Offenders

2 Dec
09:00

Dov Gabbay
(joint work with Gadi Rozenberg and Lydia Rivlin)
King's College London
Ashkelon Academic College

The lecture[1] investigates the notion of strength of arguments x in the form that it cannot be 'out' unless attacked by at least $n > 1$ active ('in') attackers. This notion is widely used in practice (e.g. the requirement of at least two independent positive referee reports/witnesses/expert opinions/etc. Such a notion is orthogonal to numerical strength used by other authors. We call it the 'the many lives approach' option, where the argument x must be 'killed' n times to be really dead. Such notion arises in the logical analysis of argument distortion found in sex offenders who like to think they are not doing wrong. We develop semantics for such argumentation networks. This small many lives change forces us into the view that the proper semantics for an argumentation network (S,R, Extra labels) is not just a subset of S but another network of the same type. We also compare with literature and we also tell you about argumentation and sex offenders.

[1] A current version of the paper can be found at:
https://www.dropbox.com/s/cf4egt26j5fc6lq/Gabbay_Rozenberg_Rivlin_ArgStrengh2016.pdf

A Plausibility-based Semantics for Defeasible Argumentation

30 Nov
17:00

Gabriele Kern-Isberner
TU Dortmund
(Joint work with Guillermo Simari, Universidad Nacional del Sur, Bahia Blanca)

Defeasible argumentation and default reasoning are usually perceived as two similar, but distinct approaches to commonsense reasoning. This talk reports on work how to combine these two fields by viewing (defeasible resp. default) rules as a common crucial part in both areas. We will make use of possible worlds semantics from default reasoning to provide examples for arguments, and carry over the notion of plausibility to the argumentative framework. Moreover, we base a priority relation between arguments on the tolerance partitioning of Pearl's system Z and obtain a criterion phrased in system Z terms that ensures warrancy in defeasible argumentation.

Probabilistic Argumentation and its Computation – A Survey

2 Dec
16:30

Beishui Liao
Zhejiang University

In classical argumentation theory, the uncertainty of arguments and/or attacks is not considered. So, it could be regarded as a purely qualitative formalism. But, in the real world, arguments and/or attacks are often uncertain. So, in recent years, the importance of combining argumentation and uncertainty has been well recognized, and probability-based argumentation is gaining momentum. In this talk, I will review some basic notions of probabilistic argumentation, as well as various formalisms and computational approaches for probabilistic argumentation.

Argument Strength and Probability

30 Nov
14:00

Henry Prakken
Utrecht University

In this talk I will discuss to what extent argument strength can be characterised in terms of probability theory. First I will argue that a principled account of argument strength can only be given at the level of structured argumentation. Then I will make some general observations about the relation between argumentation and probability theory and discuss some existing work that connects the two. I will conclude with a brief overview of Sjoerd Timmers recent work on using argumentation for explaining Bayesian networks.

On the Acceptability of Weighted Arguments

1 Dec
09:00

Leon van der Torre
University of Luxembourg

I will first discuss argument strength based on priorities of rules, and reflect on recent results that weakest link corresponds to a greedy procedure, and last link to a construction proposed by Brewka and Eiter. Then I discuss the difference between argument strength and attack strength. Finally I discuss whether weighted argumentation should lead to rankings of arguments, a ranking of extensions, or something else.

Abstracts of Contributed Talks

Axiomatic Foundations of Acceptability Semantics

2 Dec 11:00

Leila Amgoud and Jonathan Ben-Naim
IRIT – CNRS, Toulouse

An *argument* gives reason to support a claim that is questionable, or open to doubt. It is made of three components: *premises* representing the reason, a *conclusion* which is the supported claim, and a *link* showing how the premises lead to the conclusion. The link is hence the logical "glue" that binds premises and conclusions together.

An argument has an *intrinsic strength* which may come from different sources: the certainty degree of its reason [3], the importance of the value it promotes if any [6], the reliability of its source [14], ... Whatever its intrinsic strength (strong or weak), an argument may be *attacked* by other arguments. An attack amounts to undermining one of the components of an argument, and has thus a negative impact on its target. An evaluation of the *overall strength* (or *overall acceptability*) of an argument becomes mandatory, namely for judging whether or not its conclusion is reliable.

The evaluation of arguments has received great interest from the computational argumentation community. Indeed, two families of acceptability semantics were defined for this purpose: *extension* semantics and *gradual* semantics. Extension semantics were initially introduced by Dung [3]. Starting with a set of arguments and attacks between them, they return a set of extensions, each of which is a set of arguments that are acceptable together. Then, using a membership criterion, a qualitative acceptability degree is assigned to each argument. Examples of such semantics are the classical semantics of Dung (complete, stable, preferred, ...) and their different refinements (e.g. [5, 9, 12, 10]). Unlike extension semantics, gradual semantics do not compute extensions. They assign a numerical acceptability degree to each argument. Examples of such semantics are h-Categorizer [4], Bbs, Dbs [12] and those proposed in [13, 8].

Despite the great interest in semantics, there are only a few works on *foundations* of semantics. Baroni and Giacomin [4] defined axioms that a semantics would satisfy. However, those axioms are only suited for extension semantics. Furthermore, they are mainly properties of extensions and not of overall strengths of arguments. Finally, most of the axioms are based on concepts (like defence and reinstatement) whose own foundations are unclear. Amgoud and Ben-Naim [12] proposed another set of axioms for the family of gradual semantics. The axioms are on the ranking of arguments with regard to their overall strengths. While some of the axioms (like independence and abstraction) are primitive, others are much more complex (like counter-transitivity) and their own foundations need to be clarified.

Hence, existing axiomatic studies do not tell much on the foundations of acceptability semantics. Foundations are important not only for a better understanding of the evaluation process, but also for comparing semantics and identifying families of semantics that have not been explored yet.

The aim of this paper is to set up the *foundations* of acceptability semantics. It defines elementary *concepts* and *principles* on which an evaluation of arguments is based. The approach followed in the paper is axiomatic. We introduce a set of axioms, each of which describes a concept or a principle. The axioms are primitive, in that they cannot be decomposed into other axioms. We investigate the properties of semantics that satisfy the axioms. We show in particular the foundations of defence and reinstatement, two key notions of extension semantics. Finally, we analyse existing semantics against the axioms, namely extension semantics proposed by Dung [3] and the gradual h-Categorizer semantics proposed by Besnard and Hunter [4]. This analysis allows not only a better understanding of the assumptions and choices made by those semantics, but also a clear comparison between semantics of the same family, and between extension semantics and gradual ones. A detailed version of this paper appeared in [2].

References

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

[2] Amgoud, L., and Ben-Naim, J. 2016. Axiomatic foundations of acceptability semantics. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016.*, 2–11.

[3] Amgoud, L., and Cayrol, C. 2002. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence* 34(1-3):197–215.

[4] Baroni, P., and Giacomin, M. 2007. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence* 171(10-15):675–700.

[5] Baroni, P.; Giacomin, M.; and Guida, G. 2005. Scc-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence* 168:162–210.

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

[7] Besnard, P., and Hunter, A. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence* 128(1-2):203–235.

[8] Bonzon, E.; Maudet, N.; and Moretti, S. 2014. Coalitional games for abstract argumentation. In *Proceedings of Computational Models of Argument, COMMA'14*, 161–172.

[9] Caminada, M. 2006. Semi-stable semantics. In *Proceedings of the 1st International Conference on Computational Models of Argument, COMMA'06*, 121–130.

[10] Dung, P.; Mancarella, P.; and Toni, F. 2007. Computing ideal skeptical argumentation. *Artificial Intelligence* 171:642–674.

[11] Dung, P. M. 1995. On the Acceptability of Arguments and its Fundamental Role in Non-Monotonic Reasoning, Logic Programming and n-Person Games. *Artificial Intelligence* 77:321–357.

[12] Grossi, D., and Modgil, S. 2015. On the graded acceptability of arguments. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence, IJCAI'15*, 868–874.

[13] Matt, P., and Toni, F. 2008. A game-theoretic measure of argument strength for abstract argumentation. In *Proceedings of the 11th European Conference on Logics in Artificial Intelligence, JELIA'08*, 285–297.

[14] Parsons, S.; Tang, Y.; Sklar, E.; McBurney, P.; and Cai, K. 2011. Argumentation-based reasoning in agents with varying degrees of trust. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems, AAMAS'11*, 879–886.

On the notion of compensation between number and strength of attackers in ranking-based semantics

30 Nov
16:30

Leila Amgoud[1] and Jonathan Ben-Naim[1] and Dragan Doder[2] and Srdjan Vesic[3]
[1] IRIT – CNRS, Toulouse
[2] Faculty of Mechanical Engineering, University of Belgrade
[3] CRIL, CNRS – Université d'Artois, Lens

An *argumentation framework* consists of an *argumentation graph*, that is *arguments* and *attacks* between them, and a *semantics* for evaluating each argument, and thus for specifying at what extent the argument is acceptable.

The most dominant semantics in the literature are *extension semantics* that were initially proposed by [3] and further developed by other scholars. Such semantics compute extensions of arguments, each of which is a set of jointly acceptable arguments. Then, using a membership criterion, they assign a qualitative acceptability degree to each argument. An argument is either *skeptically accepted* if it is in all the extensions, or *credulously accepted* if it belongs to some but not all extensions, or *rejected* if it does not belong to any extension.

Such semantics are all based on at least two assumptions: First, a successful attack completely destroys its target. Indeed, an argument that is attacked by a skeptically accepted argument is necessarily rejected. There are applications where this assumption makes perfect sense. However, in applications like decision making or dialogues, an attack only weakens its target. 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). The second assumption underlying extension semantics is that the number of attackers is useless. Consider the following two argumentation graphs:

$\mathcal{F}_1$ $\mathcal{F}_2$

The argumentation graph $\mathcal{F}_1$ has two preferred/stable extensions $\{a\}$ and $\{b\}$, thus both a and b are credulously accepted. Similarly, $\mathcal{F}_2$ has two preferred/stable extensions $\{a, c\}$ and $\{b\}$. Thus, b is credulously accepted even if it has an additional attacker (c). This shows that the attacker c did not further weaken the argument b.

[1] introduced the so-called *ranking semantics*. Those semantics do not follow the two previous assumptions. They rather assume that if the attackers of two arguments x and y have the same strength, and x has more attackers than y, it is expected that y is stronger than x. Also, if x and y have the same number of attackers, but the attackers of x are stronger than attackers of y, argument y should be stronger than x. But what happens if x has a small amount of strong attackers, and y has a large amount of weak attackers? How to decide which of the two arguments should be stronger?

This work aims at defining a ranking-based semantics where several weak attacks may compensate one strong attack. But how many weak attacks do we need in order to compensate for one strong attack? Our work is driven by the principle that such a decision should not be taken in an arbitrary manner. This decision should not be hard-coded in the definition of the semantics. Instead, we show that it is possible to define a class of semantics based on parameter α that allows the user to choose to which extent to take into account the number, and to which extent the strength of attackers. We define a broad class of ranking semantics, called α-BBS, which allow for compensation. α-BBS assign a burden number to each argument and order the arguments with respect to those numbers.

We study formal properties of α-BBS. First, we show that they satisfy all the mandatory axioms for ranking-based semantics from the literature. We also define a new axiom to capture the notion of compensation and show that it is satisfied by our semantics. We then propose an algorithm that calculate the arguments' scores in an efficient manner and prove its correctness. We also perform experiments using ICCMA benchmark and show that the approach computes the ranking very quickly (less than 0.5 seconds per argumentation graph in average). Moreover, an approximation of the ranking can be provided at any time.

We study the conditions such a semantics should satisfy. Namely, it should be intuitive, it should satisfy the existing principles and the new principle linked to compensation, it should be easy to compute etc. We discuss the alternative definitions of ranking-based semantics that satisfy compensation with respect to those conditions.

Acknowledgement

This is the short version of the paper; the full version of the paper [2] was presented and published at KR 2016.

Leila Amgoud and Srdjan Vesic benefited from the support of the project AMANDE ANR-13-BS02-0004 of the French National Research Agency (ANR). Dragan Doder was supported by the National Research Fund (FNR) of Luxembourg through project PRIMAT.

References

[1] L. Amgoud and J. Ben-Naim. Ranking-Based Semantics for Argumentation Frameworks. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM'13)*, 2013, 134–147.

[2] L. Amgoud and J. Ben-Naim and D. Doder and S. Vesic. Ranking Arguments with Compensation-Based Semantics. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR'16)*, 2016, 12–21.

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

The Importance of Argument Strength in Policy-Making: The Case of European Union Directives

2 Dec 10:00

Corina Andone[1] and Florin Coman-Kund[2]

[1]University of Amsterdam
[2]Erasmus University of Rotterdam

This paper discusses argument strength in the context of policy-making by focusing on the arguments advanced by the European Union (EU) legislator in the preamble of directives adopted for harmonization in the internal market. The paper also aims at assessing whether the arguments advanced by the legislator are strong enough to convince the Member States to implement the directives at issue.

In the first part of the paper, we explain that arguments are considered strong in this context if they cover at least three aspects as imposed by the fundamental legal principles governing EU decision-making: (a) the legal basis on which the decision-making competence rests, (b) the principle of subsidiarity, defining the circumstances in which it is preferable for action to be taken by the Union, rather than the Member States, and (c) the principle of proportionality, specifying that the content and form of the proposed action does not exceed what is necessary to achieve the objectives set by the Treaties.

In the second part of the paper, we examine closely the way in which the EU legislator actually argues in practice and how this affects the strength of its arguments. By focussing in particular on Directive 2011/83/EU on consumer rights, we reveal a legislative practice in which the obligation to give reasons is not discharged adequately when judged from an argumentation theory perspective. We outline a pattern in which the burden of proof is evaded: the obligation to give reasons is not discharged adequately either by failing to provide arguments at all or by giving insufficient arguments. In addition, harmonization is claimed to be the best and only solution to an existing problem in the internal market, whereas other solutions that might also be available are given little if any consideration.

Finally, this paper connects theory to practice by discussing critically the approach of the Court of Justice of the European Union (CJEU) in evaluating the strength of the arguments in EU directives. As the competent authority to carry out such an evaluation, CJEU assesses the preambles of EU directives in the context of the 'check' on the duty to state reasons under Article 296 of the Treaty for the Functioning of the European Union (TFEU). Case law in this respect shows that the problematic legislative practice remains generally unsanctioned due to the rather light and flexible test used by CJEU under Article 296 TFEU.

Argument Strength – An Engineering Perspective

1 Dec 15:00

Robin E. Bloomfield[1,2] and Kate Netkachova[1,2]

[1]Adelard LLP
[2]City University London

Background

We are researchers and practitioners in the safety and security of complex, often computer based, systems. We are concerned with the evaluation and communication of risk and whether systems are trusted and trustworthy. We have developed an approach based on Claims, Argument and Evidence (CAE) and have an industrial tool (ASCE), methodology and research to support it.

Over the past decade there has been a trend towards an explicit claim-based approach to safety justification and considerable work has been done on the structuring of safety arguments [1, 2, 3, 4] and supporting standards [5]. A recent rigorous analysis of assurance cases is provided in [6]. The approach we use is based on the simple concepts of:

- Claims statements about a property of the system
- Evidence that is used as the basis of the justification of the claim
- Arguments link the evidence to the claim

In order to support the use of CAE, a graphical notation ASCAD [1] is used to describe the interrelationship of the claims, argument and evidence. In addition there are important narratives and analyses explaining and detailing the claims and arguments being made: narrative is an essential part of the case and tools such as the Adelard ASCE tool [8] explicitly support it.

Engineering justifications are complex and cannot be expressed in terms of a simple CAE triple. If we are developing a top down justification, the claims need to be expanded into subclaims until we can identify evidence that can directly support the subclaims. For example, claims such as "the system is adequately secure" are too vague and it's therefore necessary to create subclaims until the final nodes of the assessment can be directly supported or refuted by evidence. Engineering assurance arguments tend to be some 10s to 100s of nodes and have considerable supporting narrative.

We have developed an approach to structuring such arguments based on a set of archetypal CAE fragments what we have termed CAE building blocks [9]. The identification of the blocks was supported by an empirical analysis of the types of engineering arguments that are made about safety and dependability from defence, finance and medical applications. There are five basic CAE building blocks:

- Decomposition partitions some aspect of the claim
- Substitution refines a claim about an object or property into another claim about an equivalent object or property
- Concretion gives a more precise definition or specific value to some aspect of the claim
- Calculation or proof used when some value of the claim can be computed or proved
- Evidence incorporation incorporates evidence that directly supports the claim

The block structure contains enhancements to the classical CAE approach in how arguments are addressed. We identify specific aspects of the argument that allow us to logically, deductively, show that the claim is valid and call these aspects "side-warrants". Side-warrants tell us why we can deduce the top-level claim from the subclaims, and under what circumstances the argument is valid. The side-warrant is in fact a type of claim and we may wish to challenge and demonstrate it for the specific case by introducing further subclaims (e.g. a negation of a defeater) or by justifying the side-warrant directly providing narrative argument and evidence.

The basic scheme of a generic CAE block structure is shown in Figure 1. It shows n subclaims supporting an argument that justifies a top-level claim, with some of the key properties expressed as the side-warrant and supported by the system information and external backing.

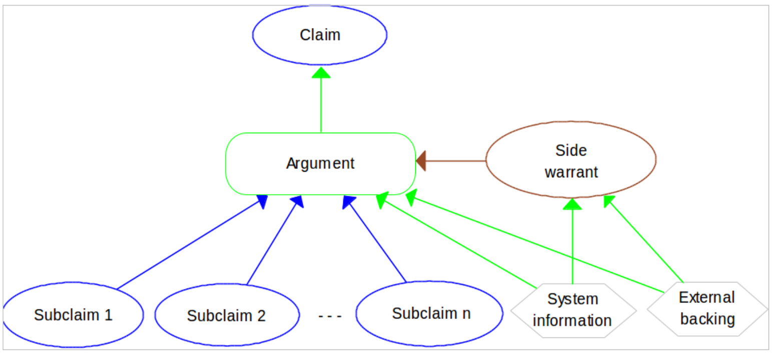


Figure 1. General CAE Block Structure.

Strength of argument

In assessing a justification for a system there are a number of sources of doubt to address including the:

1. Formulation and validity of the top-level claims,
2. Trustworthiness and relevance of the evidence provided to support the claims, One of the CAE Blocks deals explicitly with evidence incorporation and we would expect a close link between the evidence and claim (e.g. on strength of evidence we have experimented with the use of confirmation theory),
3. Individual argumentation fragments, based around a claim-argument-subclaim structure that has been developed from applying a CAE Block, and
4. The aggregation of the subclaims and intermediate arguments, the individual CAE Blocks, to justify adequate confidence in the top-level claim.

One innovation in the work reported here is to use CAE blocks to address (3). These provide a framework to challenge the arguments and recursively separate out the inductive and deductive parts of them. As we identify defeaters to the argument we may wish to add them as negated subclaims or be part of the supporting narrative.

We can systematically elaborate sources of doubt arising from the application of the approach, for example:

- Problems with the blocks instantiation
- Incorrect deductions
- Mismatch with respect to the real world

If we are to have a calculus of confidence that deals with argument strength, we need to classify different types of doubt and we have been experimenting with classification approaches.

The proposed talk

The motivation for this talk and attendance at the workshop is to see how the justification of critical engineered systems that impact all out lives might be informed by the work of the philosophical argumentation community. The talk would provide the context of the justifications we work on and how our research is trying to address issues of confidence and argument strength by providing a deductive/framework and by classifying the sources and types of doubt that undermine the strength of the argument.

References

[1] R. E. Bloomfield, P. G. Bishop, and C. C. M. Jones, P. K. D. Froome, *ASCAD Adelard safety case development manual* (London, 1998).

[2] P. G. Bishop, R. E. Bloomfield, The SHIP safety case - a combination of system and software methods, In *SRSS95, Proc. 14th IFAC Conf. on Safety and Reliability of Software-based Systems*, Bruges, Belgium, 12-15 September 1995.

[3] P. G. Bishop, R. E. Bloomfield, A methodology for safety case development, in F. Redmill, T. Anderson (eds.) *Industrial Perspectives of Safety-critical Systems: Proc. 6th Safety-Critical Systems Symposium*, 194–203, Springer, 1998.

[4] T. Kelly, The goal structuring notation – a safety argument notation, in *Proc. DSN 2004 Workshop on Assurance Cases*, 2004.

[5] ISO/IEC 15026-2:2011, *Systems and software engineering – Systems and software assurance, Part 2: Assurance case*, 2011.

[6] Rushby, J. *The Interpretation and Evaluation of Assurance Cases*, Technical Report SRI-CSL-15-01, July 2105.

[7] S.E. Toulmin, *The Uses of Argument* (Cambridge University Press, 1958.

[8] Adelard LLP, *Assurance and Safety Case Environment (ASCE)*, `http://www.adelard.com/web/hnav/ASCE/index.html`.

[9] Bloomfield, R.E., Netkachova, K: Building Blocks for Assurance Cases. In *IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*, 2014, 186 – 191.

Argumentation Ranking Semantics based on Propagation

30 Nov
16:00

Elise Bonzon[1] and Jérôme Delobelle[2] and Sébastien Konieczny[2] and Nicolas Maudet[3]

[1]LIPADE, Université Paris Descartes
[2]CRIL, CNRS – Université dArtois, Lens
[3]LIP6, CNRS –Université Pierre et Marie Curie, Paris

Argumentation is a very natural framework for representing conflicting information. A proof of its appeal is the recent development of online platforms where people participate in debates using argumentation graphs (e.g. *debategraph.org* or *arguman.org*).

The question now goes towards the reasoning part: how to automatically use these argumentation graphs that are constructed this way? Argumentation has been a very active topic in Artificial Intelligence since more than two decades now, and Dung's work on abstract argumentation framework [1] can be used to represent the graphs (even if some additional information should be also represented, like the number of people that agree/disagree with an argument and/or an attack, or a support between arguments, ...). But the main issue is about the semantics that one should use in this case. In fact classical Dung's semantics, using extensions [1] (or equivalently labellings [2]), with their dichotomous evaluation of arguments (accepted/rejected) do not seem very well suited for such applications. As discussed in [3], on such online platforms, with a big number of arguments, and a lot of individuals participating, it can be problematic (in particular quite unintuitive for the participants) to have such a drastic evaluation, that is not that informative (since there are only two levels of acceptability), or to propose several possible results (different extensions). So in this case a finer evaluation of arguments seems to be more adequate. The idea is then to have ranking-based semantics, that allow to produce a full ranking of the arguments, from the most to the least acceptable ones. This kind of semantics seems very natural, and it is then quite surprising that they have received little attention until recently [3, 4, 5, 6, 7, 12, 9, 10]. These semantics basically rely on the attacks and defenses of each argument in order to evaluate its acceptability rank.

In this work we propose a new family of semantics, that relies on attacks and defenses, like previous semantics, but that also puts a strong emphasis on non-attacked arguments. While many principles remain discussed and controversial, all semantics agree on the fact that non-attacked arguments should have the highest rank. The idea of our semantics is that these arguments should also have a greater impact on the evaluation of the other ones.

Following this idea, we propose six new semantics based on the principle of propagation defined in two steps. The first step consists in assigning a positive initial weight to each argument. The score of 1 attached to non-attacked arguments is set to be higher than the score of attacked arguments, which is an ϵ between 0 and 1. The value of this ϵ is chosen accordingly to the degree of influence of the non-attacked arguments that we want: the smaller the value of ϵ is, the more important the influence of non-attacked arguments on the order prevails. Then, during the second step, we propagate the weights into the graph in changing their polarities in order to comply with the attack relation meaning (attack or defense). For each argument, we accumulate and store the weights from its attackers and defenders in the argumentation framework. The difference between these semantics lies in the method that is chosen to differentiate non-attacked arguments and attacked ones, and in the choice of considering one or all paths between arguments.

We show that these semantics have interesting properties. In particular they satisfy the properties that should be satisfied by any ranking semantics according to [11]. Some relationships between these semantics and other ones exist : all the propagation semantics based that consider all paths between arguments coincide with the *Discussion based semantics* [12] when there is no non-attacked arguments. So they can be viewed as improvement of *Discussion based semantics* allowing to take into account the impact of non-attacked arguments. Also, by many respect one of our semantics is close to the $Tuples^*$ semantics [5]. The $Tuples^*$ semantics does not necessarily provide a total pre-order, and it cannot be applied (easily) if there is a cycle in the argumentation framework. So in a sense our semantic could be seen as an improvement of the ideas of $Tuples^*$ that allows to overcome these limitations.

Acknowledgement

This work benefited from the support of the project AMANDE ANR-13-BS02-0004 of the French National Research Agency (ANR).

References

[1] 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.

[2] Caminada, M. 2006. On the issue of reinstatement in argumentation. *Proc. of the 10th European Conference on Logics in Artificial Intelligence (JELIA'06)*: 111–123.

[3] Leite, J. and Martins, J. 2011. Social abstract argumentation. *Proc. of the 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*: 2287–2292.

[4] Besnard, P. and Hunter, A. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence* 128 (1–2):203–235.

[5] Cayrol, C. and Lagasquie-Schiex, M. 2005. Graduality in argumentation. *Journal of Artificial Intelligence Research* 23:245–297.

[6] Matt, P. and Toni, F. 2008. A game-theoretic measure of argument strength for abstract argumentation. *Proc. of the 11th European Conference on Logics in Artificial Intelligence (JELIA'08)*: 285–297.

[7] da Costa Pereira, C. and Tettamanzi, A. and Villata, S. 2011. Changing one's mind: erase or rewind?. *Proc. of the 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*: 164–171.

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

[9] Pu, F. and Luo, J. and Zhang, Y. and Luo, G. 2014. Argument ranking with categoriser function. *Proc. of the 7th International Conference on Knowledge, Science, Engineering and Management (KSEM'14)*: 290–301.

[10] Grossi, D. and Modgil, S. 2015. On the graded acceptability of arguments. *Proc. of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI'15)*: 868–874.

[11] Bonzon, E. and Delobelle, J. and Konieczny, S. and Maudet, N. 2016. A Comparative Study of Ranking-based Semantics for Abstract Argumentation. *Proc. of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*: 914–920.

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

On the possibility of rational indifference in abstract argumentation with structured arguments

1 Dec
10:00

Sjur Dyrkolbotn[1] and Truls Pedersen[2] and Jan Broersen[1]

[1] Utrecht University

[2] University of Bergen

In recent work, it has been noted that the so-called rationality postulates for argumentation are not necessarily satisfied when the weakest-link principle is used to lift priorities over rules to orderings over arguments.[2] In response, a *disjoint and strict* lifting principle has been proposed, forcing the resulting orderings over the arguments to be *anti-symmetric* [1].

We argue that this restricts us too much, also when the ordering is meant to encode a notion of argument strength. It seems perfectly possible for two distinct arguments to be equally strong, so that one may be forced to make a non-deterministic choice between them. Based on this conceptual starting point, we set out to find lifting principles that do not enforce anti-symmetry, yet still suffice to ensure that the rationality postulates are satisfied.[3]

[2]Here we take priorities to be preorders over rules. That is, a priority ordering is always transitive and reflexive.

[3]The rationality postulates were first introduced in [2].

First, we show that *connectedness* of the underlying priority ordering will suffice: if any two rules are related under the ordering, traditional weakest-link lifting will satisfy rationality. This is because connectedness (unlike anti-symmetry) is *preserved* by weakest-link lifting of a preorder: if the priority ordering is connected, the same is true for the lifted argument ordering. However, we go on to note that connectedness also seems like an excessive requirement. Why should the agnostic position be ruled out *a priori*?

This leads us to consider the following higher-level question: what feature of structured argumentation (in the style of $ASPIC^+$) is responsible for causing trouble when we work with argument orderings that are neither connected nor anti-symmetric? The answer turns out to be the following: the standard heuristic for inducing a defeat relation from a collection of structured arguments treats incomparable arguments *the same way* as it treats arguments that are symmetrically related [3, 4].[4] In both cases, rebutting attacks will succeed as defeats, in either direction.

We argue that this feature is reasonable, at least in many contexts. However, it forces us to look more carefully at the properties of the *complement* of the symmetric closure of our ordering relation. We go on to show that rationality at the argumentation level can be regained by simply imposing transitivity for this complementary relation. That is, for all A, B and C, if A, B and B, C are both unrelated, then we require that A and C must be unrelated as well. We argue that this is reasonable, especially in view of the standard approach to inducing defeat relations in $ASPIC^+$-style structured argumentation. Effectively, our paper shows that despite recent doubts, rational indifference is possible, provided we require basic reasonableness (transitivity) regardless of whether the form of indifference we encounter is that of incomparability or that of symmetry.

References

[1] Young, A. and Modgil, S. and Rodrigues, O. 2016. Prioritised default logic as rational argumentation. In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems*: 626–634.

[2] Caminada, M. and Amgoud, L. 2007. On the evaluation of argumentation formalisms. *Artificial Intelligence* 171(5–6): 286–310.

[3] Prakken, H. 2010. An abstract framework for argumentation with structured arguments. *Argument & Computation* 1(2): 93–124.

[4] Modgil, S. and Prakken, H. 2013. A general account of argumentation with preferences. *Artificial Intelligence* 195: 361–397.

---

[4] In general, there still *is* a difference between symmetry and incomparability (compositionally they behave differently, for instance). Hence, simply equating the two notions by introducing a symmetric link between any two incomparable arguments is not an appropriate solution.

Comparative Argument Strength: A Formal Inquiry

1 Dec
16:00

James Freeman

Hunter College; City University of New York

We explore whether and to what extent we may give a formal account of the comparative strength of arguments through L.J. Cohen's account of relevant variables, the Bayesian notion of a priori probability as explicated through plausibility, and Toulmin's notion of the layout of arguments. In Toulmin's model, an argument's strength is intimately connected to its warrant. For clarity, we distinguish a warrant as an inference licence from its associated generalization. As Toulmin indicates, warrants raise the question of their "authority and currency," i.e. the question of their backing. At least for warrants backed "from below," ultimately by observation of instances, the strength of a warrant is a function of the weight of its backing determined in this way:

Suppose one observes a correlation between properties P and Q. This observation reflects a default situation where there is no experimental manipulation of other relevant variables which might affect whether what satisfies P also satisfies Q. On Cohen's account, one may at least ideally determine argument strength through a canonical test. One takes account of the relevant variables. One then tests whether Q's holding is affected by whether values of any of the relevant variables hold. Say there are n relevant variables. In level 1 of the test, only values of the first variable are considered to see whether any pose counterexamples to the associated generalization of the warrant. If none, the warrant and corresponding generalization receive the first level of support, i.e. to degree $1/n$. At level 2, the generalization is tested to see whether any values of variable 2, singly or in combination with a value of variable 1, constitute a counterexample. At level 3. values of variable 3, singly or in combination with one or two values of the first two variables are considered. The test continues until reaching one of two outcomes: At some level $i+1$, where no counterexample has appeared in tests through level i, a counterexample does appear. The generalization then has support to degree i/n. By contrast, if no counterexample appears through level n, the generalization is supported to degree n/n and may be regarded as a law of nature.

There is a problem here. For some variables, the probability may be high that some of the variants of the variables will constitute counterexamples to the generalization, For other variables, the probability may be less, even low. Cohen calls this the "accepted falsificatory efficacy of [the] variants" of some variable. Suppose the falsificatory efficacy of values of variable i is high, where for variable j it is low. Then if i appears before j in the ordering of relevant variables, the generalization may pass fewer levels of the test than if j precedes i. The strength of the argument would be different, depending on the order of the relevant variables. Hence we need to order the variables according to (expected) falsificatory efficacy.

Intuitively, once a set of relevant variables is identified, this ordering requires an estimation of the a priori probability that the values of a given variable will include counterexamples to a given generalization. Answering this question requires considerations of *plausibility.* Salmon has argued for understanding a priori probability through plausibility considerations. Rescher has developed a formal theory of plausibility, and J.R. Welch has developed a theory of plausibilistic coherence. Can a plausibility account of a priori probability allow us to construct a method of assessing the falsificatory efficacy of a relevant variable? If so, we may give in turn an account of comparative argument strength. A canonical test is an idealization. In many cases, carrying the test through to level $i \leq n$ may not be practically feasible. But if a warrant has some degree of support, an argument instancing that warrant has some positive strength. Indeed, in the absence of recognized counterexamples, the strength may be sufficient to justify accepting the conclusion on the basis of the premises. Variables which might produce a counterexample have a low prior probability of doing so. Hence our inquiry has importance for one of the central questions in argument evaluation: Are the premises adequately connected to the conclusion?

A Probabilistic Analysis of Argument Cogency

2 Dec
14:30

David Godden[1] and Frank Zenker[2,3,4]
[1]Michigan State University
[2]Lund University
[3]Slovak Academy of Sciences
[4]Konstanz University

As John Woods [1, p. 15] put it: "Formal logic is a theory of logical forms; and informal logic is all the rest." Informal logicians tend to view "all the rest" as shouldering the real work in the analysis and evaluation of natural language argumentation. Indeed, many reject formal methods. In place of the proof techniques of the truth-functional calculus, typical resources rather include argument diagrams, schemes, and the fallacies. Similarly, rather than endorsing soundness (premise truth and deductive validity) as a standard of good argument, informal logicians speak of cogency (premise acceptability, relevance, inferential sufficiency, also known as the "RSA conditions").

When this anti-formalist stance arose during the 1960s, the only widely available formal apparatus had been deductive logic. The breath of formal resources available today, however, makes a continued disenchantment with them questionable. In fact, their neglect deprives informal logicians of useful resources in appraising defeasible argument. A probabilistic analysis of argument cogency clarifies this core concept of informal logic, providing a corrective to its usual applications, but also yields a sense in which a formal and an informal normative approach align. In particular, we identify content features of defeasible argument on which the RSA conditions depend, namely (i) the change in the acceptability of the reason; (ii) the reason's sensitivity and selectivity to the claim, and (iii) one's prior credence in the claim itself, together with (iv) the contextually determined thresholds of acceptability for reasons and for claims (i.e., inferential sufficiency).

This analysis of the RSA conditions corrects their orthodox informal understanding and application, for it shows that a threshold application of the orthodox RSA criteria as update-gates can lead to what on a probabilistic construal are errors.

Moreover, while satisfying the cogency conditions is entitlement establishing, cogency is nevertheless an obligatory norm, rather than a permissive one. Finally, our analysis advances the probabilistic approach to argumentation, particularly regarding the fallacies and argumentation schemes. We show how the informal treatment of schemes and fallacies as situated successes or failures to meet the RSA cogency conditions coheres with a probabilistic appraisal of schemes and fallacies. This explains how these two normative theories can in principle, if not in fact, agree, and so contributes to the cohesion and completeness of an account of argumentative norms.

Section 2 identifies the anti-formalist sentiments motivating informal logic, presents cogency as a normative standard for defeasible argument, and briefly surveys recent probabilistic treatments of argumentation. Section 3 introduces the probabilistic calculus, seeking to make its resources more accessible to informal logicians. Section 4 then offers a probabilistic analysis of the informal notions acceptability, relevance, and sufficiency. Section 5 discusses consequences of this analysis.

(Available upon request, the full paper is forthcoming with Synthese under the above title.)

References

[1] Woods, J. (2000). How philosophical is informal logic? *Informal Logic* 20: 139–167.

Defining Argument Weighing Functions

1 Dec 14:30

Thomas F. Gordon

Fraunhofer FOKUS, Berlin

Dung designed abstract argumentation frameworks [1] to model attack relations among arguments. However, a common and arguably more typical form of human argumentation, where pros and cons are weighed and balanced to choose among alternative options, cannot be simply and intuitively reduced to attacks. In [2], Doug Walton and I defined a new formal model of structured argument which generalizes Dung abstract argumentation frameworks to provide better support for argument weighing and balancing, enabling cumulative arguments and argument accrual to be handled without causing an exponential blowup in the number of arguments. Dung proposed a pipeline model of argument evaluation for abstract argumentation frameworks, where first all the arguments are evaluated and labelled, at the abstract level, and then, in a subsequent process, the premises and conclusions of the arguments are labelled, at the structured argument level. This pipeline model makes it impossible to make the weight of arguments depend on the labels of their premises. To overcome this problem, in our new model the weight of arguments and labels of statements can depend on each other, in a mutually recursive manner. The new model is a framework which can be instantiated with a variety of argument weighing functions. In this presentation, we illustrate this feature by defining a number of argument weighing functions, including: 1) simulating linked and convergent arguments, by making the weight of an argument depend on whether all or some of its premises are labelled **in**, respectively; 2) making the weight of an argument depend on one or more meta-level properties of the argument, such as the date or authority of the scheme instantiated by the argument; 3) modeling a simple form of cumulative argument, by making the weight of an argument depend on the percentage of its **in** premises; 4) making the weight of an argument depend on the percentage of its **in** "factors", from a set of possible factors, where premises represent factors; and, finally 5) making the weight of an argument depend on a weighted sum of the **in** properties of an option, in the style of multi-criteria decision analysis, where premises model properties of an option.

References

[1] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357, 1995.

[2] Thomas F. Gordon and Douglas Walton. Formalizing balancing arguments. In *Proceeding of the 2016 conference on Computational Models of Argument (COMMA 2016)*, pages 327–338. IOS Press, 2016.

The Influence of an Agent's Credibility on the Questioning Agenda in a Multi-Agent Environment

1 Dec
11:30

Paweł Łupkowski and Mariusz Urbański

Adam Mickiewicz University, Poznań

The main aim of our talk is to address the issue of how to model the influence of agent's credibility on the obtained questioning results. We are interested in this problem in the context of problem solving based on information retrieval from distributed sources with different credibility levels. We address intuitions similar than the ones presented by [7], i.e. how a problem solving may be planned and executed in multi-agent environment with the clear distinction of common knowledge needed for achieving common goal and private knowledge of agents which need not to be revealed.

Our formal framework is Inferential Erotetic Logic [12, 14] with its epistemic interpretation presented by [8]. This logic focuses on inferences whose premises and/or conclusion are questions (erotetic inferences). The epistemic component allows for defining such concepts as individual and group question as well as group and common knowledge.

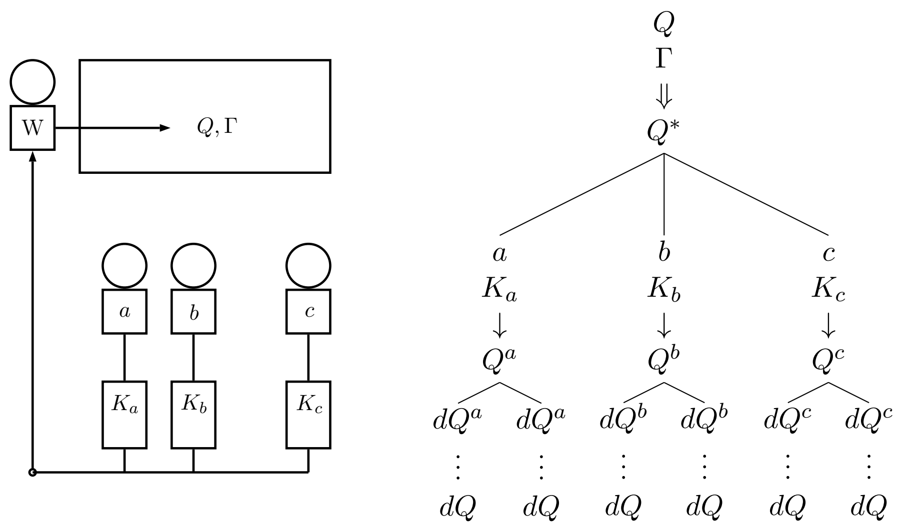


Figure 1: The blackboard architecture for the cooperative problem solving with distributed information sources

The main tool we use for modelling a questioning agenda[5] for an agent is *erotetic search scenario* (e-scenario for short, [13]). E-scenario represents a certain map of possible courses of events in a given questioning process for the initial question, with respect to initial knowledge. Each path of an e-scenario is one of the ways the process might go depending on the answers obtained to queries. E-scenarios has proven to be a powerful logical tool for modelling cognitive goal-directed processes and problem solving (see [13], [11], [5], [6], [9]). In our approach e-scenarios allows us to explain the relevance of questions appearing in interactions between agents in the way proposed by [10]. In order to tackle the multi-agent interaction we will use the *blackboard architecture* [1].

In the proposed model we are able to express a situation where a group of agents solves a complex problem in the cooperative manner. The central element is the blackboard visible for all the agents. We have one main agent, called the Writer (who writes down questions and information on the blackboard) and other agents (let us refer to them as a, b, c, etc.). As for agents from the group we assume that they have different knowledge and different credibility/expertise levels.

We assume that the group attacks a complex question which cannot be resolved by any agent from the group by her own (so called *group question*). This question is then written down on the blackboard along with the common knowledge of the group. Afterwards the initial group question is decomposed on a group level into a series of simpler questions (using the aforementioned common knowledge). In what follows, these simpler questions are analysed by group members (at this level e-scenarios for each agent are introduced). The last step is collecting the solutions to these auxiliary questions by the Writer and establishing the answer to the initial question. At this level we analyse and discuss different ways the answer might be reached with respect to the credibility of agents.

[5]For an alternative formal approach to research agendas see e.g. [3] and [4] (for argumentative contexts)

Credibility of an agent is expressed by a certain labelling. In accordance with the ideas underlying the concept of a Labelled Deductive System [2], we define an appropriate algebra of labels and a preferential ordering of labels, allowing for stratification of "quality" of information, obtained from different sources.

Acknowledgments

This work was supported by funds of the National Science Centre, Poland (DEC-2012/04/A/HS1/00715).

References

[1] Corkill, D. Blackboard Systems. *AI Expert* 6(9) (1991): 40–47.

[2] Gabbay, D. *Labelled Deductive Systems* (Clarendon Press, 1996).

[3] Genot, E. The game of inquiry: the interrogative approach to inquiry and belief revision theory. *Synthese* 171(2) (2009): 271–289.

[4] Jacot, J. and Genot, E. and Zenker, F. From reasonable preferences, via argumentation, to logic. *Journal of Applied Logic* 18 (2016): 105–128.

[5] Leszczyńska-Jasion, D. and Łupkowski, P. Erotetic Search Scenarios and Three-Valued Logic. *Journal of Logic, Language and Information* 25(1) (2016): 51–76.

[6] Łupkowski, P. *Logic of Questions in the Wild. Inferential Erotetic Logic in Information Seeking Dialogue Modelling* (College Publications, 2016).

[7] Pardo, P. and Godo, L. A temporal argumentation approach to cooperative planning using dialogues. *Journal of Logic and Computation* (2015). DOI: 10.1093/logcom/exv007.

[8] Peliš, M. *Inferences with Ignorance: Logics of Questions (Inferential Erotetic Logic and Erotetic Epistemic Logic)* (Karolinum, 2016).

[9] Urbański, M. and Paluszkiewicz, K. and Urbańska, J. Erotetic Problem Solving: From Real Data to Formal Models. An Analysis of Solutions to Erotetic Reasoning Test Task. In Paglieri, F. (ed.), *The Psychology of Argument: Cognitive Approaches to Argumentation and Persuasion* (College Publications, 2016).

[10] Roberts, C. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics* (1996): 91–136.

[11] Urbański, M. and Łupkowski, P. Erotetic Search Scenarios: Revealing Interrogator's Hidden Agenda. In Łupkowski, P. and Purver, M. (eds.), *Aspects of Semantics and Pragmatics of Dialogue. SemDial 2010, 14th Workshop on the Semantics and Pragmatics of Dialogue* (Polish Society for Cognitive Science, 2010), 67–74.

[12] Wiśniewski, A. *The Posing of Questions: Logical Foundations of Erotetic Inferences* (Kluwer AP, 1995).

[13] Wiśniewski, A. Erotetic Search Scenarios. *Synthese* 134 (2003): 389–427.

[14] Wiśniewski, A. *Questions, Inferences, and Scenarios* (College Publications, 2013).

The Ellsberg Paradox and the Weight of Arguments

2 Dec
15:00

William Peden
Durham University

Confirmation theory has traditionally focused on assessing the balance of evidence: "Given some premises E and a conclusion H, how much does E support H?" However, there are many other aspects of argumentation, and we might be able to answer some paradoxes in confirmation theory and decision-theory by studying these other aspects using formal methods.

The Ellsberg Paradox is a classic paradox of decision. It challenges the standard normative decision theory framework. In this paper, I explore a possible answer to the Ellsberg Paradox: the standard decision theory framework generates the Ellsberg paradox because it ignores an important aspect of arguments, which is what John Maynard Keynes called the "weight" of arguments. I argue that we can provide a formal rationalization of our beliefs about the Ellsberg Paradox by modifying the standard decision-theory framework to recognise that rational agents have a preference for acting on greater rather than lesser evidence.

In standard decision theory, we model a person as having (1) Bayesian probabilities P for each of a set of possible outcomes O given some decisions D and (2) utilities U for each member of the set of possible outcomes. By multiplying P and U for each O, and summing them for each D, we can calculate the expected utility for every possible decision that the modelled person faces.

In the Ellsberg Paradox, we are confronted by a betting choice between two possible decisions A and B, where these decisions have equal expected utility, but B is based on more detailed information about the betting set-up. For example, imagine you are offered a choice between drawing balls from either Box A or Box B. You are betting that the ball that will be drawn is not green. You know that each box has 90 balls. Box A has between 0 and 60 green balls, 0 and 60 red balls, and exactly 30 white balls. Box B has 30 green balls, 30 red balls, and 30 white balls. Which box should you choose? Empirically, most people have a preference for Box B over Box A. However, using standard decision theory and some uncontroversial assumptions, the expected utility of choosing Box B is equal to the expected utility of choosing Box A, because the Bayesian probabilities for not drawing a green ball from either box are identical.

One possible rationalization of our preference for Box B is that the argument that the ball will not be green cites more evidence. The argument from our evidence about the betting set-up to

H1: ‘The ball drawn will not be green if drawn from Box B.’

has more weight than the argument from evidence to

H2: ‘The ball drawn will not green if drawn from Box A.’

because our evidence tells us more about the relative frequencies of balls in in Box B than Box A.

If we could successfully formalize the notion of the weight of arguments in this rationalization, then a simple addition to the standard decision theory model might answer the Ellsberg Paradox: ceteris paribus, if the argument for an action A has a greater weight than the argument for an action B, then we ought to do A, even if A and B have equal expected utility.

Imprecise probability theory has become a popular way of formalising the concept of the weight of arguments, which today is often called the “weight of evidence”. I argue that Henry Kyburg’s theory of “evidential probability” can be used to formalize this concept in simple scenarios like the Ellsberg Paradox. I finish with a novel reconciliation of evidential probability theory, in which probabilities are typically imprecise, and standard decision theory, in which probabilities are precise.

Overall, I reach three points. Firstly, a minor addition to standard decision theory can evade the Ellsberg Paradox. Secondly, evidential probability theory is a useful addition to a formal epistemology. Thirdly, the Ellsberg Paradox is an example of how we can sometimes answer paradoxes by studying the multiple aspects of argument assessment.

Arguing About Constitutive and Regulative Norms

2 Dec
16:00

Gabriella Pigozzi[1] and Leon van der Torre[2]

[1]Université Paris-Dauphine

[2]University of Luxembourg

Norms regulate people’s behaviour. It is thus fundamental to have approaches to represent and reason about the different types of norms. Regulative norms indicate what is obligatory or permitted. Deontic logic formalises concepts like obligation and permission and its history has been marked by many puzzles and paradoxes. Many of the deontic logic paradoxes contain obligations conditional on a violation, as the gentle murderer paradox, which says that one should not kill but, if you kill, you should do it gently. A contrary-to-duty obligation expresses what one should do when obligations have been violated. This kind of sentences have often counterintuitive consequences. In formal deontic logic, permissions are studied less frequently than obligation. For a long time, it was naively assumed that it can simply be taken as a dual of obligation, just as possibility is the dual of necessity in modal logic. Perhaps most importantly, Bulygin observed that an authoritative kind of permission must be used in the context of multiple authorities and updating normative systems: if a higher authority permits you to do something, a lower authority can no longer prohibit it.

Normative reasoning distinguishes three types of norms: constitutive, regulative and permissive norms. Deontic logic has been concerned mainly with regulative norms, but the logic of constitutive norms is a subject of study on its own. Constitutive norms are rules that create the possibility of or define an activity. For example, according to Searle, the activity of playing chess is constituted by action in accordance with these rules. As another example, a signature may count as a legal contract, and a legal contract may define a permission to use a resource and an obligation to pay. Searle points out that, unlike regulative norms, constitutive rules do not regulate actions but define new forms of behaviour. Constitutive norms link brute facts (like the signature of a contract) to institutional facts (a legal contract) and are usually represented as count-as conditionals: X counts as Y in context C. Searles analysis insisted on the contextual nature of constitutive norms: a signature counts-as a legal contract when posed on a paper stating the terms of such a contract. If, on the other hand, I exercise writing my signature on a white sheet, that does not constitute a legal contract.

We are not aware of any argumentation analysis of constitutive norms, though there are various papers considering arguments among regulative and permissive norms, and several frameworks have been proposed for legal argumentation. Recently, Beirlaen and Straßer presented an argumentation system for detaching conditional obligations based on Dungs grounded semantics. They define two ways in which deontic arguments may attack one another and present some mechanisms for conflict-resolution. Our work is related to Beirlaen and Straßer, though our aim here is to explore how conflicts among constitutive and regulative norms can be resolved.

An important challenge in defining constitutive norms like counts-as conditionals is defining their relation with regulative norms like obligations and permissions. Boella and van der Torre use the notion of a logical architecture combining several logics into a more complex logical system, called logical input/output nets (or lions). Here, we focus on the analysis of the types of conflicts (and on their resolution) that can arise between constitutive and regulative norms. In this paper we pursue the following:

1. How can we classify the different ways in which conflicts can be resolved among constitutive and regulative norms? There may be conflicts between facts, between facts and obligations, among obligations, and combinations thereof.

2. How to resolve conflicts among constitutive and regulative norms using argumentation? For example, we must be careful not to allow for a normative version of wishful thinking, where obligations defeat factual statements.

Arguments are often presented as natural deduction derivations or as pairs (support, conclusion). In normative systems, arguments may be represented by triples (brute, institutional, deontic), where constitutive norms derive institutional facts from brute facts, and regulative norms derive deontic facts from institutional facts. For instance, a triple (signature, contract, payment). Such expressions were recently introduced in a deontic logic of Sun and van der Torre, and we therefore use their logic.

At first sight, it may seem that we can simply define an attack of the argument (brute1,institutional1,deontic1) on (brute2,institutional2,deontic2) iff the set of formulas {brute1,institutional1,deontic1,brute2,institutional2,deontic2} is inconsistent. However, this definition would not be satisfactory for at least the following reasons:

Violations: Consider an argument (brute1,institutional1,deontic1) where the set of formulas brute1,institutional1,deontic1 is already inconsistent, representing a violation. Such an argument would attack all other arguments, whereas for contrary-to-duty reasoning it is pertinent that violations do not have such kind of interference effects;

Wishful thinking: According to this definition, it would be possible that an obligation attacks a fact, which would be a kind of normative wishful thinking.

Deontic dilemmas: Moreover, for some applications this definition may be too strong. For example, one may require the brute and institutional facts to be consistent, but allow for the representation of conflicting obligations.

In this paper, we provide an analysis of the resolution of conflicts between constitutive and regulative norms, which leads to three properties regarding conflicts, wishful thinking and reinstatement. We introduce an argumentation theory for constitutive and regulative norms, and we show that for one particular kind of definition of attack, all properties are satisfied.

Group Polarization and Argument Strength

1 Dec
12:00

Carlo Proietti
University of Lund

Group-induced attitude polarization ([6]) occurs "when an initial tendency of individual group members toward a given direction is enhanced following group discussion" ([4]). This phenomenon occurs very often in real-life scenarios such as political debate via social media ([7]). Polarization often leads subgroups towards opposite directions (bi-polarization effects) and therefore speaks against the assumption that debate among informed individuals should lead to consensus and be truth-conducive. The fundamental question to ask is whether polarization may happen in situations of perfect communication within a group and with individuals performing rational information updates. The literature on *Argumentation Frameworks* ([1]) provides enough tools to capture the notion of *rational update of information* and therefore allows asking this question in a precise way.

Large field experiments conducted in the 1970s isolated *Persuasive Arguments Theory* (PAT) as the most effective explanatory clue for polarization. The PAT explanation assumes that individuals become more convinced of their view when they hear novel and persuasive arguments in favor of their position, and therefore "Group discussion will cause an individual to shift in a given direction to the extent that the discussion exposes that individual to persuasive arguments favoring that direction" ([4]). The PAT explanation posits that polarization may arise by a rational process due to individuals refining their argumentative knowledge. However, the exact mechanisms of how this process may unfold are still unclear. To understand polarization we need to decompose it into its basic ingredients, i.e. (a) a plurality of agents, (b) a debated issue a, (c) possibly different prior opinions held by the agents about the debated issue, (d) *pro* and *contra* arguments possibly related with each other (refutation, support, counterattack etc.), (e) the update, by the agents, of their argumentative basis and (f) a measure of the individuals' belief about (b), this being defined on the basis of (c). Elements (a)-(e) can be encoded via purely qualitative *Bipolar Argumentation Frameworks* (BAF) introduced by [1], while (f) requires a *measure* that is provided by quantitative approaches to justification in AF.

The polarization mechanism can be illustrated, on a qualitative level, by means of a simple example involving BAF. A BAF consists of a triple $\mathcal{B} = (\mathcal{A}, \mathcal{R}^a, \mathcal{R}^s)$, where $\mathcal{A}$ is a finite and non-empty set of arguments, $\mathcal{R}^a$ and $\mathcal{R}^s$ are binary relations over $\mathcal{A}$, called the attack and the support relation. $a\mathcal{R}^a b$ (resp. $a\mathcal{R}^s b$) means "a attacks b" (resp. "a supports b"). Here we use BAF with a dynamic turn to understand the steps of an argumentative debate among n agents. Indeed, given a BAF $\mathcal{B}$, the information available to the participants to a debate can be represented as a n-tuple $\mathcal{B}_1, \ldots, \mathcal{B}_n$ of BAFs where each $\mathcal{B}_1$ (for $1 \leq i \leq n$) is a subgraph of $\mathcal{B}$ and is the information of agent i. The result of a debate/exchange of arguments between two agents j and k can be framed as an operation on the graphs $\mathcal{B}_j$ and $\mathcal{B}_k$. Many such operations are possible. The most obvious one to meet the standards of a rational interaction is union. In general, given a vector $(\mathcal{B}_1, \ldots, \mathcal{B}_n)$ of BAFs we define the update after information exchange as $\mathcal{B}^*_i = (\bigcup_{i=1}^n \mathcal{B}_i, \bigcup_{i=1}^n \mathcal{R}^a_i, \bigcup_{i=1}^n \mathcal{R}^s_i)$ for each i. The updated information of j after exchanging with k is then $\mathcal{B}^*_j = \mathcal{B}_j \cup \mathcal{B}_k$. It is very easy to see, even in this purely qualitative framework, that polarization may easily happen throughout debate. Consider a simple example of an exchange on a specific issue a with two agents 1 and 2 where $\mathcal{B}_1 = (\{a\}, \emptyset, \emptyset)$ and $\mathcal{B}_2 = (\{a, b\}, \emptyset, (b, a))$. $\mathcal{B}_1 \cup \mathcal{B}_2$ is clearly $(\{a, b\}, \emptyset, (b, a))$. $\mathcal{B}_2$ remains unaltered while $\mathcal{B}_1$ gets now enriched with the support relation (b, a). In other words, agent 1 gets more convinced of a by the support of b. The group therefore "shifts" in the direction of more support to a.

As a further step, one can endow a BAF with a measure of the strenght of its arguments and its relations of supports and attack ([5]). This automatically provides a probabilistic measure of the *likelihood* of a single argument and of a set of arguments. By this enrichment one can provide a more precise quantitative definition of **polarization as the increased/decreased likelihood of a debated issue after update**. I show, by elaborating on the previous example, that such polarization arises by the described mechanism. I also briefly outline some alternative possibilities to capture quantitative polarization with alternative AF semantics ([3]).

References

[1] C. Cayrol and M.C. Lagasquie-Schiex. On the Acceptability of Arguments in Bipolar Argumentation Frameworks. *Lecture Notes in Computer Science*, 3571: 378–389, 2005.

[2] P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77 (2): 321–357, 1995.

[3] D. Grossi and S. Modgil On the Graded Acceptability of Arguments. *Proceedings of the IJCAI 2015*: 868–874, 2015.

[4] D.J. Isenberg. Group Polarization: A critical review and a Meta-Analysis. *Journal of Personality and Social Psychology* 50 (6): 1141–1151, 1986.

[5] H. Li, N. Oren and T.J. Norman Probabilistic Argumentation Frameworks *Lecture Notes in Computer Science* 7132: 1–16, 2011.

[6] J.A. Stoner A comparison of individual and group decision involving risk MA thesis, Massachusetts Institute of Technology, 1961.

[7] C. Sunstein. *Why societies need Dissent.* Cambridge, Harvard University Press, 2003.

Simple Rules for Probabilistic Commonsense Reasoning

30 Nov
15:00

Adam Purtee and Lenhart Schubert
University of Rochester

A long-sought goal of AI is the mechanization of human-like reasoning and argumentation based not only on firm knowledge supporting deduction but also generic conditional knowledge supporting tentative or probabilistic inferences. Arguably, most of the general knowledge people possess is of this type. We propose probabilistic "amplifying rules", "Bayes rules", and "categorization rules", combined with an algebraic conception of probability, as a basis for commonsense probabilistic reasoning, i.e., the generation of new, probabilistically qualified propositions from a commonsense knowledge base. The proposed framework allows for both "hard" (e.g., universally quantified) knowledge and knowledge in the form of the above types of rules, expressing that under certain specified conditions, a conclusion is warranted to some degree, or the members of some spectrum of alternative conclusions are warranted to varying degrees.

Our rules are meant to be quantitative formalizations of generic statements about the world. Such statements, in sentential form, come in two distinct types: ones that generalize over individuals that realize *kinds* (such as the kind, dogs, or the kind Rottweilers); and ones that generalize over events or situations (e.g., dogs often bark when agitated; people usually use paint rollers to paint walls; etc.). We can view many generic sentences as statistical claims about the world (sometimes quite explicitly so, as in *One out of four times, the airlines will lose your checked baggage when you travel*). We regard our rules as grounded in statistical claims as well, although, when applied to particular instances, the probabilities they lead to are no longer statistical but rather degrees of belief. Our proposed rules are able to capture a wide variety of statistical knowledge expressed in natural language.

More specifically, amplification/attenuation rules are used when a direct causal (or anti-causal) relationship is known to hold. E.g., watering the lawn daily is known to cause grass to grow, whereas sowing salt into the soil has the opposite effect. The net effect of an a-rule is monotonic, in the sense that even when their corresponding parameters are arbitrarily small, the likelihood of the consequent is monotonically increased (or decreased, in the case of attenuation rules).

An example application is the claim that for each common friend between a pair of persons, the likelihood that the pair are friends increases. We can capture this with a single amplification rule, and the consequences thereof are that as the number of common friends increases, so does the likelihood of transitive friendship, and that the strength of the boost in friendship likelihood decays with the distance, i.e., the number of intermediary friendships necessary to establish a transitive connection.

Two key processes are portrayed in this example: the accumulation of evidence from common friendships, and the reduced evidential strength derived from distant transitive connections in comparison with more direct ones. These complementary processes find expression in our approach as evidence combination and forward chaining, and we have developed formal rules for such reasoning with each of our three classes of proposed rules.

As another example, we may consider that while a typical dog may be friendly, upon learning that it is a Labrador, we may grow more confident in its friendliness, and yet further knowledge that is has been ill-treated and starved may cause us to revise our conclusion. Our method allows for defeasible reasoning by accumulation of evidence in favor of and evidence opposed to particular conclusions.

Categorization rules are intended to capture probability distributions over mutually exclusive alternatives, and are particularly well-suited to modeling proportional partitioning of kinds (of entities or events) into sub-kinds.

Bayesian rules are so-named to evoke the familiar Bayesian rule of conditioning, upon which they are based. The direction of their effect can be positive or negative, depending both on the rule parameter and the prior probability of the modified event.

Empirically, our rules are able to exactly match the outputs of an artificially constructed, medical diagnosis-themed Bayesian network, and in principle can represent more general Bayesian networks as well.

Further, we verified that with a simple, two-rule baseline, we are able to match performance of Richardson and Domingos [1] on the UW-CSE dataset, which consists of predications about individual members of a computer science department, publication authorship, teaching roles, and teaching-assistant roles, using these to infer advisor-advisee relations by means of Markov Logic Networks (MLNs). We obtain an area under the precision-recall curve of 0.225 as compared to the best reported results in that work of 0.220. While methods enhanced with structure-learning techniques later improved upon the original MLN results, our results provide an empirical verification of the utility of our approach. We emphasise that these results are preliminary, and represent baseline performance with a-rules only.

Our work connects two distinct threads of research into logical representations suitable for capturing the kinds of knowledge expressed through natural language [3], as well as an algebraic characterization of probability which is closely connected to logical and-or-not networks [2].

The existing approaches that come closest to enabling such general reasoning as our method are nonmonotonic reasoning (NMR) and production systems (also called rule-based systems or expert systems). NMR methods extend deductive methods to allow conclusions to be drawn tentatively in the absence of contrary evidence. However, they fail to allow for the quite subtle *quantitative* intuitions that people have about the world around them, thanks to both "statistical experience" and verbally transmitted knowledge.

Our framework for commonsense reasoning refines NMR methods by allowing for rules whose consequent is affirmed with some degree of certainty (or with a "spectrum" of consequents affirmed with various degrees of certainty). In relation to production systems, our rules operate on logical, formally interpretable formulas, are restricted to a set of simple, natural types, allow for alternatives in their conclusions, and interact in well-motivated ways when applied in sequence or in parallel.

Being built upon an extension of first-order logic, our method is compatible with quantified and ground deduction, and supplies transparent provenance for arguments about belief strengths. We strictly generalize path-based methods of nonmonotonic reasoning which only operate with "usually is-a" and "usually is-not-a" relationships. Our numerical parameters have natural interpretations as conditional probabilities, and thanks to its expressivity, our representation is well-suited to working with knowledge expressed through natural language.

Acknowledgements

The work was supported in part by DARPA grant W911NF-15-1-0542.

References

[1] Richardson, M. and Domingos, P. 2006. Markov logic networks. *Machine Learning* 62(1–2): 107–136.

[2] Schubert, L. 2004. A new characterization of probabilities in Bayesian networks. *Proceedings of the 20th conference on Uncertainty in Artificial Intelligence* (AUAI Press): 495-503.

[3] Schubert, L. and Hwang, C.H. 2000. Episodic logic meets Little Red Riding Hood: a comprehensive natural representation for language understanding. In Iwanska, L. and Shapiro, S, (eds.), *Natural Language Processing and Knowledge Representation* (MIT Press): 111–174.

Argument Evaluation Based on Proportionality

2 Dec 11:30

Marcin Selinger
University of Wroclaw

The aim of this paper is to present and discuss a model of numerical evaluation of complex argumentation structures. The model was elaborated in [3] and extended in [4]. It is based on what I call the principle of proportionality, which says that the strength of argument should vary proportionally to the values assigned to its components.

1. On different levels of arguments' complexity these components are: (i) premises collected in order to (ii) support conclusions in linked arguments, (iii) entire linked arguments constituting convergent and serial arguments, and finally (iv) complex pro and contra argumentation structures forming conductive arguments (following [6] conductive arguments are understood as pro-contra arguments). This syntax can be formalized by means of usual set-theoretical notions. Namely, the argumentation structures can be regarded as finite and non-empty sets of pro and contra sequents, i.e. triples of the form (P, c, d), where P is a finite and non-empty set of premises, c is a conclusion, and d is a Boolean value denoting whether the premises are pro or contra the conclusion. On this ground, basic syntactic concepts (first premise, intermediate premise, final conclusion) and properties of arguments (divergence, coherence, circularity) can be defined (cf. [3, 4]). Independently of certain technical differences this framework seems to be in line with some other formal approaches to the syntax of arguments as e.g. Walton & Gordon's Carneades Argumentation System (cf. [5]) or Prakken's $ASPIC+$ (cf. [2]).

2. The acceptability (credibility) of sentences is expressed within the probabilistic scale, i.e. it can vary in the range of values between 0 and 1 (1 stands for 'fully acceptable', 0 for 'fully inacceptable', and 1/2 for 'undecided'). Formally, the process of evaluation of argument strength can be described as extending of some partial function assigning acceptability values to some sentences of some language (i.e. evaluation function), which must be defined for the first premises of an examined argument, to the set containing all the sentences occurring in this argument, in particular its final conclusion. The strength of an argument can be measured as just the acceptability assigned to its final conclusion by the new, extended evaluation function.

The distribution of acceptability values proceeds from the first premises to the final conclusion of an examined argument, in accordance with its structure and with the principle of proportionality. In the order determined by the levels of argument complexity the procedure can be described as follows:

(i) the acceptability of the (set of the) premises of linked arguments is calculated as the multiplication of the values assigned to each of these premises separately;

(ii) the acceptability of (the conclusions of) linked arguments is calculated as the multiplication of the acceptability of their premises and their conditional acceptability, which is a parameter expressing the relevance between the premises and the conclusion, i.e. the weight assigned to the type of inference under consideration (different types of inference that determine conditional acceptability can be distinguished as corresponding to argumentation schemes (cf. [7]); obviously, for deductive arguments conditional acceptability equals 1; for defeasible arguments it must be estimated using any accessible (e.g. logical, statistical or empirical methods, by expert opinions, or possibly by convention);

(iii) the acceptability of convergent arguments is calculated using a special, commutative and associative operation of addition $\oplus$ which is defined by the formula:

$$x \oplus y = 2x + 2y - 2xy - 1$$

where x and y are acceptability values assigned to their converging counterparts;

(iv) contra-arguments, i.e. arguments denying some conclusions are interpreted as the pro-arguments supporting the negations of these conclusions; the acceptability of conductive arguments is calculated using subtraction, i.e. the acceptability values of their acceptable contra-counterparts are subtracted from the acceptability values of their acceptable pro-counterparts; 1/2 is added to this value in order to obtain the final result within the scale (i.e. between 0 and 1).

The above computations have their graphic illustrations reflecting the idea of proportionality. Illustrated proportions are shown with help of the Thales theorem applied to line segments (intervals) whose lengths represent computed values (cf. [3, Fig. 2]; see also [4, Fig. 4]).

3. Attack relation can be defined within the outlined model. Such a definition was proposed in [4]. It consists of three parts, each of them covering one of three traditionally distinguished kinds of attack, namely, attack on a premise, conclusion or relationship between them. The definition can serve us as a link between logic and dialectic of argument. It can also be viewed as a specification of the abstract argumentation theory [1], where attack relation is a primitive notion.

4. Since the proposed model of evaluation is numerical, it is naturally sensitive to the double counting fallacy, in particular while computing convergent arguments. In simple cases calculations can be easily adjusted by means of logical analysis. In dubious and problematic cases this model can still be useful when it is combined with other approaches. For instance, the operation $\oplus$ gives us an upper bound, whereas the preponderance of evidence proof standard (cf. [6]) can give us a lower bound for the strength of some convergent argument being considered.

References

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

[2] Reed, C. (2014). Structured arguments and their aggregation: A reply to Selinger. Argumentation 28(3): 395–399.

[3] Selinger, M. (2014). Towards formal representation and evaluation of arguments. *Argumentation* 28(3):379–393.

[4] Selinger, M. (2015). A formal model of conductive reasoning. In *Proceedings of the 8th ISSA, Amsterdam 2014*: 1331–1339.

[5] Selinger, M. & Koszowy, M. (2016). Considering Carneades as a framework for informal logic: A reply to Walton and Gordon. *Informal Logic* 36(2):217–237.

[6] Walton, D. & Gordon, T.F. (2015). Formalizing informal logic. *Informal Logic* 35(4), 508-538.

[7] Walton, D., Reed, C. & Macagno, F. (2008). *Argumentation Schemes* (Cambridge University Press).

Modeling the Strength of Arguments in Dynamic Epistemic Probabilistic Logic

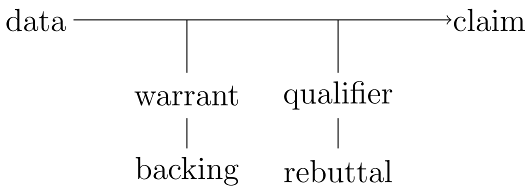
1 Dec 11:00

Chenwei Shi and Sonja Smets

Institute for Logic, Language and Computation, University of Amsterdam

In this work we focus on modeling the strength of arguments in a dynamic epistemic probabilistic logic [6]. The motivating idea that underlies our formal approach is that arguments can be viewed as providing a justification for an agent's belief update. In our approach we follow Toulmin's scheme [7] when modeling the structure of arguments.

We start with introducing the Toulmin scheme (as depicted in the following graph) for the structure of an argument via an example.



Example 1 *Ben's doctor told him that he might have a horrible disease d. Ben was very shocked, since he heard that the chance of getting this disease was small, say* 4% *chance. The doctor's argument for his diagnosis goes as follows: I received a report on your blood test, it shows that your blood is short of a very important constituent c. (Given you are really short of c, the error rate of the test is* 20% *and given you are not short of c, the error rate of the test is* 1%*).*

According to the study, (you have 80% *chance of lacking c given that you are suffering from disease d and you have* 40% *chance of lacking c given you do not have disease d), if your blood is in lack of c, it is most likely that you have disease d. So you might have disease d.*

In this example, the doctor makes the diagnosis that Ben might get disease d given the report on Ben's blood test. The diagnosis is the claim in Toulmin's scheme and the "lack of c" according to the test report is what we call "data". Note that the data consists of two parts, the symptom "lack of c" shown in the report and the test which supports the report. The correlation between the symptom and the disease d, backed by the medical study, is the warrant for the inference from the lack of c to disease d. Note that the expressions "most likely" and "might", as the qualifier in Toulmin's scheme, are used to qualitatively specify the force of the claim based on the data.

The argument of the doctor lends credence to his diagnosis. How strong is the argument of the doctor? For Ben, the stronger is the doctor's argument, the more credible is the doctor's diagnosis. *Supported by this idea, we take an argument to be the trigger for an agent's belief change. Thus its strength can be measured by the extent of the agent's belief change.*

Dynamic epistemic logics [1, 3] provide us with a tool of modeling and reasoning about different types of information change in multi-agent scenarios . The flow of information is triggered by so-called events, which are themselves the object of study in this logical setting. Later developments in [2, 6] have provided an extended logical setting that allows us not only to express the qualitative features of epistemic and doxastic states and events but also their quantitative features. Such a framework makes it possible to express the likelihood for an event to happen in a certain situation as well as the subjective credence that an agent can assign to both propositions and events.

In this paper we follow [6] and use a single-agent probabilistic model $\mathcal{M}$ to model the agent's belief state. Moreover, to fully encode Toulmin's scheme, we modify the probabilistic update model $\mathcal{A}$ and adjust its interpretation. The updated model $(\mathcal{M} \otimes \mathcal{A})$ can characterize the influence of the argument on the agent's belief state. In line with [1] the notation $\otimes$ is used to compute how the event model $\mathcal{A}$ acts on the given input model $\mathcal{M}$. As such, our work follows the close connection between Toulmin's scheme, logic and information dynamics that has been pointed out by Johan van Benthem in [4, 5].

In the remainder of this abstract, we explain how the probabilistic event models are used to analyze arguments and their strength, while further details of the analysis will be offered in the full paper. For example, the definitions of the strength of an argument, the attack relation and the defeat relation between arguments, an answer to the question "can weaker arguments defeat stronger arguments?" and so on.

Modeling Toulmin's Scheme Let $\text{RV} = \{A, B, C, \ldots\}$ whose members are finite sets of propositions which are mutually exclusive and collectively exhaustive. More precisely, take any $A = \{a_1, \ldots, a_n\} \in \text{RV}$, $\vdash \bigvee_{i=1}^{n} a_i \leftrightarrow \top$ and $\vdash \bigwedge_{i=1}^{n} a_i \to \bot$. Let GRV be a set which includes RV and is closed under taking the product (if $X, Y \in \text{GRV}$, then $X \times Y \in \text{GRV}$).

Definition 1 (Probabilistic Update Model (PUM)) *Probabilistic update models are structures* $\mathcal{A} = \langle \text{Premise}_{\mathcal{A}}, \text{Ground}_{\mathcal{A}}, \text{Claim}_{\mathcal{A}}, \text{Warrant}_{\mathcal{A}}, \text{Backing}_{\mathcal{A}} \rangle$ *where (we ignore the subscript when there is no ambiguity)*

- $\text{Premise} \in \text{GRV}$ *is a set of possible values the premise of the argument can take;*

- Ground : Data $\to \mathbb{R}^+ \cup \{0\}$ *assigns to each value in* Premise *a real number, representing the support each value get from the ground.*

- Claim $\in$ GRV *is a set of possible values the conclusion of the argument can take;*

- Warrant : Claim $\times$ Data $\to [0,1]$ *is a conditional probability function where* $\sum_{d_i \in \text{premise}} \text{Warrant}(d_i|c_j) = 1$ *for each* $c_j \in$ Claim*;*

- Backing : $\{\text{Warrant}\} \to [0,1]$ *assigns a number to the function* Warrant *which measures the reliability of the warrant given the backing.*

We use an PUM to capture the ingredients of an argument. In particular, the set Premise and the function Ground constitutes the data in Toulmin's scheme. And the set Claim, the functions Warrant and Backing formalize the corresponding parts in Toulmin's scheme. Note that we abstract away from almost all the contents of ground, warrant and backing in an argument and only specify the quantitative relations between them and other parts in Toulmin's scheme. However, these quantitative relations are not generally specified in a real argument or even realized by the arguer. Moreover, qualifier and rebuttal are not explicitly modeled. We will address these issues in the full paper.

To illustrate how PUM works, we formalize the doctor's argument in the example 1:

$$\mathcal{A} = \langle \text{Premise} = \{loc, \overline{loc}\}, \text{Ground} = (80, 1), \text{Claim} = \{d, \overline{d}\},$$
$$\text{Warrant} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}, \text{Backing} = 1 \rangle$$

where loc represents the symptom that there is a lack of c, and d represents that the agent gets disease d. Note that the ground for the symptom, the test, is not represented explicitly in PUM. It is captured by the function Ground which gives a ratio between possible values of the premise. $\text{Warrant}(loc|d) = 0.9$ and $\text{Warrant}(loc|\overline{d}) = 0.09$ specifies the correlation between loc and d, which is the warrant and embodies the qualifier "most likely". Ben completely trust the doctor and the medical study, so Backing = 1.

Acknowledgement

S. Smets acknowledges the support from the European Research Council via ERC Grant no. 283963.

References

[1] Baltag, A. and L.S. Moss and S. Solecki. The Logic of Common Knowledge, Public Announcements, and Private Suspicions. In Gilboa, I. (ed.), *Proceedings of the 7th conference on theoretical aspects of rationality and knowledge (TARK 98)*, 1998, 43–56.

[2] Baltag, A. and Smets, S. Probabilistic Dynamic Belief Revision. In van Benthem et al. (eds.), *A Meeting of the Minds: Proceedings of the Workshop on Logic, Rationality and Interaction Vol. 8*, 2007.

[3] Baltag, A. and Smets, S. A qualitative theory of dynamic interactive belief revision. *Texts in logic and games* 3(2008): 9–58.

[4] Van Benthem, J. *Logic and Argumentation* (Institute for Logic, Language and Computation, 1994).

[5] Van Benthem, J. One logician's perspective on argumentation. *Cogency* 1 (2009): 13–26.

[6] Van Benthem, J. and Gerbrandy, J. and Kooi, B. Dynamic Update with Probabilities. *Studia Logica* 93 (2009): 67–96.

[7] Toulmin, S. *The Uses of Argument* (Cambridge University Press, 2003).

How Specific Arguments Defeat General Dogmas: Lack of Parsimony in Molecular Biology

2 Dec
12:00

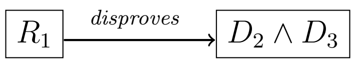
Vlasta Sikimić
University of Belgrade

Scientific argumentation, as a very fine-grained type of argumentation, is a fruitful field for analysing and modelling different argumentative strategies. For instance, formal argumentation framework for modelling explanatory and justificatory strength of rivalling theories has been proposed in [4]. Because of the variety of methods and criteria for adopting a scientific theory, we focus on the field-specific behaviour. Our general approach is a data-driven and field-specific formal analysis of scientific argumentation in molecular biology. We find the cases in molecular biology particularly suitable for such an analysis because of their precise argumentative steps, as well as their complex behaviour when it comes to reaching conclusions which violate the law of parsimony, i.e. Ockham's razor.

We discuss cases in molecular biology where three main dogmas of molecular biology were disproved. These dogmas provided simpler and more general explanations of the phenomena. However, over the years, dogmas were defeated by the explanations, which violate the law of parsimony. The first dogma (D_1) states that only DNA encodes genetic material, RNA transmits it, while only proteins have a catalytic function [2]. The second dogma (D_2) states that all infectious diseases are caused by an organism. Finally, the third dogma (D_3) says that all organism need to have intact nucleic acids. These general principles were disproved by showing that protease-resistant protein (PrP) causes Scrapie disease, RNA can have a catalytic function and that at least one organism is very resistant against the DNA destructive treatment [2].

In order to defeat the dogmas, elaborative argumentative processes that lasted several years were required. Specifically, accepting the argument that a protein can cause an infectious disease and that it can influence folding of other proteins lasted from 1967, when the first experiments were conducted and when the protein hypothesis was made, over the famous results of Prusiner in 1982 [1], until 1990 when mice were infected with the disease in a laboratory the result which persuaded the scientific community [3]. These results are considered to be cutting edge, because they disproved D_2, and Prusiner was finally awarded a Nobel Prize in 1997. Until the 1990, seven major argumentative steps were made. In opposition to such argumentative schemes, results that consist of simple isolation of bacteria or viruses in infected organisms are immediately and widely accepted with far less additional experiments required by the biological community.

When analysing the argumentation structure in the disputes against three dogmas, we made the following choices. Because of the specific nature of the research in molecular biology, experimental results are considered to be truthful unless they are deliberately faked. We abstract from the latter possibility. Thus, we consider hypotheses and results in favour or against them, where hypotheses can contradict each other, in other words can attack each other, while results are firm and truthful. Therefore, results cannot contradict each other but can disprove hypotheses.[6] This leads to an unambiguous logical behaviour. Moreover, explanatory power of different hypotheses does not play a key role in the cases considered, since they are based on the results that directly disprove them. For instance, one of the very first experiment (R_1) has show that Scrapie causing agent does not react to destruction of DNA and RNA. From this result, it follows that either D_2 or D_3 is wrong. In other words, R_1 disproves the conjunction $D_2 \wedge D_3$:



D_2: All infectious diseases are caused by an organism.
D_3: All organism need to have intact nucleic acid.
R_1: Scrapie agent does not need to have intact nucleic acid.

We will present the argumentative structure of the three mentioned arguments against the three dogmas and situate them in the theoretical argumentative framework. Ultimately, we will conclude that an argumentative behaviour where a specific argument defeats more general one is a consequence of the non-parsimonious nature of biology. The lack of parsimony in molecular biology is usually demonstrated by mutation scenarios where the simplest direct transition is frequently not accurate, but intermediate mutations are present. In the cases considered, general dogmas have been directly disproven by specific cases, demonstrating that the simplicity and generality of a hypothesis did not lead to its accuracy.

References

[1] Stanley B. Prusiner (1982). Novel Proteinaceous Infectious Particles Cause Scrapie. *Science 216:* 136-144.

[6] This feature of argumentative structure in molecular biology makes it simpler for modelling in comparison to other scientific arguments, for example ones coming from classical physics.

[2] Jan Sapp (2003). *Genesis: The Evolution of Biology.* Oxford: New York: Oxford University Press.

[3] Claudio Soto (2011). Prion hypothesis: the end of the controversy? *Trends in Biochemical Sciences 36(3):*151-158.

[4] Dunja Šešelja & Christian Straßer (2013). Abstract argumentation and explanation applied to scientific debates. *Synthese 190(12):* 2195-2217.
