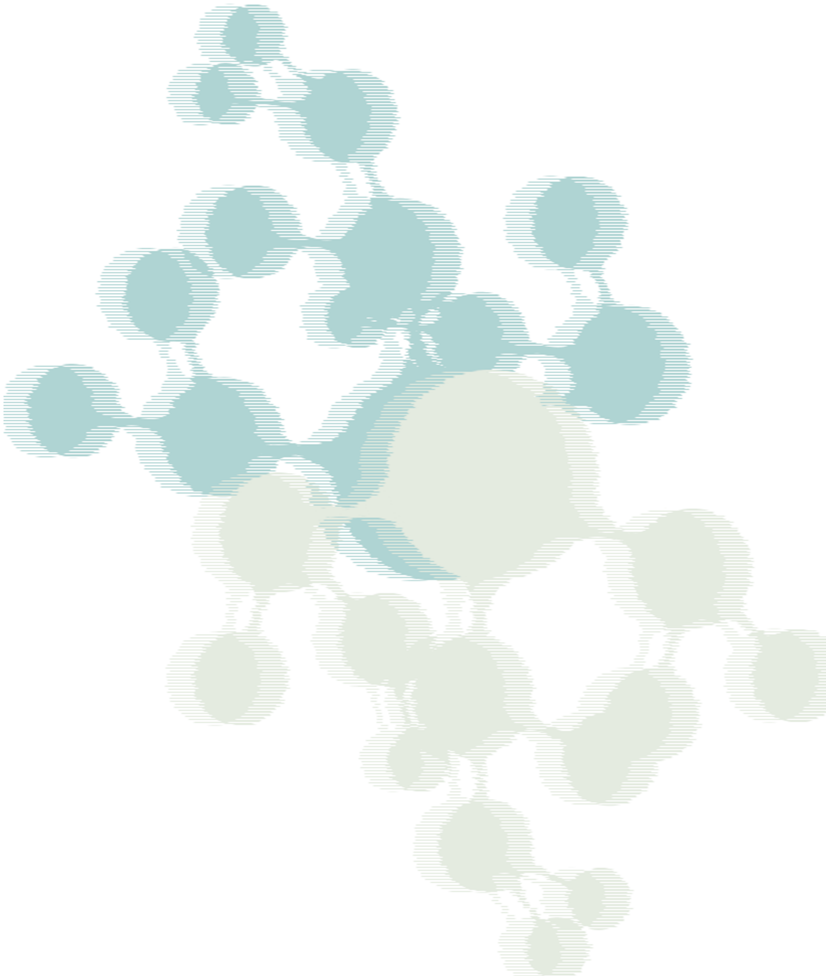
# Bridging the Gap between Formal Argumentation and Actual Human Reasoning

**Program and Abstracts**

October 4 - 5, 2018

Institute for Philosophy II

Ruhr-University Bochum



RUB

Research Group for
Non-Monotonic Logic
and Formal Argumentation

# Organization and Sponsor

The conference is organized by:

- ☺ Ofer Arieli
- ☺ AnneMarie Borg
- ☺ Marcos Cramer
- ☺ Jesse Heyninck
- ☺ Pere Pardo Ventura
- ☺ Christian Straßer

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

`https://homepage.ruhr-uni-bochum.de/defeasible-reasoning/ArgRea-2018.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*.

The conference is financed from the Sofja Kovalevskaja award and by the Ruhr University Research School PLUS (funded by Germany's Excellence Initiative [DFG GSC 98/3]).

Unterstützt von / Supported by

Alexander von Humboldt
Stiftung / Foundation

RUB
RESEARCH SCHOOL

# Conference Venue

The venue of the conference is the **Veranstaltungszentrum**, located on the university campus, underneath the Mensa. The conference will be held in room 'Saal 4'.

- ☺ **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, €2,70). 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 left and then go down the stairs. Continue walking straight until you go down more stairs and pass the Auditorium building on your left. Continue walking until you reach the "Mensa" building, again walking down some stairs, and you will finally reach an elevator that you can take (you can also walk down more stairs instead). You will go to "Ebene 04 Convention Center" by pressing the "04" button in the elevator. The room Saal 4 is on the left once you enter the convention center. (Around 7 min. walk in total.)

# Program

| | **Thursday October 4th, 2018** |
|---|---|
| 09:15 - 09:30 | Opening |
| 09:30 - 10:30 | Chris Reed<br>*Argument Technology: The First 100,000 Users* |
| 10:30 - 11:00 | Break |
| 11:00 - 11:35 | Floris Bex, Emmanuel Hadoux, Jos Hornikx and Claudia Schulz<br>*Human-Aware Computational Argumentation: A Workshop Report* |
| 11:35 - 12:10 | Trevor Bench-Capon and Katie Atkinson<br>*Taxonomising Argument Types* |
| 12:10 - 12:45 | Abdelraouf Hecham, Pierre Bisquert and Madalina Croitoru<br>*On a Flexible Representation for Defeasible Reasoning Variants* |
| 12:45 - 14:30 | Lunch |
| 14:30 - 15:30 | Adam Wyner<br>*Talking about EMIL: Extracting Meaning from Inconsistent Language* |
| 15:30 - 16:00 | Break |
| 16:00 - 16:35 | Federico Cerutti<br>*On Formal Argumentation and Scientific Enquiry* |
| 16:35 - 17:10 | Claudia Schulz, Jan Kiesewetter, Michael Sailer, Elisabeth Bauer, Martin Fischer, Frank Fischer and Iryna Gurevych<br>*The Theory of Scientific Reasoning and Argumentation in Practice* |
| 17:10 - 17:30 | Break |
| 17:30 - 18:30 | Ofer Arieli<br>*Structured Argumentation Frameworks and Reasoning with Maximal Consistency* |
| 19:30 | Conference Dinner |

| | **Friday October 5th, 2018** |
|---|---|
| 09:30 - 10:30 | Serena Villata<br>*Artificial Argumentation for Humans* |
| 10:30 - 11:00 | Break |
| 11:00 - 11:35 | Alice Toniolo<br>*Argumentation-Based Support for Human Sensemaking of Conflicting Information* |
| 11:35 - 12:10 | Trevor Bench-Capon and Katie Atkinson<br>*Structured Arguments Unchained* |
| 12:10 - 12:45 | Emmanuelle-Anna Dietz Saldanha and Antonis Kakas<br>*Cognitive Argumentation and the Suppression Task* |
| 12:45 - 14:30 | Lunch |
| 14:30 - 15:30 | Antonis Kakas<br>*Reconciling Formal and Informal Reasoning* |
| 15:30 - 16:00 | Break |
| 16:00 - 16:35 | Emil Weydert<br>*Inference Trees over a Conditional Language as Proxies for Real-World Arguments* |
| 16:35 - 17:10 | George Butler, Juliette Rouchier and Gabriella Pigozzi<br>*An Opinion Diffusion Model with Deliberation: A Tool for Democracy and Governance Analysis* |
| 17:10 | Closing |

# Abstracts of Invited Talks

## Structured Argumentation Frameworks and Reasoning with Maximal Consistency

Oct. 4
17:30

Ofer Arieli
The Academic College of Tel-Aviv

A common view in maintaining the consistency of a given set of assertions is that the main information of the set is carried by its consistent subsets and that such subsets should be as large as possible in order not to lose data. Reasoning with maximal consistent subsets of premises has gained a considerable interest since its introduction by Rescher and Manor [5]. A number of applications of this approach and its extensions have been considered for different AI-related areas, such as knowledge-base integration systems, consistency operators for belief revision, computational linguistics, and many others.

In this talk we consider several variations of this kind of reasoning, for each one we introduce two complementary computational methods that are based on structured (logical) argumentation theory:

1. a declarative method, based on Dung's semantics [3] for the underlying (structured) argumentation framework, and
2. a computational (proof-theoretical) method, based on generalized Gentzen-style sequent calculi [4], applied to the relevant framework.

The first approach is a common method to interpret argumentation frameworks by means of extensions, that is: sets of arguments that can be collectively accepted by the reasoners. The second approach is based on what we call *dynamic derivations*, which are intended for explicating actual reasoning in an argumentation framework. Unlike "standard" proof methods, the idea here is that an argument can be challenged (and possibly withdrawn) by a counter-argument, and so a certain sequent may be considered as not derived at a certain stage of the proof, even if it were considered derived in an earlier stage of the proof.

The outcome of the presentation is thus an indication of the strong link, in different settings, between reasoning with consistent sets and argumentative reasoning. A by-product of this are soundness and completeness results of dynamic proof systems with respect to several of Dung's semantics.

This is a joint work with AnneMarie Borg, Jesse Heyninck and Christian Straßer. It is mainly based on the results reported in [1] and [2].

## References

[1] O. Arieli, A. Borg, and C. Straßer. Reasoning with maximal consistency by argumentative approaches. 2018. Forthcoming in JLC.

[2] O. Arieli, A. Borg, and J. Heyninck. Structured argumentation frameworks and reasoning with maximal consistency. Submitted as a chapter in a book of the Studies in Logic series.

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

[4] G. Gentzen. Investigations into logical deduction, 1934. In German. English translation: 'The Collected Works of Gerhard Gentzen', M. E. Szabo, editor, North-Holland, 1969.

[5] N. Rescher and R. Manor. On inference from inconsistent premises. *Theory and Decision*, 1:179–217, 1970.

---

# On Formal Argumentation and Scientific Enquiry

Oct. 4
16:00

Federico Cerutti
Cardiff University

Epistemology is central to western philosophy: the pessimistic cave story of Plato as well as the optimistic of *anamnesis* are examples of it. When it comes to using computer science to support epistemology we cannot avoid to look at Leibniz and his *Caluculs Ratiocinator*—e.g. [5, p. 654]—as a precursor of several approaches aimed at creating a language for representing every piece of available knowledge and then applying logical reasoning to infer new knowledge. While extremely powerful in specific contexts, those approaches are not widely adopted in scientific enquiry due to their general lack of robustness against highly uncertain and only partially observable phenomena.

We argue in favour of a *Regulæ Philosophandi Ratiocinator* (cf. [7, p. 387]), as we first did in [1], that implements modern and widely adopted theories of epistemology. In particular, according to Popper, the advancement of scientific knowledge is based upon a process of conjecture and refutation [8], an inherently argumentative process. We base our argument on work performed in recent years [9, 2, 3] and show how existing theories of computational argumentation can already provide (limited) support for scientific enquiry in real domains. While we abstain from discussing approaches to argument mining [6], language clearly plays a role in formulation of theories. However, "although clarity is valuable in itself, exactness or precision is not [. . . ]. Words are significant only as instruments for the formulation of theories, and verbal problems are tiresome: they should be avoided at all cost" [8, p. 28]. We will therefore summarise our experience in terms of natural language interfaces to formal argumentation [4].

We show that scientific enquiry can be supported by formal argumentation, that is uniquely equipped to implement the process of conjecture and refutation discussed by Popper [8]. While we base our speculation only on—so far—anecdotal evidence, they seem convincing enough to suggest that we can soon be equipped to build a *Regulæ Philosophandi Ratiocinator*, a machinery implementing general principles of formal science.

## References


[1] Federico Cerutti. On scientific enquiry and computational argumentation. In *Proceedings of the 18th Workshop on Computational Model of Natural Argument*, 2018.

[2] Federico Cerutti, Timothy J. Norman, Alice Toniolo, and Stuart E. Middleton. Cispaces.org: from fact extraction to report generation. In *Proceedings of COMMA 2018*, 2018.

[3] Federico Cerutti and Gavin Pearson. Supporting scientific enquiry with uncertain sources. In *Proocedings of FUSION 2018*, 2018.

[4] Federico Cerutti, Nava Tintarev, and Nir Oren. Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation. In *ECAI 2014*, pages 207–212, 2014.

[5] Gottfried Wilhelm Leibniz. *Philosophical Papers and Letters*. Springer Netherlands, Dordrecht, 1976.

[6] Marie-Francine Moens. Argumentation mining: Where are we now, where do we want to be and how do we get there? In *FIRE '13*, 2013.

[7] Isaac Newton. *Philosophiæ naturalis principia mathematica*. Apud G. & J. Innys, 3rd edition, 1726.

[8] Karl R. Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, 5th edition, 1989.

[9] Alice Toniolo, Timothy J. Norman, Anthony Etuk, Federico Cerutti, Robin Wentao Ouyang, Mani Srivastava, Nir Oren, Timothy Dropps, John A. Allen, and Paul Sullivan. Agent Support to Reasoning with Different Types of Evidence in Intelligence Analysis. In *AAMAS 2015*, pages 781—-789, 2015.

# Reconciling Formal and Informal Reasoning[1]

Oct. 5
14:30

Antonis Kakas
University of Cyprus

For decades formal and informal reasoning were considered as intrinsically different processes of human thought. While the first one governs our strict mathematical or scientific reasoning the other relates to the common sense reasoning that humans carry out at large in their everyday life. Due to this quite different "application arena" our epistemological or philosophical studies tend to given them a separate identity. In addition, one school of thought had considered the informal human reasoning as an inferior form of reasoning exactly because of its significant difference with formal classical reasoning. This school would go as far as advocating that humans ought to learn to reason closer to that captured by formal logic based on the supposition that this would have a beneficial effect.

Yet, if we look at the origins of the study of logic, far back with Aristotle, we can see that in his books, collectively called "Organon", informal and formal reasoning co-exist and the latter constitutes a continuation of the former. Aristotle starts with the study of Argumentation and its dialectic process, which is akin to informal reasoning and later moves to Syllogistic reasoning, which forms the basis of modern formal logic.

Recent work on the computational argumentation in Artificial Intelligence (AI) has shown that indeed formal and informal reasoning can be unified under the umbrella of dialectic argumentative reasoning. Both types of reasoning are forms of argumentation, with formal reasoning on one end of the spectrum, where arguments and their relative evaluation is carried one in a rigid and strict way, whereas informal reasoning is carried out in a structurally equivalent framework of argumentation, but where we admit a high degree of flexibility in the way that this is applied.

In other words, argumentation is a primary and foundational notion on which we can uniformly build all forms of reasoning. We call this Argumentation Logic. This central position of argumentation is also strongly supported by work in Cognitive Psychology where direct evidence is given that argumentation is "native" to human reasoning. Argumentation Logic, therefore reconciles formal and informal reasoning and offers the possibility to sufficiently formalize the human forms of common sense reasoning and decision making into a logical system that can possess similar cognitive or thinking faculties that are common in the natural intelligence of people. Such a framework, which we call Cognitive Argumentation, would then allow us to develop human-like systems which can have a symbiotic relationship with their human users.

---

[1]This work is based on a long standing collaboration with P. Mancarella and F. Toni, and recently with E. Dietz Saldanha.

## Argument Technology: The First 100,000 Users

Oct. 4
9:30

Chris Reed
University of Dundee

The establishment and growth of computational models of argument over the past 25 years has been a uniquely fruitful collaboration crossing disciplinary boundaries from humanities, through social and computational sciences, to engineering. More recently, argument mining in particular has enjoyed phenomenal success, exploding from nothing to a worldwide industry in barely half a decade. Despite ingenious theoretical creativity and ground-breaking innovative collaborations, the area has, however, largely struggled to deliver products and applications to end users. Why should this be? What are the challenges and dynamics that have stymied better uptake, and what can be done to improve this translational aspect? This talk draws on the experiences from the Centre for Argument Technology in the development of philosophical and linguistic theory, in the engineering of integrative and interoperable systems, and in working with a variety of partners. The goal is to unpack some of the connections between formal and human argumentation, and to explore the challenges and opportunities laying ahead of the field of argument technology.

---

## Argumentation-Based Support for Human Sensemaking of Conflicting Information

Oct. 5
11:00

Alice Toniolo
The University of St.-Andrews

Models of argumentation have increasingly been employed in human decision-making systems to facilitate good reasoning. Sensemaking of conflicting and incomplete information is one application where argumentation-based tools have the potential to help users reduce the cognitive load in identifying hypotheses about a situation. To improve the effectiveness of systems that employ computational models of argumentation, however, there is a real need to evaluate their use in human decision support. In this talk, we seek to better understand the link between human reasoning, argumentation schemes and preferred extensions in supporting sensemaking of conflicting information. An application will be presented in the context of intelligence analysis which employs argumentation schemes to construct hypotheses about the world and counteract cognitive biases. Preferred extensions are linked to different possible world explanations and help analysts reduce the cognitive effort in identifying what is coherent in a situation. However, using argumentation-based tools to support reasoning about the world opens questions on how people engage with and understand these approaches. A set of experiments with human participants is presented to investigate the use of argumentation schemes and preferred extensions in identifying plausible explanations. Initial results show that argumentation schemes are a reliable method to structure inferences and draw plausible conclusions from incomplete information with potential for supporting the identification of biases. On the other hand, preferred extensions can be seen as capturing different possible world explanations affecting the degree of believability of a conclusion. Results from the experiments show that the degree of believability of a conclusion may be associated with the number of preferred extensions in which the conclusion is credulously accepted with similar heuristics as those employed in understanding probabilities.

---

# Artificial Argumentation for Humans

Oct. 5
9:30

Serena Villata
Université Côte d'Azur, CNRS, Inria, I3S, France

Since the early years of the field, Artificial Intelligence has the goal to understand the principles governing intelligent behavior and to encode such principles in so-called *intelligent machines*. In the latest years, progress in AI seems to be accelerating, e.g., given the recent results in Machine Learning, Natural Language Processing (NLP) and Knowledge Engineering, leading to important investments in AI also from main information technology companies. Alas, all that glitters is not gold, and together with the increasing popularity of AI and the expectations on it, new concerns are also rising around the development of intelligent machines. We are at a crossroads: on the one hand, AI can be enormously beneficial for human flourishing, but on the other hand, we need to take care about the design of AI machines in order to reach a so-called *good AI hybrid society* [4, 3]. In this society, intelligent machines have the capability to form teams with humans.

Argumentation is the process by which arguments are constructed, compared, evaluated in some respect and judged in order to establish whether any of them are warranted. The idea of *argumentation* as the process of creating arguments for and against competing claims, was a subject of interest to philosophers and lawyers. In recent years, however, there has been a growth of interest in the subject from formal and technical perspectives in Computer Science, and a wide use of argumentation technologies in practical applications [1]. The field of artificial argumentation plays an important role in Artificial Intelligence research. The reason for this is based on the recognition that if we are to develop robust intelligent machines able to act in mixed human-machine teams, then it is imperative that they can handle incomplete and inconsistent information in a way that somehow emulates the way humans tackle such a complex task.

My research focused on different problems that I believe stand in the way of reaching this ambitious goal. I started from the observation that, in their deliberation process, humans use argumentation either internally, by evaluating arguments and counterarguments, or externally, by entering into a debate where arguments are exchanged. The three pillars of the development of argumentation-enhanced intelligent machines are, from my point of view: *(i)* modeling and reasoning on socio-cognitive components like trust using computational models of argument which are able to deal with incomplete and conflicting information, *(ii)* mining argument structures in natural language text [2] to detect, e.g., potential fallacies, recurrent patters, and inner strength, and *(iii)* analyzing and understating the role of emotions in real world argumentative situations (e.g., debates) to inject such information in the computational models of argument to better cast incomplete and inconsistent information when emotions play a role.

Argumentation-enhanced intelligent machines passes through the use of argumentation technologies to support the transparency of the deliberation process (*why* the machine deliberated in a certain way), and to support the extraction and reasoning on argumentation structures from different settings (e.g., clinical trials, social media posts, political debates) which, being generated by humans, require a high capability to deal with incompleteness and inconsistency. Humans argue. Machines should be able to argue too if we aim to achieve mixed teams in a hybrid society.

## References

[1] Katie Atkinson, Pietro Baroni, Massimiliano Giacomin, Anthony Hunter, Henry Prakken, Chris Reed, Guillermo R. Simari, Matthias Thimm, and Serena Villata. Towards Artificial Argumentation. *AI magazine*, 38(3), 2017.

[2] Elena Cabrio and Serena Villata. Five years of argument mining: a data-driven analysis. In Jérôme Lang, editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden.*, pages 5427–5433. ijcai.org, 2018.

[3] Corinne Cath, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo, and Luciano Floridi. Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. *Science and Engineering Ethics*, 24(2):505–528, 2018.

[4] Stuart J. Russell. Provably beneficial artificial intelligence. *Exponential Life, The Next Step*, 2017.

---

# Talking about EMIL: Extracting Meaning from Inconsistent Language

Oct. 4
14:30

Adam Wyner
Swansea University

Formal and computational theories of argumentation along with argument mining are leading towards automated processing and reasoning with inconsistent, linguistically expressed knowledge. However, there is a gap between the coarse-grained (propositional logic) knowledge-bases of computational theories (perhaps derived via argument mining) and the requirements of inference from knowledge-bases expressed in natural language, which are fine-grained (predicate logic). To draw theory, implementation, and natural language closer together, we couple the development of a computational foundation for argumentation with an adaptation of a controlled natural language system. For the computational foundation, we provide a direct semantics. For the natural language processing, we adapt an existing controlled natural language system with the expression 'it is usual that'. Overall, we can input arguments expressed in a controlled natural language, translate them to a fine-grained, formal knowledge base, represent the knowledge in a rule language, reason with the rules, generate argument extensions, and finally reconvert the arguments in the extensions into natural language. Methodologically, we take an 'engineering' approach to fine-grained argument analysis.

---

# Abstracts of Contributed Talks

## Taxonomising Argument Types

Oct. 4
11:35

Trevor Bench-Capon and Katie Atkinson
Department of Computer Science, University of Liverpool

Abstract argumentation frameworks were introduced by Dung [2] to explore the interactions between arguments viewed at the most abstract level. He identified a single relation between arguments, "attack", the idea being that an argument would be defeated by its attacker, unless that attacker was itself defeated. Subsequently preference and value based frameworks distinguished between *attack* and *defeat*, so that it was possible to resist an attack if the attacked argument was preferred [4]. Other researchers have been interested in support as well as attack. For some, support for an argument is a second relation between abstract arguments, but others define it in terms of arguments which defeat its attackers. Others have explored structured arguments, in which an argument has several components, including (at least) a conclusion and premises [5], and support is expressed in terms of these components.

What is relatively unexplored is a degree of abstraction between fully structured and fully abstract argumentation. At this level several types of arguments and attacks can be identified. While types of argument and restrictions on the ways in which they can be attacked have been mentioned, there has been little systematic exploration of the roles, relations and effects of such types. We distinguish between practical arguments, where the conclusion suggests an action to be performed, and theoretical arguments, which suggest that a certain statement is true. Key here is the "direction of fit" [7]: in theoretical reasoning beliefs are made to fit the world, but in practical reasoning the world is changed so as to fit what is desired. Both kinds of arguments may be strict, valid without exceptions, or defeasible, normally (or typically, or presumptively) valid. The strict/defeasible distinction has rarely been made explicitly for practical arguments. In some planning systems they appear to be assumed to be strict when forming the plan and then performance is monitored to see whether replanning is needed, although others treat them as defeasible [1] during plan formation.

For attacks we start from the well-known distinction between *rebuttals*, arguments for a contrary of the conclusion; *underminers*, arguments that a premise does not hold; and *undercutters*, which challenge the applicability of the inference rule used [5]. For practical arguments we further distinguish between standard rebuttals, which argue that an action should not be performed, and *alternatives*, which argue that a different, incompatible, action should be performed. We also add a *counter example* attack, which is used to show that an argument is not strict.

To facilitate integration of practical and theoretical reasoning we give a common semantics in terms of Action Based Alternating Transition Systems (AATS) [8]. AATS are a variety of state transition diagram in which each transition corresponds to a 'joint' action which is formed by every agent performing some (independently chosen) action. AATS have been used to supply a semantical structure for practical argumentation [1], which additionally requires each transition to be labelled with the social values promoted and demoted by the transition. For theoretical reasoning, the states of the AATS can be regarded as possible worlds, with the initial state as the actual world. We can then provide the conditions to instantiate all arguments available in a given world (or set of worlds) and the attacks on them in terms of an AATS. While seventeen ways of attacking practical arguments were defined in AATS terms in [1], these were not been characterised as rebuttals, undercutters and underminers.

We now examine which types of argument are subject to which kinds of attacks. While it has been noted that strict theoretical arguments can neither be rebutted nor undercut, there is no explicit discussion of other arguments types: for example, the attacks that can be used against strict practical arguments. We also consider properties of the different kinds of attack: for example rebuttals always give rise to a mutual attack between the arguments concerned, whereas undercutters are always uni-directional.

It is also important to analyse how attacks are resolved: often there is an appeal to preferences. But whereas resolution of practical rebuttals is in terms of subjective choice, often expressed in terms of the social values promoted by the arguments [1], the choice for theoretical arguments is more constrained (cf. [6]). For theoretical arguments it will be necessary to consider non-monotonic logics, such as circumscription [3], and probabilities. Choice can arise if the degree of risk an agent is willing to accept needs to be considered. These considerations may give rise to additional kinds of arguments, such as preference arguments, and possible extensions to the set of propositions constituting a state (e.g. to represent *ab* predicates for circumscription), or even the AATS (e.g. to explicitly represent audiences).

A key contribution of this work is to propose a means of integrating strict and practical reasoning in a principled fashion by basing both on the AATS structure. There may well be computational implications of this analysis, but here our focus is on strategies for attacking an argument: given that attack options are limited, a particular kind of argument can only be attacked in particular ways, and often some kinds of attack will be preferred over others. Also some attacks need to be used in combination. Thus to attack a strict theoretical argument in the absence of underminers, one will first find a counter example, and then need an undercutter or rebuttal to attack the resulting defeasible argument. In the case of a rebuttal, an argument to prefer the attacker will also be needed. Understanding of these strategies will improve the naturalness of computational dialogues.

# References

[1] K. Atkinson and T. Bench-Capon. Practical reasoning as presumptive argumentation using Action based Alternating Transition Systems. *Artificial Intelligence*, 171(10-15):855–874, 2007.

[2] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial in-*

*telligence*, 77(2):321–357, 1995.

[3] J. McCarthy. Circumscription—a form of non-monotonic reasoning. *Art. Int.*, 13:27–39, 1980.

[4] S. Modgil and T. Bench-Capon. Metalevel argumentation. *Journal of Logic and Computation*, 21(6):959–1003, 2011.

[5] S. Modgil and H. Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument and Computation*, 5(1):31–62, 2014.

[6] H. Prakken. Combining sceptical epistemic reasoning with credulous practical reasoning. In *Computational Models of Argument: Proceedings of COMMA 2006*, pages 311–322, 2006.

[7] J. R. Searle. *Rationality in action.* MIT press, 2003.

[8] M. Wooldridge and W. Van Der Hoek. On obligations and normative ability: Towards a logical analysis of the social contract. *Journal of Applied Logic*, 3(3-4):396–420, 2005.

---

# Structured Arguments Unchained

Oct. 5
11:35

Trevor Bench-Capon and Katie Atkinson

Department of Computer Science, University of Liverpool

Within computational argument, structured argumentation currently relies heavily on *modus ponens*. For example, ASPIC+ chains arguments together to form inference trees [5], Definition 3.6. Arguments are structured into premises and a conclusion, with the conclusion being acceptable if the premises are acceptable. The base case is that the conclusion is in the knowledge base. Otherwise, arguments are chained together, supporting premises by showing that they are themselves the conclusion of a sub-argument. This corresponds to a standard AI formalism: arguments are derived from a standard knowledge base, comprising rules (strict and defeasible in [5]) and facts (*axioms* and *premises* in [5]), with the arguments corresponding to applications of rules. Where the premises are themselves derived as consequences of rules, they form sub-arguments, giving rise to a tree of arguments, corresponding to chaining the rules.

Attacks in ASPIC+ may be *rebuttals* (where the arguments have contradictory conclusions); *underminers* (where the conclusion of the attacker is the negation of a premise of the other argument) and *undercutters* (where the conclusion of one argument is that the rule used in the other is not applicable).

Argument schemes offer a similar pattern. The schemes given in [7] take the form a set of premises, a defeasible conclusion, and a set of *critical questions* which can be used to challenge the applicability of the scheme, or the truth of the premises or the conclusion in the particular context. Attacks on instantiations of argument schemes are based on these *critical questions*. Critical questions can give rise to any of the three kinds of attack [6].

Premises, Rule and Conclusion is indeed a common form of argumentation, but actual argumentation can take other forms. This is apparent from Natural Deduction systems (such as [4]) in which *modus ponens* is only one of many rules.

- A common way of arguing is *reductio ad absurdum*. Here we need a pair of arguments with at least one premise in common and with the conclusion of one argument the negation of the other. This enables us to conclude that (at least) one of the premises is false. Further arguments may be needed to establish which of the premises is to be considered false.

- A second example is argument from cases. Here we offer a disjunction which exhausts all possibilities, and then a set of arguments showing that the conclusion can be shown from each of the disjuncts.

- We should also consider *modus tolens*. This does not invariably apply to defasible arguments, and in ASPIC+ is handled by explicitly adding the contrapositions to the knowledge base [2] where it is desired, especially for strict rules. It may, however, be useful to see it as a separate kind of argument.

- Another rule of Natural Deduction is *Conditional Proof*, in which an argument from an assumption to a conclusion is taken to establish the rule that *if assumption, then conclusion*. This could be used to attack an undercutter.

We suggest therefore that we need to extend the definition of argument in ASPIC+ (and similar frameworks) to cater for arguments with these very different structures. That this matters can be seen from the following example. Suppose we have a knowledge base comprising two rules:

> **sr1**: $p \wedge q \to r$. **sr2**: $p \wedge s \to \neg r$ and $p, q$ and $s$ as premises.

Now, using the ASPIC+ definition, we get two arguments:

> **A**: premises $\{p, q\}$ and conclusion $r$ and **B**: premises $\{p, s\}$ and conclusion $\neg r$

ASPIC+ yields two arguments in a relationship of rebuttal. Whether we accept $r$ or $\neg r$ will depend on whether $A$ is preferred to $B$ or *vice versa*. But the more normal response to $A$ and $B$ is that they show that the common premise $p$ cannot be true. So we at least need the option of denying $p$ and seeing the *reductio* argument ($C$) as undermining both $A$ and $B$. Had $p$ not been a premise, but established by a fourth argument $D$, then $C$ would have been a rebuttal of $D$. Thus ignoring the possibility of arguments such as $C$ can lead to the wrong conclusions in some cases.

These additional forms of argument greatly expand the possible strategic considerations in various kinds of dialogue, perhaps especially persuasion dialogues. Moreover, by including these forms of argument which are closely related to rules in Natural Deduction systems e.g. [4], we bring computational argument closer to informal argumentation.

Some of these diferent kinds of arguments have been noted before: [3] discussed *reductio* and [1] reasoning by cases. These, however, treated the structures separately, and the proposed solutions are directed towards the specific case under consideration, and so they do not address the whole problem generally. What is required is a thorough, uniform, treatment, able to be integrated with the existing ASPIC+ framework, of arguments which have arguments as premises, and arguments with other structures (such as *conditional proof*, which has an argument as conclusion), and how these interact with each other and existing argument/subargument chains.

# References


[1] Mathieu Beirlaen, Jesse Heyninck, and Christian Straßer. Reasoning by cases in structured argumentation. In *Proceedings of the Symposium on Applied Computing*, pages 989–994. ACM, 2017.

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

[3] Jérémie Dauphin and Marcos Cramer. ASPIC-END: Structured argumentation with explanations and natural deduction. In *Theory and Applications of Formal Argumentation*, pages 51–66. Springer, 2017.

[4] Edward John Lemmon. *Beginning logic*. CRC Press, 1971.

[5] Sanjay Modgil and Henry Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument and Computation*, 5(1):31–62, 2014.

[6] Bart Verheij. Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial intelligence and Law*, 11(2-3):167–195, 2003.

[7] Douglas Walton, Christopher Reed, and Macagno Fabrizio. *Argumentation schemes*. Springer, 2008.

# Human-Aware Computational Argumentation: A Workshop Report

Oct. 4
11:00

Floris Bex[1], Emmanuel Hadoux[2], Jos Hornikx[3] and Claudia Schulz[4]

[1]Department of Information and Computing Sciences, Utrecht University & Tilburg Institute for Law, Technology and Society, Tilburg University, The Netherlands

[2]Department of Computer Science, University College London, UK

[3]Radboud University, Nijmegen, The Netherlands

[4]Ubiquitous Knowledge Processing (UKP) Lab, TU Darmstadt, Germany

Argumentation is an omnipresent method of human communication. Politicians argue for their election manifestos, colleagues argue about the best way of solving a task, and we even argue with ourselves before making an important decision. But what exactly is an argument, what does it mean that two arguments are in conflict, and how can we determine who wins a debate? In the past 20 years, these questions have been investigated from a computational point of view within the field of Artificial Intelligence. Numerous theories of computational argumentation have been proposed, which formalise, for example, how arguments may be constructed from underlying knowledge, how contradictory information in arguments may lead to conflicts between arguments, and which sets of arguments may be deemed acceptable in a debate with conflicting arguments. Amongst others, the theories have proven useful for aiding decision making, for robot communication, and for providing human understandable explanations of algorithm solutions. Even though the theories of computational argumentation provide sound theoretical systems, there is only little work on whether or not they indeed encode concepts found in human argumentation.

In order to tackle this question, we have organised a workshop that occurred in the Lorentz centre in Leiden, The Netherlands, the 14$^{\text{th}}$-18$^{\text{th}}$ of May 2018. We have invited researchers from different disciplines: Computer Science, Psychology and Philosophy with a different take on the definition of argumentation. This workshop aimed at bridging the gap between computational and human aspects of argumentation by exploiting the strength of each discipline represented. The goal was to understand the overlap of computational and human argumentation in more detail, to form interdisciplinary collaborations, and to formulate cross-discipline research questions that will advance the study of human-aware argumentation in upcoming years.

When organising such a workshop, the first difficulty to overcome is the use of the same terminology for different concepts. Therefore, to create a common ground for all participants, the first day of the workshop featured talks reviewing argumentation research in four different fields: logic-based artificial intelligence, psychology, communication studies, and technology. They each gave a different insight on the definition of argument to the research of the other fields. During the three following days, speakers presented specific aspects of argumentation in their respective domain in the mornings, leading to very productive working group sessions in the afternoons. The success of these group sessions was partly based on carefully crafted interdisciplinary group allocations, taking into account the diverse strengths and expertise of each participant. During the working sessions, each group worked on a concrete task designed with the objective of initiating a momentum leading up to fruitful plenary discussions of each group's achievements. Concretely, the outcomes of the group sessions were:

1. a better understanding of the overlaps between the different fields studying argumentation represented at the workshop,
2. a paper draft for the fictive journal Empirical Studies in Argumentation, investigating research questions bridging the gap between human and computational argumentation,
3. a SWOT[2] analysis of an existing or newly designed argument technology,
4. a concrete plan for the organisation of a subsequent workshop aimed at fuelling newly formed collaborations and ideas.

These outcomes gave birth to concrete future steps needed to grow the new research area of human-aware computational argumentation. Participants from different fields were excited to work together on achieving these steps. Specifically, the first future step is to use the same dataset, discussions gathered from the Change My View subreddit[3], with each participant own method and objective. The aim is to find complementary methods. For instance, argument mining analysis may contribute to strategic dialogues analysis down the process line. In addition, two follow-up workshops will be organised, one with the objective of joint grant proposal writing and reviewing, the other aiming at drafting joint research papers.

---

# An Opinion Diffusion Model with Deliberation: A Tool for Democracy and Governance Analysis

Oct. 5
16:35

George Butler, Juliette Rouchier and Gabriella Pigozzi
Université Paris-Dauphie-LAMSADE

Collective decision-making, democracy and persuasion go hand in hand. In the current Western world, democracy is uttered by politicians so often that one would believe it to be a religion, as if it was a normative requirement for collective action and policy choices. If so were true, then this requirement should be accompanied with criteria and tools for evaluating decisional outcomes and methods of aggregation, and procedures of deliberation that guarantee democracy through their fair and unequivocal application (Peter, 2007). In effect, holding to the requirement should prove blind to the object to which it is applied and provide assurance in the grasping of some kind of legitimate "truth" - be it pragmatic, consensual or epistemic, despite that these considerations may be philosophically or semantically opposed as they are procedural.

Indeed, one tends to generalize the concept of democracy and ignore the distinction between the procedural and consequential goals that entail its application.

[2] https://en.wikipedia.org/wiki/SWOT_analysis
[3] https://www.reddit.com/r/changemyview/

For example, vote-centric democratic theory (e.g. Condorcet, de Borda), a liberal-individualistic way of understanding democratic procedures, views democracy as an arena where fixed preferences, opinions and interests compete via seemingly fair aggregation operators. Communication-centric democratic theory, in contrast, the type of democracy we are interested on in this paper, focuses on the democratic elements that encompass the communicative processes of opinion and will formation that precede voting, and in particular, on the concepts of Action-Arena, argumentation and persuasion through reason. Deliberative democratic theory studies democracy through this scope and provides some empirical evidence of the effects of deliberation on individual and public opinion and thus, on collective decision-making (Chambers, 2003; Fishkin & Luskin, 2005). Even so, in the literature of opinion diffusion, excepting (Gabbriellini & Torroni, 2013), deliberation and rational discussion are very frequently abstracted away and considered one with pair-wise opinion-based interactions among individual agents (Deffuant et al., 2002; Jager & Amblard, 2005; Meadows & Cliff 2014).

In this paper, we bring forward an agent-based model that combines pair-wise interaction among individuals and collective deliberation. It considers public decision-making processes and policy decisions as its main objects of interest in that such decisions are legitimate because they have gone through a process of fair public scrutiny in accordance to pure epistemic proceduralism (Peter, 2007). Notwithstanding, we also embrace an epistemic approach to democracy for we show interest in the epistemic quality and outcome of the collective decisions agents make under simple procedural constraints. We regard the criteria of success for an instance of deliberation as being partly internal and partly external to the decision making process. Success relies, in the first place, on consequences - on how well these processes help agents make ”correct” and ”rational” collective decisions - in second place, on the fairness of the decision-making process (Anderson, 2006). Further, we partake the question of legitimacy of public policy and collective decision-making through the scope of Policy Analytics - in a formal way, and thought through models (Tsoukias et al. 2013). Our model aims to display to what extent simple deliberation policy design, intrinsic for policy acceptance, can have an effect on observables that shape policy decisions which, in turn, may shape policy cycles.

Abstract argumentation in the system we present here allows a relatively intuitive, yet expressive modeling of deliberation processes. It helps reduce the complex problem of acceptance of arguments and policy, since policies here come as arguments, to solving computational problems in graphs through the use of semantics. Moreover, it demands no particular knowledge of the structure of the arguments and the way they relate with each other, and the way individuals think about a policy problem. Semantics, in their stead, model different criteria of acceptability that may depend on the nature of the problems policies responds to. Here, policies can only be accepted if they survive the process of scrutiny - if they are accepted for a given semantics.

At its actual stage and in technical terms, the model we present strives to create a sustainable ontological link between abstract argumentation and opinion diffusion models. In defining such link, we aim to describe a realistic process of collective decision-making coherent with Western ideals of governance, where deliberation is a necessary condition for democracy and policy choices. We draw inspiration from Tsoukias' Policy Analytics approach, Habermas' Theory of Truth and Moscovici and Doise's results on how consensus within groups is reached to define legitimacy as a result of public decision-making, and characterize changes in paradigm. Moreover, we show interest in how deliberation reformulates stylized facts in classic opinion diffusion models and on how governance, or as we put it, constraints on deliberation procedures, affect the overall aggregated individual opinion and democracy through the scope of epistemic democracy. For this, we characterize the effect of the "strong" public space in opinion dynamics and on the process of collective decision through argumentative and opinion-based voting schemes. Ultimately, we seek to create a tool that helps policy-makers analyze deliberation protocols and make studied decisions about them, whenever doing so is an important issue (e.g costly, on emergency situations).

At this point in the modeling, we show by using linear regressions that deliberation has a significant overall impact on the distribution of opinions and on the diversity of these. Further, we bring up evidence on the fact that these depend heavily on the strength and probability of occurrence of the argumentation dynamics and, in the same line, show that effects of deliberation parameters have a significant impact on opinion distributions given that agents are disengaged and pair-wise interactions are not very frequent, yet not when the system is less complex (e.g. voting in accepting deliberated results). We find that, with respect to classical opinion diffusion models, deliberation allows for higher variance of opinion in populations, whilst producing lower opinion-shifting statistics.

We shed light on the fact that collective rationality and epistemic accuracy, society's ability to accept winning deliberated motions and to correctly judge propositions, respectively, depend on how institutions organize debates. The number of debates allowed, as well as how many agents may participate in them, seldom constrains epistemic accuracy, whereas collective rationality is guaranteed by the possibility of voting for deliberated motions. In a similar manner, increasing the proportional requirements in voting for accepting motions, throws the model into a path similar to that of the classical opinion-diffusion models. Last but not least, we point out how important framing is in public opinion by showing that opinion dynamics are conditioned by the topics discussed and to how agents argue.

---

# Cognitive Argumentation and the Suppression Task

Oct. 5
12:10

Emmanuelle-Anna Dietz Saldanha[1] and Antonis Kakas[2]

[1]International Center for Computational Logic, TU Dresden

[2]Department of Computer Science, University of Cyprus

One of the original aims of Artificial Intelligence was to formalize human and commonsense reasoning [5, 6]: To develop what today we call human-like AI. In order to address this challenge it is natural to appeal to logic and use or adapt the main logical formalism of Classical Logic (CL) that successfully underlies our scientific reasoning. However, various psychological experiments have shown that humans do not reason according to CL and some other logical formalism is needed. Although several alternatives have been proposed, such as non-monotonic or many-valued logics, the emphasis had shifted away from their application to real case studies of human reasoning. On the other hand, cognitive scientists have investigated human thinking extensively and have empirically identified many characteristics exposed in human reasoning, such as human biases and presuppositions implied by natural language.

Consider for instance the so-called *suppression task* [3] a well-known psychological study on reasoning where the experiment was carried out as follows: Three groups of participants were asked to derive conclusions given variations of a set of premises. Group I was given the following two premises:[4]

$$\textit{If she has an essay to finish, then she will study late in the library.} \qquad (e \rightarrow \ell)$$
$$\textit{She has an essay to finish.} \qquad (e)$$

The participants were asked what necessarily had to follow assuming that the above two premises were true. They could choose between the following three answer possibilities:

$$\textit{She will study late in the library.} \qquad (\ell)$$
$$\textit{She will not study late in the library.} \qquad (\neg\ell)$$
$$\textit{She may or may not study late in the library.} \qquad (\ell \vee \neg\ell)$$

96% of the participants concluded that *She will study late in the library.* Next to the two premises that were received by group I, group II was given additionally the following premise:

$$\textit{If she has a textbook to read, then she will study late in the library.} \qquad (t \rightarrow \ell)$$

Still, 96% of the participants concluded that *She will study late in the library.* Finally, group III received, next to the two premises that were given to group I, additionally the following premise:

$$\textit{If the libary stays open, then she will study late in the library.} \qquad (o \rightarrow \ell)$$

In this group only 38% concluded that *She will study late in the library.* The results of this experiment show that previously drawn conclusions seem to be suppressed given additional information, i.e. participants seemed to reason non-monotonically. A natural explanation why participants in group III did not conclude that *She will study late in the library*, is because they were not sure whether *The library stays open*, which is a necessary requirement for her to study late in the library. In the first two groups the majority of the participants did not have this doubt, as they have not been made *aware* of the possibility that the library may not be open and thus this option was not relevant for them.

Consider again group I and assume that $e$, $\ell$, $\neg\ell$ and $e \rightarrow \ell$ are the only available premises, premises from which we can construct arguments. Assume further that $e$

[4]The participants received only the natural language sentences and not the abbreviated notation on the right hand side.

and $e \to \ell$ are premises that are preferred over $\ell$ and $\neg\ell$ because they are explicitly stated to be true. We can construct an argument $\Delta^e_{e\to\ell}$ consisting of $e$ and $e \to \ell$, which supports $\ell$. The only argument $\Delta_{\bar{\ell}}$ that we can construct for $\neg\ell$ consists of $\neg\ell$ itself. According to the specified preference among the premises, $\Delta^e_{e\to\ell}$ defends against $\Delta_{\bar{\ell}}$, but not vice versa. Thus there is an acceptable argument for $\ell$ but not for $\neg\ell$. For group II we additionally have the premises $t \to \ell$, $t$ and $\neg t$. Yet we can construct another argument, $\Delta_{t,t\to\ell}$, for $\ell$ consisting of $t$ and $t \to \ell$. As $\Delta^e_{e\to\ell}$ is still preferred over $\Delta_{\bar{\ell}}$, the result stays the same. In both cases, $\ell$, the premise supported by the acceptable arguments, corresponds to the majority's conclusion, that *She will study late in the library*.

For group III, the case seems to be understood differently: The participants might have understood the conditional statement $o \to \ell$ as *If the library does not stay open, then she will not study late in the library* ($\neg o \to \neg\ell$). This interpretation of the statement together with the possibility that *The library does not stay open* ($\neg o$), supports the conclusion that *She will not study late in the library*. Furthermore, it seems that this argument is at least as preferred as $\Delta^e_{e\to\ell}$.

We will define an argumentation framework suitably adapted for the task of capturing logical reasoning as an instance of preference-based argumentation [8, 7, 1, 4, 2]. We will then present a formalization of all three cases of the suppression task discussed here, together with the other three cases that include negation (i.e. where the second premise instead is *She does not have an essay to finish*). Finally, we will evaluate whether the premises supported by the acceptable arguments in the formal framework correspond to the majority's conclusions.

# References


[1] Leila Amgoud, Yannis Dimopoulos, and Pavlos Moraitis. Making decisions through preference-based argumentation. In Gerhard Brewka and Jérôme Lang, editors, *KR*, pages 113–123. AAAI Press, 2008.

[2] Leila Amgoud and Srdjan Vesic. A new approach for preference-based argumentation frameworks. *Ann. Math. Artif. Intell.*, 63(2):149–183, 2011.

[3] Ruth M. J. Byrne. Suppressing valid inferences with conditionals. 31:61–83, 1989.

[4] Antonis C. Kakas and Pavlos Moraitis. Argumentation based decision making for autonomous agents. In *Proc. of 2nd Int. Joint Conf. on Autonomous Agents & Multiagent Systems, AAMAS*, pages 883–890. ACM, 2003.

[5] John McCarthy. Programs with common sense. In *Proceedings of Teddington Conference on Mechanization of Thought Processes*, pages 75–91, London, 1959. Her Majesty's Stationary Office.

[6] John McCarthy. Elaboration tolerance. In *4th International Symposium on Logical Formalizations of Commonsense Reasoning*, 1998.

[7] Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.

[8] Henry Prakken and Giovanni Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 7(1):25–75, 1997.

---

# On a Flexible Representation for Defeasible Reasoning Variants

Oct. 4
12:10

Abdelraouf Hecham[1], Pierre Bisquert[2], Madalina Croitoru[1]
[1]University of Montpellier
[2]IATE, INRA

Defeasible reasoning [9] is used to evaluate claims or statements in an inconsistent setting. It has been successfully applied in many settings when reasoning in presence of inconsistency is necessary. We focus here, as per the EU H2020 Glopack project, on the task of reasoning from different inconsistent points of views. An inherent characteristic of defeasible reasoning is its systematic reliance on a set of intuitions and rules of thumb, which have been long debated between logicians [7]. For example, could an information derived from a contested claim be used to contest another claim (i.e. ambiguity handling)? Could “chains” of reasoning for the same claim be combined to defend against challenging statements (i.e. team defeat)? etc. The main available defeasible reasoning tools are ASPIC+ [10], DEFT [6], DeLP [4], DR-DEVICE [1], and Flora-2 [11]. Table 1 shows that no tools can support all features.

Table 1: Defeasible features supported by tools.

| Tool | Blocking | Propagating | Team Defeat | No Team Defeat |
|---|---|---|---|---|
| ASPIC+ | - | ✓ | - | ✓ |
| DEFT | - | ✓ | - | ✓ |
| DeLP | - | ✓ | ✓ | - |
| DR-DEVICE | ✓ | - | ✓ | - |
| Flora-2 | ✓ | - | ✓ | - |

We propose a new logical formalism called *Statement Graph* (SGs) that *captures all features showed in Table 1* via a flexible labelling function. The SG can be seen as a generalization of Abstract Dialectical Frameworks (ADF) [2] that enrich ADF acceptance condition.

Furthermore, the flexibility of SG labeling allows the representation of semantics beyond defeasible reasoning. We evaluate our approach with respect to human reasoning and show how Statement Graphs can be used to capture other forms of human reasoning, namely, the *suppression task*. A state of the art psychological study [3] shows that people tend to change (suppress) previously drawn conclusions when additional information becomes available even if from a logical point of view, the new information should not affect reasoning. This suppression effect has longly been studied in cognitive computer science and can be presented in different forms. We are interested in the modus-ponens suppression tasks. This study shows that, much like non-monotonic reasoning, conclusions can be suppressed in human reasoning in presence of additional information. Since humans tend to rely on background knowledge, the first step of human reasoning is reasoning towards an appropriate logical representation of the situation. A possible explanation for the modus-ponens suppression is the fact that humans consider unsupported counter-arguments as valid attacks. While such reasoning behaviour cannot be expressed in defeasible logic it could be captured by the labelling function of the SGs. We defined a labelling function that is able to represent the suppression effect under the plain logical representation of the situation presented to participants and tested its computational efficiency. The "**modus-ponens suppression task**" is explained in Ex. 1.

**Example 1** *Consider the following situation 1 [3]:*

1. *"If Lisa has an essay to write, she will study late in the library".*
2. *"Lisa has an essay to write".*

- *Will Lisa study late in the library?*

*Most subjects (96%) conclude that she will study late in the library. However, if subjects receive an additional information (situation 2):*

3. *"If the library stays open, she will study late in the library".*

*Only a minority (38%) concludes she will study late in the library.*

This study shows that, much like non-monotonic reasoning, conclusions can be suppressed in human reasoning in presence of additional information. A possible explanation for the modus-ponens suppression effect is that humans consider as possibly valid unsupported counter-arguments (attacks), they think that the library might be closed and therefore cannot conclude that Lisa will study in the library. This can be represented by Statement Graph's labellings (as detailed in [5]) and shown in Figure 1 and 2.

SG and its labelling have been implemented in a tool called ELDR [5]. In future work we plan on studying if SGs can be used to represent other non-monotonic reasoning such as the selection task and other types of conflict-tolerant reasoning such as Repair Semantics [8].

# Acknowledgments

The authors have been funded by the EU H2020 773375 Glopack project.

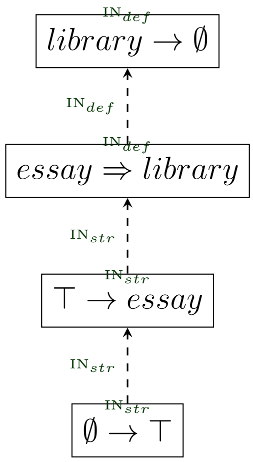



Figure 1: SG of Situation 1.

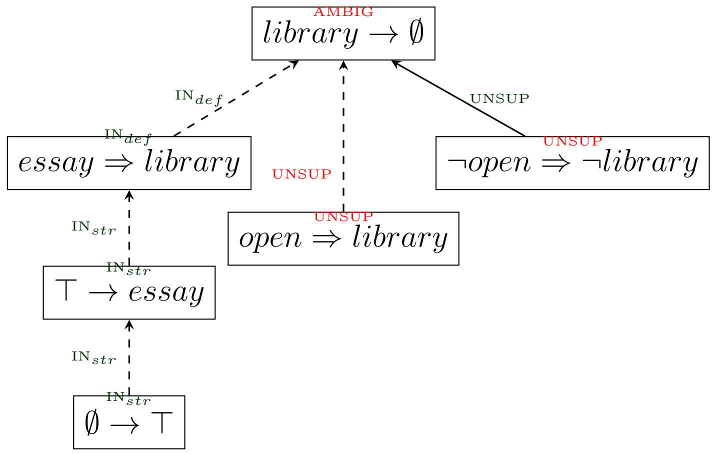



Figure 2: SG of Situation 2.

# References


[1] Nick Bassiliades, Grigoris Antoniou, and Ioannis Vlahavas. A defeasible logic reasoner for the semantic web. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 2(1):1–41, 2006.

[2] Gerhard Brewka and Stefan Woltran. Abstract dialectical frameworks. In *Twelfth International Conference on the Principles of Knowledge Representation and Reasoning*, 2010.

[3] Ruth M J Byrne. Suppressing valid inferences with conditionals. *Cognition*, 31(1):61–83, 1989.

[4] Alejandro J García and Guillermo R Simari. Defeasible logic programming: An argumentative approach. *Theory and practice of logic programming*, 4(1+2):95–138, 2004.

[5] Abdelraouf Hecham, Pierre Bisquert, and Madalina Croitoru. On a Flexible Representation of Defeasible Reasoning Variants. In *Proceedings of the 17th AAMAS Conference*, 2018.

[6] Abdelraouf Hecham, Madalina Croitoru, and Pierre Bisquert. Argumentation-Based Defeasible Reasoning For Existential Rules. In *Proceedings of the 16th AAMAS Conference*, pages 1568—-1569, 2017.

[7] John F Horty, D S Touretzky, and R H Thomason. A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. In *Proceedings of the Tenth IJCAI Conference*, pages 476–482, 1987.

[8] Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati, Marco Ruzzi, and Domenico Fabio Savo. Inconsistency-tolerant semantics for description logics. In *International Conference on Web Reasoning and Rule Systems*, pages 103–117. Springer, 2010.

[9] Donald Nute. *Defeasible reasoning: a philosophical analysis in prolog.* Springer, 1988.

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

[11] Guizhen Yang, Michael Kifer, and Chang Zhao. Flora-2: A Rule-Based Knowledge Representation and Inference Infrastructure for the Semantic Web. *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, pages 671–688, 2003.

---

# The Theory of Scientific Reasoning and Argumentation in Practice

Oct. 4 16:35

Claudia Schulz[1], Jan Kiesewetter[2], Michael Sailer[3], Elisabeth Bauer[3], Martin Fischer[2], Frank Fischer[3] and Iryna Gurevych[1]

[1]Ubiquitous Knowledge Processing (UKP) Lab, Technische Universität Darmstadt, Germany

[2]Institut für Didaktik und Ausbildungsforschung in der Medizin, Klinikum der Universität München, LMU München, Germany

[3]Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie, LMU München, Germany

## Motivation

Analysing human argumentative reasoning behaviour can be advantageous in many applications, such as automatic feedback on students' essays [10, 9], the identification of most helpful or deceptive reviews [8, 2], and aiding eRulemaking [7]. To analyse the argumentative process in text or speech, an understanding of what makes an argument or which steps are involved in argumentation is needed. Some works present their own definition of argument, whereas others try to apply existing formalisations of argumentation (see [5]) such as Toulmin's model of argument [11] or Freeman's theory of argumentation [4]. Arguably, using existing formalisation of argumentation is beneficial as it facilitates a common understanding when analysing argumentation. It furthermore provides an evaluation of the theory as to its applicability to, and representativeness of, human argumentation. In addition, and most importantly, it provides a bridge between the *formalisation* of argumentation and *actual human* argumentation.

## The Theory of Scientific Reasoning and Argumentation (SRA)

Humans use argumentative reasoning not only for persuasion but also for problem-solving. One form of problem-solving crucial in many professions is *diagnosis*: physicians determine a patient's disease based on clinical tests, teachers recognise behavioural disorders in children based on observations, and engineers debug errors in machines or programs based on their analyses of log files or flight recorders. The analysis of the reasoning underlying diagnosis is thus of importance across disciplines for educational applications aiming to understand and improve students' diagnostic reasoning skills.

Building upon findings in education and psychology, Fischer et al. [3] propose a *theory of scientific reasoning and argumentation* (SRA). In contrast to formalisations of argumentation in terms of components such as premise and conclusion used in the context of persuasion, SRA formalises *epistemic activities* involved in problem-solving: problem identification, questioning, hypothesis generation, construction and redesign of artefacts, evidence generation, evidence evaluation, drawing conclusions, and communicating and scrutinising.

Bridging the gap between theory and human argumentation, we choose SRA to analyse students' argumentation when diagnosing. This provides a unified theory to study argumentative reasoning when diagnosing in different disciplines. We here focus on teacher and medical education.

# Analysing Argumentation in Diagnostic Reasoning Texts with SRA

To simulate professional diagnosis, various professional scenarios are outlined to the students, detailing both relevant and irrelevant information about a virtual patient (medicine) or pupil (teacher education). The students' task in each scenario is to decide on a diagnosis and to then write an explanation on how they came up with this diagnosis. These *(diagnostic) reasoning texts* are the object of our argumentation analysis, i.e. we aim to identify epistemic activities in the reasoning texts.

Since the texts contain highly domain specific terminology, we recruited domain-experts for the identification of epistemic activities. These experts simultaneously identified epistemic activity segments and their type. However, the original definitions of epistemic activities could not be applied one-to-one in the context of reasoning texts.

## Building the Bridge between Theory and Human Argumentation

We find that four of the epistemic activities rarely occur in reasoning texts and thus focus on the four frequently used ones: hypothesis generation (HG), evidence generation (EG), evidence evaluation (EE), and drawing conclusions (DC). Their general-purpose definitions furthermore had to be interpreted in our context of reasoning texts stimulated by scenario simulations as follows. See Figure 3 for an example of epistemic activities identified based on our interpretations thereof.

**EG:** Due to the scenario setup, students cannot generate evidence in the original sense, i.e. by performing tests and analyses, since such evidence is already given in the scenario's information. We thus interpret EG as statements describing the

First I wanted to see if the problem was new, so I checked the teacher's observations.
*As it was the same back then*, I ruled out a trauma or another dramatic event.
I was then undecided between autism and ADHD, *since his social behaviour seems to be problematic and that's a sign for both diagnoses*.
In the end, I settled on ADHD *since his script seems chaotic and unorganised* and *because he seems to have some friends despite his difficult behaviour*.

Figure 3: Exemplary diagnostic reasoning text from the teacher education domain, annotated with epistemic activity segments: evidence generation, *evidence evaluation*, drawing conclusions, hypothesis generation.

explicit activity of obtaining evidence from the scenario information or by recalling own knowledge.

**EE:** Many students do not explicitly evaluate evidence concerning its degree of relevance in supporting or refuting a potential answer. We thus interpret the mentioning of evidence as an active selection of information considered relevant and define EE in this manner. Compared to the original definition, we also drop the restriction that EE is targeted at supporting or refuting an answer, since not all students state an answer (hypothesis or conclusion) in their reasoning texts.

We also found that some of the epistemic activities were difficult to distinguish based on their definition by Fischer et al. [3].

**HG versus DC:** In theory, HG is the identification of *possible* solutions often not based on evidence, whereas DC involves aggregating evidence to come to a *final* decision. However, in practice the distinction is less clear. Some students state a possible diagnosis based on evidence at the beginning of the reasoning text or a certain diagnosis without any evidence, other students state a final decision without explicit reference to evidence or with uncertainty. We therefore define the difference between HG and DC based on the role they play in the reasoning process: HG *initiates* whereas DC *terminates* (a part of) the reasoning.

**EE versus DC:** When students generate new knowledge by evaluating given information, it is often difficult to distinguish EE and DC. We thus define DC as an evaluation leading to knowledge that forms an answer to the problem (diagnosis), whereas EE as an evaluation that may lead to knowledge about certain aspects of the problem, e.g. more information about evidence.

# Results and Insights

Using these interpretations of the theory of SRA, we find that the domain-experts can reliably identify epistemic activities in diagnostic reasoning texts (agreement of 0.67 and 0.65 Krippendorff's $\alpha_U$ [6] between the experts in medicine and teacher education, respectively). This indicates that the argumentation formalism chosen, that is SRA, is suitable for analysing argumentative reasoning in diagnostic reasoning texts across domains.

Having built a bridge between the theory and actual human argumentation in one way (from theory to human argumentation), our context-specific interpretations of the definitions of epistemic activities may be useful to go the opposite direction in the future. In other words, the further development of the theory of SRA may be informed by the findings of our analysis of human argumentative reasoning. It is interesting to note that SRA, and in particular the epistemic activities HG and DC, resemble abductive reasoning and inference to the best explanation [1]. A detailed comparison of these theories is part of future work.


# Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (BMBF) under the reference 16DHL1040 (FAMULUS).


# References


[1] Daniel G. Campos. On the distinction between peirce's abduction and lipton's inference to the best explanation. *Synthese*, 180(3):419–442, 2011.

[2] Oana Cocarascu and Francesca Toni. Detecting deceptive reviews using Argumentation. In *Proceedings of the 1st International Workshop on AI for Privacy and Security (PrAISe'16)*, pages 1–8, 2016.

[3] Frank Fischer, Ingo Kollar, Stefan Ufer, Beate Sodian, Heinrich Hussmann, Reinhard Pekrun, Birgit Neuhaus, Birgit Dorner, Sabine Pankofer, Martin R. Fischer, Jan-Willem Strijbos, Moritz Heene, and Julia Eberle. Scientific Reasoning and Argumentation: Advancing an Interdisciplinary Research Agenda in Education. *Frontline Learning Research*, 4:28–45, 2014.

[4] James B. Freeman. *Dialectics and the Macrostructure of Arguments: A Theory of Argument Structure*. De Gruyter, 1991.

[5] Ivan Habernal and Iryna Gurevych. Argumentation Mining in User-Generated Web Discourse. *Computational Linguistics*, 43(1):125–179, 2017.

[6] Klaus Krippendorff. On the Reliability of Unitizing Continuous Data. *Sociological Methodology*, 25:47–76, 1995.

[7] John Lawrence, Joonsuk Park, Katarzyna Budzynska, Claire Cardie, Barbara Konat, and Chris Reed. Using argumentative structure to interpret debates in online deliberative democracy and erulemaking. *ACM Transactions on Internet Technology*, 17(3):25:1–25:22, 2017.

[8] Haijing Liu, Yang Gao, Pin Lv, Mengxue Li, Shiqiang Geng, Minglan Li, and Hao Wang. Using Argument-based Features to Predict and Analyse Review

Helpfulness. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP'17)*, 2017.

[9] Yi Song, Michael Heilman, Beata Beigman Klebanov, and Paul Deane. Applying Argumentation Schemes for Essay Scoring. In *Proceedings of the 1st Workshop on Argumentation Mining (ArgMining'14)*, pages 69–78, 2014.

[10] Christian Stab and Iryna Gurevych. Recognizing Insufficiently Supported Arguments in Argumentative Essays. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL'16)*, pages 980–990, 2016.

[11] Stephen E. Toulmin. *The Uses of Argument.* Cambridge University Press, 1958.

---

# Inference Trees over a Conditional Language as Proxies for Real-World Arguments

Oct. 5
16:00

Emil Weydert

CSC, University of Luxembourg

**Abstract**

In a first approximation, real-world arguments may be modeled as inference trees built from general, semantically justified defeasible local inference steps. The base language allows the expression of first-order default knoowledge and the priorities governing attacks/defeats can be extracted from the local default bases. Modeling summaries of scientific discussions helps to illustrate the strengths and weaknesses of such an approach.

Designing formal models of real-world argumentation may be seen as being part of a broader research program, namely to model general forms of goal-directed human reasoning in the presence of uncertainty, inconsistency, and multiple communicating agents. The popularity of the argumentation paradigm has several reasons. First, the dialectical, incremental way humans tend to organize and express their reasoning activities. Secondly, the simplicity of the structured argument building process (modus-ponens-driven, strict/defeasible dichotomy). Third, the transparency of the aggregative inference and conflict resolution process based on simple attack notions and definitions of acceptable extensions (well-behaved argument sets). However, it is one thing to specify formal models inspired from human argumentation, and quite another one to actually model and analyze real instances of human deliberation on non-trivial issues, even in a mainly normative context.

From a general perspective, formal argumentation is an inferential task concerned with, on one side, the logical modeling of instances of strict and defeasible real-world reasoning and justifications, and on the other side, their evaluation and prioritzation when confronted to conflicting outcomes. Abstract argumentation, as promoted and refined by Dung and its followers, abstracts away from the first concern. Structured argumentation, most prominently implemented within expressive

ASPIC-style frameworks [MP 13], takes into account both. But it still stands very much in the tradition of Dung by relying on its evaluation strategies and also by failing to exploit the deep logical content of the arguments, except for establishing the inter-argument attack structure. This may not be enough to do justice to the inferentially relevant plausibilistic and logical information included in real-world arguments.

If we focus on the macro-inferential structure (ignoring contingent logico-linguistic details), real-world arguments can often be reconstructed as trees constituted by general, usually defeasible inference steps, linking various forms of logically relevant information, including (but not restricted to) conditionals, which will be our focus here. They may also rely on implicit assumptions as determined by linguistic conventions, pragmatic context, or common domain assumptions. These inferential moves are clearly not restricted to instances of strict or defeasible modus ponens. What one can - and should - assume is that they are justifiable by principles of rational inference. This includes modus ponens, but also more general - and speculative - macro patterns of nonnmonotonic inference, as well as domain-transcending argumentation schemes [WRM 08]. In this context it is important to distinguish between contingent conditional information, inhabiting the object-level, and instances of reasonable inference steps, anchored at the meta-level and pushing forward the reasoning process.

**General local inference steps:** To bridge the gap between real and formal argumentation, we propose to investigate in an intermediate step inferential trees composed of inference triples $(i, \Gamma, \psi)$, where $i$ indicates the inference type (e.g. a specific form of strict or defeasible inference), $\Gamma$ is a finite set of premises which are assumed to include complexer forms of factual and conditional knowledge, and $\psi$ is the inferred local claim which may serve as an input to further inference steps. Correctness presupposes that $\Gamma \mid\!\sim_i \psi$ holds, inferential adequacy that $\mid\!\sim_i$ meets certain plausibility requirements. However, for the local steps one may only consider restricted, cognitively manageable instances of such inferential relationships (inspired by actual formal reconstructions). Examples include reasoning by cases and defeasible contraposition.

- This is P or R. Ps are As, Rs are As. Hence (we may assume) this is A.
- As are Bs. This is not B. Hence (we may assume) this is not A.

The first defeasible pattern can be validated by System P, arguingly the simplest plausibilistic default inference notion. The second one can be justified by stronger default reasoning methods, like System Z, or JLZ [Wey 03]. The semantics of the chosen inference relations may be grounded in the quasi-probabilistic ranking semantics or derivatives thereof. This choice reflects the fact that it defines a well-behaved family of default formalisms and also allows to directly specify extension semantics for abstract argumentation frameworks [Wey 14]. Computational complexity is here a lesser issue because the focus is on small local premise bases. Although real-world arguments may have many other inference driving components, we will here concentrate on conditional information where the broad experience from default reasoning can be exploited. This looks like a reasonable first step. While we target a uniform plausible inference account, our framework should be general enough to embed other heterogeneous, hybrid approaches inspired by default reasoning, like the work in [GS 11].

**First-order expressivity:** Real-world arguments are also often based on some kind of first-order reasoning. A prototypical driver of defeasible inference steps at this level is generic quantification, which is also the correct logical form of “Birds fly”. It is important to emphasize here that first-order default information cannot be adequately grasped by universal quantification over open conditionals involving first-order formulas. We need conditionals binding variables. While the monotonic logic of these quantifiers is reasonably well-understood since the nineties [Del 98], it is not straightforward to extend propositional default reasoning accounts to the first-order context, especially - as already observed in nonmonotonic probabilistic reasoning - if there are binary predicates around, hard to avoid in real-world examples. Fortunately, this issue seems to be manageable for the local inference steps within arguments.

**Attack zoo:** The approach sketched above offers specific possibilities to specify attacks. On one hand, rebuttal and undermining need strict attack concepts for facts and conditionals, which are semantically benign. On the other hand, we do not want to rely on user-specified priorities if these not expressed within the formal arguments (e.g. statements about relative plausibility, source reliability, or legal hierarchies possibly). To implement undercut, i.e. attacks on the inference steps themselves, we may make use of the intrinsic preferences available for preferential inference notions, which amounts to override local inference steps, possible because of their generality.

Such inferential tree arguments over first-order conditional languages semantically grounded and evaluated within the ranking measure framework may serve as finer-grained proxies for real-world argumentation. To evaluate and better understand the potential and limitations of such a proposal, we will illustrate it by the inferential formalization of arguments found in short summaries in the Science magazine discussing hot scientific issues. Insofar time allows, we will also take a look at possible applications to the foundations of mathematics.

# References

[Del 98] J.P. Delgrande. On first-order conditional logics. Artificial Intelligence, 105:105-137, 1998. [GS 11] G. Kern-Isberner, G.R. Simari. A Default logical semantics for defeasible argumentation. FLAIRS Conference, 2011. [MP 13] S. Modgil, H. Prakken. A general account of argumentation with preferences. Artificial Intelligence, 195:361-397, 2013.
[Wey 03] E. Weydert. System JLZ - Rational default reasoning by minimal ranking constructions. Journal of Applied Logic, 1(3-4):273308. Elsevier, 2003.
[Wey 14] E. Weydert. A plausibility semantics for abstract argumentation frame-

works. Proc. of the 15th NMR-WS, Arxiv 2014.
[WRM 08] D. Walton, C. Reed, F. Macagno. Argumentation Schemes, Cambridge University Press, 2008.

---