Argumentation Theory in Formal and Computational Perspective

Frans H. van Eemeren
University of Amsterdam
f.h.vaneemeren@uva.nl

Bart Verheij
University of Groningen
bart.verheij@rug.nl

Abstract

Argumentation has been studied since Antiquity. Modern argumentation theory took inspiration from these classical roots, with Toulmin's 'The Uses of Argument' (1958) and Perelman and Olbrechts-Tyteca's 'The New Rhetoric' (1969) as representants of a neo-classical development. In the 1970s, a significant rise of the study of argumentation started, often in opposition to the logical formalisms of those days that lacked the tools to be of much relevance for the study of argumentation as it appears in the wild. In this period, argumentation theory, rhetoric, dialectics, informal logic, and critical thinking became the subject of productive academic study. Since the 1990s, innovations in artificial intelligence supported a formal and computational turn in argumentation theory, with ever stronger interaction with non-formal and non-computational scholars. The present article sketches argumentation and argumentation theory as it goes back to classical times, following the developments before and during the currently ongoing formal and computational turn.

1 Introduction

Argumentation has been studied since Antiquity. Several 20$^{\text{th}}$ century developments in the study of argumentation (in particular since the 1950s) were initiated by concerns that the formal methods of the time, especially classical formal logic, were not fully adequate for the study of argumentation. In recent years, such concerns have been addressed, and partially answered, using innovations in formal and computational methods, in particular in computer science and in artificial intelligence.

We can speak of a formal and computational turn in the study of argumentation. This article sketches argumentation and argumentation theory as it goes back to classical times, following the developments before and during the currently ongoing formal and computational turn. While doing so, we explain what the study of argumentation, generally known as *argumentation theory*, involves. Our exposé is based on the *Handbook of Argumentation Theory* that we recently co-authored with Bart Garssen, Erik C.W. Krabbe, A. Francisca Snoeck Henkemans and Jean Wagemans [Amgoud *et al.*, 2008, in particular Chapters 1 and 11].[1]

In Section 2, 'Argumentation and argumentation theory before the formal and computational turn', we define argumentation in the way this concept has been used in argumentation theory before the formal and computational turn; starting from this definition we explain what argumentation theory is about and describe its main aims. We introduce crucial concepts that play a major role in argumentation theory, and give an overview of prominent theoretical approaches. In Section 3, 'Formal and computational argumentation theory: precursors and first steps', we start the discussion of formal and computational approaches to argumentation by addressing precursors and first steps made, in particular in non-monotonic logic and defeasible reasoning. Section 4, 'Argumentation and the structure of arguments in formal and computational perspective', is about the formalization of argument attack, the structure of arguments, argument schemes and dialogue. In Section 5, 'Specific kinds of argumentation in formal and computational perspective', we discuss argumentation with rules, cases, values and evidence. We conclude the article by looking back at the formal and computational turn in argumentation theory using the crucial concepts of argumentation theory before that turn, and by an outlook into the future of argumentation theory.[2]

[1]Relevant journals include: Artificial Intelligence, Artificial Intelligence and Law, Autonomous Agents and Multi-Agent Systems, Computational Intelligence, International Journal of Cooperative Information Systems, International Journal of Human-Computer Studies, Journal of Logic and Computation, and The Knowledge Engineering Review. Contributions have also been made to journals that deal primarily with argumentation, such as Argumentation and Informal Logic. A journal devoted explicitly to the interdisciplinary area of AI is Argument and Computation. The biennial conference series COMMA is devoted to the study of computational models of argument. The first was held in Liverpool in 2006, followed by conferences in Toulouse (2008), Desenzano del Garda (2010), Vienna (2012), Pitlochry (2014), and Potsdam (2016). See `http://www.comma-conf.org/`. ArgMAS (Argumentation in Multi-Agent Systems) and CMNA (Computational Models of Natural Argument) are related workshops.

[2]The article has been written as a chapter for the *Handbook of Formal Argumentation* (Volume 1: Foundations) (`http://formalargumentation.org/`).

2 Argumentation and argumentation theory before the formal and computational turn

Argumentation, a phenomenon we are all familiar with, arises in response to, or in anticipation of, a real or imagined difference of opinion. It comes into play in cases when people start defending a view they assume not to be shared by others. Not only the need for argumentation, but also the requirements argumentation has to fulfil and the structure of argumentation are connected with a context in which doubt, potential opposition, and perhaps also objections and counterclaims arise.

A definition of argumentation suitable to be used in argumentation theory should connect with commonly recognized characteristics of argumentation. It is important to realize however that there are striking differences between the meaning of the pivotal word 'argumentation' in English usage and the meaning of its lexical counterparts in other languages.[3] A first relevant difference is that the meaning of argumentation in the latter naturally includes both argumentation as a process and argumentation as a product. Second, unlike the English word 'argumentation', its non-English counterparts pertain exclusively to a constructive effort to convince the addressee of the acceptability of one's standpoint, so that argumentation is immediately associated with reasonableness.[4] Third, in the non-English counterparts 'argumentation' is taken to refer only to the constellation of propositions put forward in defence of a standpoint without including the standpoint,[5] so that standpoint and argumentation are viewed as separate entities, which facilitates the study of their relationship [van Eemeren & Grootendorst, 1984, p. 18]. Note that—as we will see below—since the formal and computational turn discussed below, attention for argumentation that goes against a standpoint has increased.

Next to the meaning of the non-English counterparts, which captures some vital characteristics, there are also some general characteristics of argumentation that are independent of any specific language that are taken into account in defining the term argumentation in argumentation theory. To begin with, argumentation is a *communicative act complex*,[6] whose structural design reflects the functional intent of the communicative moves that are made. Next, argumentation is an *interactional act*

[3]For instance, in French 'argumentation,' in German 'Argumentation,' in Italian 'argomentazione,' in Portuguese 'argumentação,' in Spanish 'argumentación,' in Dutch 'argumentatie,' and in Swedish 'argumentation.'

[4]This does not mean, of course, that in practice argumentation cannot be abused, so that there is no matter of acting reasonably.

[5]According to Tindale [1999, p. 45], it is 'the European fashion' to refer to the premises of an argument as the argumentation and to the conclusion by using another term, such as standpoint.

[6]Because argumentation can also be non-verbal, for instance, visual, it is defined here—more generally—as a 'communicative' rather than a 'verbal' ('linguistic') act complex.

complex directed at eliciting a response that indicates acceptance of the standpoint that is defended, so that it is always part of an explicit or implicit dialogue with the addressee. Further, as a rational activity of reason, argumentation involves putting forward a constellation of propositions the arguer can be held accountable for, so that it is not just an expressive but creates commitments. Finally, in making an appeal to common critical standards of reasonableness in trying to convince the addressee, the arguer approaches the addressee as a rational judge who judges reasonably.[7]

Based on these starting points, defining argumentation starts from ordinary usage and is next made more precise and explicit in order to adequately serve its purpose in argumentation theory:

Argumentation is a communicative and interactional act complex aimed at resolving a difference of opinion with the addressee by putting forward a constellation of propositions *the arguer can be held accountable for* to make the standpoint at issue acceptable to *a rational judge who judges reasonably*.[8]

Argumentation theory is the umbrella term used to denote the study of argumentation in all its manifestations and varieties, irrespective of the intellectual backgrounds, primary research interests and angles of approach of the theorists. Other general labels, such as informal logic and rhetoric, refer to specific theoretical perspectives on the study of argumentation (and usually also include other research interests than argumentation).

Because the standpoints at issue in a difference of opinion and the argumentation advanced to support them can pertain to all walks of life and all kinds of subjects, the scope of argumentation theory is very broad. It ranges from argumentative discourse in the public and the professional sphere to argumentative discourse in the personal or private sphere. The types of standpoints supported by argumentation may vary from descriptive standpoints to evaluative and prescriptive standpoints. It is in particular worth noting that argumentation is certainly not used only for truth-finding and truth-preservation.[9]

Scholars are often drawn to studying argumentation by their practical interest in improving the quality of argumentative discourse where this is called for. In order

[7]Although the terms rational and reasonable are often used interchangeably, we think that it is useful to make a distinction between acting 'rationally' in the sense of using one's faculty of reason and acting 'reasonably' in the sense of utilizing one's faculty of reason in an appropriate way.

[8]The term argumentation refers to the whole constellation of propositions put forward in defence of the standpoint. Because each of the propositions constituting the constellation has its own share in providing grounds for accepting the standpoint at issue, in principle, these propositions by themselves also have an argumentative function. This is expressed terminologically by calling them the reasons that make up the argumentation as a whole.

[9]Generally, in discussing a claim to acceptance, argumentation has in fact no major role to play when a decisive solution can readily be offered by other means.

to be able to realize this ambition, they have to combine an empirical orientation towards how argumentative discourse is conducted with a critical orientation towards how it should be conducted. To give substance to this challenging combination, they need to carry out a comprehensive research programme that ensures that argumentative discourse will not only be examined descriptively as a specimen of verbal communication and interaction ('pragmatics') but also be measured against normative standards of reasonableness ('normative pragmatics') [van Eemeren, 1990].

In order to combine critical and empirical insights systematically, in argumentation theory argumentation scholars make it their business to bridge the gap between the normative dimension and the descriptive dimension of argumentative discourse. The complex problems that are at stake are to be solved with the help of a research programme with five interrelated components [van Eemeren & Grootendorst, 2004, pp. 9–41].[10] On the one hand, the programme has a *philosophical* component, in which a philosophy of reasonableness is developed, and a *theoretical* component, in which, starting from this philosophy, a model for argumentative discourse is designed. On the other hand, the programme has an *empirical* component, in which argumentative reality as it manifests itself in communicative and interactional exchanges is investigated. Next, in the pivotal *analytical* component of the research programme, the normative and the descriptive dimensions are systematically linked together by a theoretically motivated and empirically justified reconstruction of argumentative discourse. Finally, in the *practical* component the problems that occur in the various kinds of argumentative practices are identified, and methods are developed to tackle these problems.

In developing a philosophy of reasonableness argumentation theorists reflect in the philosophical component upon the rationale for the view of reasonableness that is to underlie their theoretical approach. Depending on the conception of reasonableness they favour, in the theoretical component standards for the validity, soundness or appropriateness of argumentation are adopted and theoretical models are developed based on these conceptions. Because the model of argumentation is in this case a normative instrument for assessing the quality of argumentation put forward in argumentative reality, the model constitutes a point of orientation for the empirical research that is to be carried out in argumentation theory but does not constitute a test of the model. The model indicates which factors and processes are worth investigating and to what extent the norms prevailing in argumentative reality agree with the theoretical standards, but deviations are not necessarily an indication of any wrongness in the model.[11]

[10]The five components of a fully-fledged research programme in argumentation theory were introduced in van Eemeren [1987].

[11]Only in case of a purely descriptive theory the empirical research could be aimed at testing

Analytical research in argumentation theory is aimed at the reconstruction of argumentative discourse as it occurs in argumentative reality from the perspective of the model of argumentation that is chosen as the theoretical starting point. Whichever theoretical background they may have, argumentation theorists engaging in analytical research need to develop appropriate tools and methods for reconstructing argumentative discourse. Practical research in argumentation theory, finally, is aimed at analyzing the (spoken and written) argumentative practices that can be distinguished in the various communicative domains from the perspective of argumentation theory and developing instruments for intervention in argumentative discourse where this is due. The instruments for enhancing the quality of argumentative practices may consist of designs for the formats of communicative activity types or of methods for improving arguers' skills in analysing, evaluating and producing argumentative discourse.

In the end, the general objective of argumentation theory is a practical one: to provide adequate instruments for analysing, evaluating and producing argumentative discourse. Ultimately the *raison d'être* of the other components of the research programme carried out in argumentation theory is that they enable the systematic development of such instruments. When taken together, philosophical and theoretical insights into argumentative discourse, analytically connected with empirical insights, are to lead to methodical applications of argumentation theory to the various kinds of argumentative practices.

In pursuing their objective of improving the analysis, evaluation and production of argumentative discourse, argumentation theorists take account of the point of departure of argumentation, consisting of the explicit and implicit material and procedural premises that serve as the starting point, and the layout of the argumentation displayed in the constellation of propositions explicitly or implicitly advanced in support of the standpoints at issue. Both the point of departure and the layout of argumentation are to be judged by appropriate standards of evaluation that are in agreement with all requirements a rational judge who judges reasonably should comply with. This means that the descriptive and normative aims of argumentation theory as a discipline can be specified as follows:[12]

1. Giving a descriptive account of the components of argumentative discourse that constitute together the point of departure of argumentation;

the model, but so far no fully-fledged argumentation theory without a critical dimension has been developed.

[12]The descriptive aims of argumentation theory are often associated with the 'emic' study of what is involved in justifying claims and what are good reasons for accepting a claim viewed from the 'internal' perspective of the arguers while the normative aims are associated with the 'etic' study of these matters viewed from the 'external' perspective of a critical theorist.

2. Giving a normative account of the standards for evaluating the point of departure of argumentation that are appropriate to a rational judge who judges reasonably;

3. Giving a descriptive account of the components of argumentative discourse that constitute together the layout of argumentation;

4. Giving a normative account of the standards for evaluating argumentation as it is laid out in argumentative discourse that are appropriate to a rational judge who judges reasonably.

2.1 Crucial concepts

Certain theoretical concepts are indispensable in developing instruments for methodically improving the quality of the analysis, evaluation and production of argumentative discourse. Among them are the notions of 'standpoint,' 'unexpressed premise,' 'argument scheme,' 'argumentation structure,' and 'fallacy.' All of them are immediately connected with central problem areas in argumentation theory.

Standpoints We use the term standpoint (or point of view) to refer to what is at issue in argumentative discourse in the sense of what is being argued about.[13] In advancing a standpoint the speaker or writer assumes a positive or negative position regarding a proposition. Because advancing a standpoint implies undertaking a positive or negative commitment, in view of the aim of resolving a difference of opinion, whoever advances a standpoint is obliged to defend their standpoint if challenged to do so by the listener or reader. The standpoints at issue in a difference of opinion can be descriptive, evaluative or prescriptive, but in all cases they can be reconstructed as a claim to acceptability (in case of a positive standpoint) or unacceptability (in case of a negative standpoint) regarding the proposition the standpoint pertains to.[14]

Unexpressed premises Unexpressed premises are often pivotal missing links in transferring acceptance from the premises that are explicitly put forward in the argu-

[13]The terms claim, conclusion, thesis and debate proposition are used to refer from different theoretical angles to virtually the same concept as the term standpoint. Terms such as belief, opinion and attitude usually refer to related concepts that are in relevant ways different from a standpoint.

[14]For an overview of the various approaches to standpoints see Houtlosser [2001].

mentation to the standpoint that is defended.[15] Such partly implicit argumentation, which is quite usual in ordinary argumentative discourse, is called enthymematic. The identification of elements left implicit in enthymematic argumentation is in practice usually unproblematic, but in some cases it can be a problem. According to most argumentation theorists, then carrying out a logical analysis does not suffice. Starting from a logical analysis, a pragmatic analysis needs to be carried out in which the analyst tries to identify the unexpressed premise by determining on the basis of the available contextual and background information to which implicit proposition the arguer can be held committed to.[16]

Argument schemes An argument(ation) scheme is an abstract characterization of the way in which in a particular type of argumentation a reason used in support of a standpoint is related to that standpoint in order to bring about a transfer of acceptance from that reason to the standpoint. Depending on the kind of relationship established in the argument scheme, specific kinds of evaluative questions—usually referred to as critical questions—are to be answered in evaluating the argumentation. These critical questions capture the specific pragmatic rationale for bringing about the transition of acceptance.[17]

Argumentation structures The argumentation structure of a piece of argumentative discourse characterizes the 'external' organization of the argumentation that is advanced: how do the reasons put forward in a particular argumentation hang together and in what way exactly do they relate to the standpoint at issue? In argumentation theory, various ways of combining reasons have been distinguished that characterize the different kinds of argumentation structures that can be instrumental in defending a standpoint.[18]

[15]Depending on the theoretical background of the theorists, other terms are used to refer to an unexpressed premise: implicit, suppressed, tacit, and missing premise, reason or argument, but also warrant, implicature, supposition, and even assumption, inference and implication.

[16]For an approach in which a logical analysis is used as a heuristic tool in carrying out a pragmatic analysis see van Eemeren and Grootendorst [1992, pp. 64–67]; [2004, pp. 117–118]. For the various kinds of resources that can be used in accounting for the reconstruction see van Eemeren [2010, pp. 16–29]

[17]For an overview of the study of argument schemes, see Garssen [2001]; for attempts at formalization and the computational implications, see Walton, Reed and Macagno [2008, Ch. 11 and 12]. A recent development is the study of what have been called prototypical argumentative patterns. These consist of constellations of argumentative moves in which a particular argument scheme or combination of argument schemes is used [van Eemeren, 2017].

[18]Different terminological conventions have been developed for naming the combinations of reasons and the divisions of the various types of structures are not always exactly the same. For an overview of the study of argumentation structures see Snoeck Henkemans [2001].

Fallacies The difference of opinion at issue in argumentative discourse will not be resolved satisfactorily if contaminators of the argumentative exchange enter the discourse that are not detected. Such contaminators, which may be so treacherous that they go unobserved in the argumentative exchange, are known as fallacies. Virtually every normative theory of argumentation includes a treatment of the fallacies. The degree to which a theory of argumentation makes it possible to give an adequate treatment of the fallacies can even be considered as a litmus test of the quality of the theory.[19]

2.2 Prominent theoretical approaches

Ancient dialectic and rhetoric—in combination with syllogistic logicare the forbears of modern argumentation theory.[20] The Aristotelian concept of dialectic is best understood as the art of inquiry through critical dialogue. In a dialogue that is dialectical in the Aristotelian sense the adequacy of any particular claim is supposed to be cooperatively assessed by eliciting premises that might serve as commonly accepted starting points, then drawing out implications from those starting points and determining their compatibility with the claim in question. Where contradictions emerge, revised claims might be put forward to avoid such problems. This method of regimented opposition amounts to a pragmatic application of logic, a collaborative method of putting logic into use so as to move from conjecture and opinion to more secure belief.

Aristotle's rhetoric deals with the principles of effective persuasion leading to assent or consensus. It bears little resemblance to modern-day persuasion theories heavily oriented to the analysis of attitude formation and attitude change but largely indifferent to the problem of the invention of persuasive messages [Eagly & Chaiken, 1993; O'Keefe, 2002]. In Aristotle's rhetoric, the emphasis is on the production of effective argumentation for an audience when the subject matter does not lend itself to a logical demonstration of certainty. When it comes to logical demonstration, the syllogism is the most prominent form; the enthymeme, thought of as an incomplete syllogism whose premises are acceptable to the audience, is its rhetorical counterpart. As yet, there is no unitary theory of argumentation available that encompasses the dialectical and rhetorical dimensions of argumentation and is universally accepted. The current state of the art in the argumentation theory (as it developed before the

[19] For a more detailed overview of the study of fallacies see van Eemeren [2001].

[20] Although ancient dialectic and rhetoric are often discussed as if both of them were unified wholes, contributions to their development have been made by various scholars and their views were by no means always in harmony. In order to be accurate, we must therefore always indicate precisely to whose views exactly we are referring.

recent formal and computational turn) is characterized by the co-existence of a variety of theoretical perspectives and approaches, which differ considerably from each other in conceptualization, scope and theoretical refinement. Every fully-fledged theoretical approach to argumentation represents in fact a particular specification of what it means for a rational judge to judge reasonably and provides a definition of (crucial aspects of) the type of validity favoured by the theorist.

Some argumentation theorists, especially those having a background in linguistics, discourse analysis or rhetoric, have a goal that is primarily (and sometimes even exclusively) descriptive. They are interested in finding out how in argumentative discourse speakers and writers try to convince or persuade others. Other argumentation theorists, often inspired by logic, philosophy or insights from law, study argumentation primarily for normative purposes. They are interested in developing validity or soundness criteria that argumentation must satisfy in order to qualify as rational or reasonable. Currently, however, most argumentation theorists seem to recognize that argumentation research has a descriptive as well as a normative dimension and that in argumentation theory both dimensions must be combined.[21]

Most modern approaches to argumentation are strongly affected by the perspectives on argumentation developed in Antiquity. Both the dialectical perspective (which nowadays usually incorporates the logical dimension) and the rhetorical perspective are represented prominently. Approaches to argumentation that are dialectically oriented tend to focus primarily on the quality of argumentation in defending standpoints in regulated critical dialogues. They put an emphasis on guarding the reasonableness of argumentation by means of regimentation. It is noteworthy that in the rhetorically oriented approaches to argumentation putting an emphasis on factors influencing the effectiveness of argumentation, effectiveness is usually viewed as a 'right to acceptance' that speakers or writers are, as it were, entitled to on the basis of the qualities of their argumentation rather than in terms of actual persuasive effects.[22]

[21]The infrastructure of the field of argumentation theory in terms of academic associations, journals and book series reflects to some extent the existing division in theoretical perspectives. The American Forensic Association (AFA), associated with the National Communication Association, and its journal *Argumentation & Advocacy* concentrate on argumentation, communication and debate. The Ontario Society for the Study of Argumentation (OSSA), the Association of Informal Logic and Critical Thinking (AILACT) and the electronic journal *Informal Logic* focus on informal logic. The International Society for the Study of Argumentation (ISSA), the journals *Argumentation* and *Journal of Argumentation in Context*, and the accompanying book series Argumentation Library and Argumentation in Context aim to cover the whole spectrum of argumentation theory. Other international journals relevant to argumentation theory are *Philosophy and Rhetoric*, *Logique et Analyse*, *Controversia*, *Pragmatics and Cognition*, *Argument and Computation*, and *Cogency*.

[22]Research aimed at examining the actual effectiveness of argumentation is usually called persuasion research. In practice, it generally amounts to quantitative empirical testing of the ways in

In modern argumentation theory a remarkable revival has taken place of both dialectic and rhetoric. Unlike in Aristotle's approach, however, there is a wide conceptual gap between the two perspectives on argumentation, going together with a communicative gap between their protagonists. In recent times, some argumentation scholars have come to the conclusion that the dialectical and rhetorical views on argumentation are not per se incompatible. It has even been argued that re-establishing the link between dialectic and rhetoric will enrich the analysis and evaluation of argumentative discourse [van Eemeren, 2010, especially Ch. 3].

In giving a brief overview of the current theoretical approaches, we first turn to two 'neo-classical' proposals developed in the 1950s: the Toulmin model and the 'new rhetoric'. In dealing with argumentation both aim to counterbalance the formal approach that modern logic provides for dealing with analytic reasoning.

In *The uses of argument*, first published in 1958, Toulmin [2003] reacted against the then dominant logical view that argumentation is just another specimen of the reasoning that the formal approach is qualified to deal with. As an alternative, he presented a model of the 'procedural form' of argumentation aimed at capturing the functional steps that can be distinguished in the defence of a standpoint by means of argumentation. The procedural form of argumentation is, according to Toulmin, 'field-independent', meaning that the steps that are taken are always the same, irrespective of the subject that is being discussed.[23]

In judging the validity of argumentation, Toulmin gives the term validity a different meaning than it has in formal logic. The validity of argumentation is in his view primarily determined by the degree to which the (usually implicit) warrant that connects the data advanced in the argumentation with the claim at issue is acceptable—or, if challenged, can be made acceptable by a backing. What kind of backing may be required in a particular case depends on the field to which the standpoint at issue belongs. This means that the criteria used in evaluating the validity of argumentation are in Toulmin's view 'field-dependent'. Thus, Toulmin puts the validity criteria for argumentation in an empirical and historical context.

In their monograph *The new rhetoric*, also first published in 1958, Perelman and Olbrechts-Tyteca [1969] regard argumentation—in line with classical rhetoric—as sound if it adduces or reinforces assent among the audience to the standpoint at issue. The audience addressed may be a 'particular' audience consisting of a specific person or group of people, but it can also be the 'universal' audience—the (real or imagined) audience that, in the arguer's view, embodies reasonableness.

which argumentation and other means of persuasion lead to changes of attitude in the recipients [O'Keefe, 2002].

[23]It is noteworthy that Toulmin's model of the argumentative procedure is in fact conceptually equivalent to the extended syllogism known in Roman-Hellenistic rhetoric as epicheirema.

Besides an overview of the elements of agreement that can in argumentation serve as points of departure (facts, truths, presumptions, values, value hierarchies and topoi[24]), Perelman and Olbrechts-Tyteca provide an overview of the argument schemes that in the layout of argumentation can be used to convince or persuade an audience. The argument schemes they distinguish remain for the most part close to the classical topical tradition. Apart from argumentative techniques of 'association', in which these argument schemes are employed, Perelman and Olbrechts-Tyteca also distinguish an argumentative technique of 'dissociation.' Dissociation divides an existing conceptual unity into two separate conceptual unities.

In spite of obvious differences between Toulmin's approach to argumentation and that of Perelman and Olbrechts-Tyteca, there are also some striking commonalities. Starting from an interest in the justification of views by means of argumentative discourse, both emphasize that values play a part in argumentation, both reject formal logic as a theoretical tool, and both turn for an alternative model to juridical procedures. A theoretical connection between the Toulmin model and the new rhetoric could be made by viewing the various points of departure distinguished in the new rhetoric as representing different types of data in the Toulmin model and its argument schemes as different types of warrants or backings.

Of the approaches to argumentation that have been developed more recently, formal dialectic, coined and instigated by Hamblin [1970], remains closest to formal logic, albeit logic in a dialectical garb. The scholars responsible for the revival of dialectic in the second part of the twentieth century treat argumentation as part of a formal discussion procedure for resolving a difference of opinion by testing the tenability of the 'thesis' at issue against challenges. Apart from the ideas about formal dialectic articulated by Hamblin, in designing such a procedure they make use of the 'dialogue logic' of the Erlangen School [Lorenzen & Lorenz, 1978], but also from insights advanced by Crawshay-Williams [1957]; Næss [1966]. The most complete proposal was presented by Barth and Krabbe [1982] in *From axiom to dialogue*. Their formal dialectic describes systems for determining by means of a regimented dialogue game between the proponent and the opponent of the thesis whether the proponent's thesis can be maintained given the premises allowed as 'concessions' by the opponent.

Building on the proposals for a dialogue logic made by the Erlangen School, Barth and Krabbe's formal dialectic offers a translation of formal logical systems into formal rules of dialogue. In *Commitment in dialogue*, Walton and Krabbe [1995] integrate the proposals of the Erlangen School with the more permissive kind of dialogues promoted in Hamblin's [1970] dialectical systems. After having provided a

[24]Perelman and Olbrechts-Tyteca use the Latin equivalent loci.

classification of the main types of dialogue, they discuss the conditions under which in argumentation commitments should be maintained or may be retracted without violating any of the rules of the type of dialogue concerned.

Related approaches can be found in some of the proposals made by formal and informal logicians. Out of dissatisfaction with the treatment of argumentation in logical textbooks, and inspired by the Toulmin model (and to a much lesser extent the new rhetoric), a group of Canadian and American philosophers have propagated since the 1970s an approach known as informal logic. The label informal logic refers in fact to a collection of logic-oriented normative approaches to the study of reasoning in ordinary language which remain closer to the practice of argumentation than is usually the case in formal logic. Informal logicians aim in the first place at developing adequate norms for interpreting, assessing and construing argumentation.

Since 1978, the journal *Informal Logic*,[25] started and edited by Blair and Johnson (later joined by others), has been the speaking voice of informal logic and the connected educational reform movement dedicated to 'critical thinking'. In their textbook *Logical self-defense*, Johnson and Blair [2006] have indicated what they have in mind when they speak of an informal logical alternative to formal logic. They explain that the premises of an argument have to meet the criteria of 'acceptability', 'relevance' and 'sufficiency'. Other informal logicians have adopted these three criteria, albeit sometimes under slightly different names (e.g., [Govier, 1987]).

Freeman [2005] provides, from an epistemological perspective on informal logic, a comprehensive theory of premise acceptability. Generally, however, informal logicians remain in the first place interested in the premise-conclusion relations in arguments (e.g., [Walton, 1989]). Most of them maintain that argumentation should be valid in some logical sense, but generally they do not stick to the formal criterion of deductive validity. Woods and Walton [1989] claim that each fallacy requires its own theoretical treatment, which leads them to applying a variety of logical systems in their theoretical treatment of the fallacies. Johnson [2000] also takes a predominantly logical approach, but he complements this approach with a 'dialectical tier', where the arguer discharges his or her dialectical obligations, for instance, by anticipating objections, and dealing with alternative positions. In Finocchiaro's contributions to informal logic, too, the logical and the dialectical approach are combined, albeit that the emphasis is more strongly on the dialectical dimension, and historical and empirical dimensions are added (e.g., [Finocchiaro, 2005]). The rhetorical perspective has received less attention from informal logicians. A notable exception is Christopher Tindale [1999; 2004].

In modern times, the study of rhetoric has fared considerably better in the United

[25] At first named Informal Logic Newsletter.

States than in Europe. Not only has classical rhetoric from the nineteenth century onwards been represented in the academic curriculum, but also has the development of modern rhetorical approaches been more prolific. In the last decades of the twentieth century, the image that rhetoric had acquired of being irrational and even anti-rational has been revised. Paying tribute to Perelman and Olbrechts-Tyteca's new rhetoric, in various countries various scholars have argued for a rehabilitation of the rhetorical approach. In spite of the unlimited extension in the United States in the 1960s of the scope of Big Rhetoric 'to the point that everything, or virtually everything, can be described as 'rhetorical' [Swearingen & Schiappa, 2009, p. 2], Wenzel [1987] emphasized the rational qualities of rhetoric. In Europe, Reboul [1990] and Kopperschmidt [1989a] argued at about the same time for giving rhetoric its rightful position in the study of argumentation beside dialectic.

Although all of them may be described as rhetoricians in the broad sense, the American scholars from the field of (speech) communication currently engaged in argumentation theory do not share a clearly articulated joint perspective. Their most obvious common feature is a concern with the connection between claims and the people engaged in some kind of argumentative practice. The American debate tradition in particular has had an enormous influence on American argumentation studies. More or less outside the immediate debate tradition, Zarefsky [2006; 2009], Leff [2003] and Schiappa [2002] have contributed profound historical rhetorical analyses. Fahnestock [1999; 2009] dealt theoretically with rhetorical figures and stylistics.

Concentrating on the public features of communicative acts, Jackson and Jacobs [1982] initiated a research programme to study argumentation in informal conversations. Their joint research aims at understanding the reasoning processes by which individuals make inferences and resolve disputes in ordinary conversation. A related empirical angle in American argumentation research is the study of argument in natural settings, such as school board meetings, counseling sessions and public relations campaigns, to produce 'grounded theory'—a theory of the specific case.

A Toulminian concept that has strongly influenced American argumentation scholarship is the notion of 'field'. Toulmin [1972] describes fields as 'rational enterprises', which he equates with intellectual disciplines, and explores how the nature of reasoning differs from field to field. This treatment led to vigorous discussion about what defines a 'field of argument': subject matter, general perspective, world-view, or the arguer's purpose—to mention just a few of the possibilities. The concept of fields of argument encouraged recognition that the soundness of arguments is not something universal and necessary, but context-specific and contingent. Instead of the term fields, Goodnight prefers the term spheres, referring to 'the grounds upon which arguments are built and the authorities to which arguers appeal' [1982, p. 216]. He uses 'argument' to mean interaction based on dissensus, so that the grounds of arguments lie in doubts and uncertainties. In a similar vein as Habermas [1984], Goodnight [2012] distinguishes between three spheres of argument: the 'personal' (or 'private') sphere, the 'public' sphere, and the 'technical' sphere.

Meanwhile, starting in the 1970s, in Europe a descriptive approach has developed in which argumentation is viewed as a linguistic phenomenon that not only manifests itself in language use, but is also inherent in most language use. In a number of publications (almost exclusively in French), the protagonists of this approach, Ducrot and Anscombre, have presented a linguistic analysis to show that almost all verbal utterances lead the listener or reader—often implicitly—to certain conclusions, so that their meaning is crucially argumentative. In *L'argumentation dans la langue* [Anscombre & Ducrot, 1983] they refer to the theoretical position they adopt as radical argumentativism. Their approach is characterized by a strong interest in words that can serve as argumentative 'operators' or 'connectors', giving linguistic utterances a specific argumentative force and argumentative direction (e.g., 'only', 'no less than', 'but', 'even', 'still', 'because', 'so'). Anscombre[1994] observes that the argumentative principles that are at issue here are on a par with the topoi from classical rhetoric.

It has become a tradition among a substantial group of European researchers, primarily based in the French-speaking world, to approach argumentation from a descriptive linguistic angle. Some of them continue the approach started by Ducrot and Anscombre. Others, such as Plantin [1996] and Doury [1997], build on this approach but are also—and often more strongly—influenced by conversation analysis and discourse analysis. Other researchers, based in Switzerland, who favour a linguistic approach, but allow also for normativity, are Rigotti [2009], Rocci [2009], and Greco Morasso [2011]. They combine their linguistic approach with insights from other approaches, such as pragma-dialectics.

The pragma-dialectical theory of argumentation developed in Amsterdam combines a dialectical and a rhetorical perspective on argumentation and is both normative and descriptive. As van Eemeren and Grootendorst [1984] explain, pragma-dialecticians view argumentation as part of a discourse aimed at resolving a difference of opinion on the merits by methodically testing the acceptability of the standpoints at issue. The dialectical dimension of the approach is inspired by normative insights from critical rationalism and formal dialectics, the pragmatic dimension by descriptive insights from speech act theory, Gricean pragmatics and discourse analysis.

The various stages argumentative discourse must pass through to resolve a difference of opinion on the merits by a critical exchange of speech acts are in the pragma-dialectical theory laid down in an ideal model of a critical discussion [van Eemeren & Grootendorst, 2004)] Viewed analytically, there should be a 'confrontation stage', in which the difference of opinion comes about, an 'opening stage', in which the point of departure of the discussion is determined, an 'argumentation stage', in which the standpoints at issue are defended against criticism, and a 'concluding stage', in which it is determined what the result of the discussion is. The model of a critical discussion defines the nature and the distribution of the speech acts that have a constructive role in the various stages of the resolution process. In addition, the standards of reasonableness authorizing the performance of particular speech acts in the various stages of a critical discussion are laid down in a set of dialectical rules for critical discussion. Any violation of any of the rules amounts to making an argumentative move that is an impediment to the resolution of a difference of opinion on the merits and is therefore fallacious [van Eemeren & Grootendorst, 1992].[26]

Because argumentative discourse generally diverges for various reasons from the ideal of a critical discussion, in the analysis of the discourse a reconstruction is required to achieve an analytic overview of all those, and only those, speech acts that play a potential part in resolving a difference of opinion on the merits. Van Eemeren, Grootendorst, Jackson and Jacobs [1993] emphasize that the reconstruction should be guided by the theoretical model of a critical discussion and faithful to the commitments that may be ascribed to the arguers on the basis of their contributions to the discourse. Because the reconstruction of argumentative discourse as well as its evaluation can be made more pertinent, more precise, and also better accounted for if, next to the maintenance of dialectical reasonableness, the simultaneous pursuit of rhetorical effectiveness is taken into account, van Eemeren and Houtlosser [2002] developed the notion of strategic manoeuvring. This notion makes it possible to integrate relevant rhetorical insights systematically in the pragma-dialectical analysis and evaluation [van Eemeren, 2010].

3 Formal and computational argumentation theory: precursors and first steps

Today much research addresses argumentation using formal and computational methods. Precursors can be found in the fields of non-monotonic logic and logic programming, and first steps were made by philosophers addressing defeasible reasoning.

[26]The extent to which the rules for critical discussion are capable of dealing with the defective argumentative moves traditionally designated as fallacies is viewed as a test of their 'problem-solving validity'. For experimental empirical research of the 'intersubjective acceptability' of the rules for critical discussion that lends them 'conventional validity' see van Eemeren, Garssen and Meuffels [2009].

3.1 Non-monotonic logic

A relevant field predating the formal and computational study of argumentation is non-monotonic logic [Antonelli, 2010]. A logic is non-monotonic when a conclusion that, according to the logic, follows from certain premises need not always follow when premises are added. In contrast, classical logic is monotonic. For instance, in a standard classical analysis, from premises 'Edith goes to Vienna or Rome' and 'Edith does not go to Rome', it follows that 'Edith goes to Vienna', irrespective of possible additional premises. The standard example of non-monotonicity used in the literature of the 1980s concerns the flying of birds. Typically, birds fly, so if you hear about a bird, you will conclude that it can fly. However, when you next learn that the bird is a penguin, you retract your conclusion. In a non-monotonic logic, a balance can be sought between the advantage of drawing a tentative conclusion, which is usually correct, and the risk of having to withdraw the conclusion in light of new information.

A prominent proposal in non-monotonic logic is Raymond Reiter's [1980] logic for default reasoning, using default rules. Reiter's first example of a default rule expresses that birds typically fly:

$$\text{BIRD}(x) : \text{M FLY}(x) \ / \ \text{FLY}(x)$$

The default rule expresses that, if x is a bird, and it is consistent to assume that x can fly, then by default one can conclude that x can fly. Other influential logical systems for non-monotonic reasoning include circumscription, auto-epistemic logic, and non-monotonic inheritance; each of them discussed in the representative overview of the study of non-monotonic logic at its heyday by Gabbay, Hogger and Robinson [1994].

3.2 Logic programming

A development related to non-monotonic logic is logic programming. The general idea underlying logic programming is that a computer can be programmed using logical techniques. In this view, computer programs are not only considered procedurally as recipes for how to achieve the program's aims, but also declaratively, in the sense that the program can be read like a text, for instance, as the rule-like knowledge needed to answer a question. In the logic programming language Prolog (the result of a collaboration between Colmerauer and Kowalski; see [Kowalski, 2011]), these are examples of facts and a rule [Bratko, 2001]:

parent(pam, bob)
female(pam)
mother(X, Y) :- parent(X, Y), female(X)

This small logic program represents the facts that Pam is Bob's parent, and that Pam is female, and the rule that someone's mother is a female parent. Given this Prolog program, a computer can as expected derive that Pam is Bob's mother. In the interpretation of logic programs, the closed world assumption plays a key role: a logic program is assumed to describe all facts and rules about the world. For instance, in the program above it is assumed that all parent relations are given, so 'parent(tom, bob)' cannot be derived. By what is called negation as failure, it will be considered false that Tom is Bob's parent. If we add 'parent(tom, bob)' it becomes derivable that Tom is Bob's parent, showing the connection between logic programming's negation as failure and non-monotonic logic.

3.3 Themes and impact of non-monotonic logics

The study of non-monotonic logics gave hope that logical tools would become more relevant for the study of natural reasoning. To some extent this hope has been fulfilled, since certain themes that before were at the boundaries of logic, were now placed in the centre of attention. Examples of such themes are defeasible inference, consistency preservation, and uncertainty. In the handbook edited by Gabbay, Hogger and Robinson [1994], Donald Nute discusses defeasible inference that can be blocked or defeated in some way [Nute, 1994, p. 354]. Interestingly, Donald Nute speaks of the presentation of sets of beliefs as reasons for holding other beliefs as advancing arguments. David Makinson [1994, p. 51] describes consistency preservation as the property that the conclusions drawn on the basis of certain premises can only be inconsistent in case the premises are inconsistent. Henry Kyburg [1994, p. 400] distinguishes three kinds of inference involving uncertainty: classical, deductive, valid inference about uncertainty; an 'inductive' kind where a conclusion can be false even when the premises are true (hence distinct from the idea of induction as going from the specific to the general, and closer to what today is often called 'defeasible'); and a kind of inference with uncertainty that gives probabilities of particular statements.

The study of non-monotonic logic has been very successful as a research enterprise, and coincided with innovations in computer programming in the form of logic-based languages such as Prolog, and to commercial applications: today's knowledge-based expert systems—in wide-spread use—often include some elementary form of non-monotonic reasoning.

At the same time, non-monotonic logic did not fulfil all expectations of the artificial intelligence community in which it was initiated. Matthew Ginsberg [1994], for instance, notes—somewhat disappointedly—that the field put itself "in a position where it is almost impossible for our work to be validated by anyone other than a member of our small subcommunity of Artificial Intelligence as a whole" [1994, p. 28–29]. His diagnosis of this issue is that attention shifted from the key objective of building an intelligent artefact to the study of simple examples and mathematics. This leads him to plead for a more experimental, scientific attitude as opposed to a theoretical, mathematical focus.

3.4 Defeasible reasoning

In 1987, the publication of John Pollock's paper 'Defeasible reasoning' in *Cognitive Science* marked a turning point. The paper emphasized that the philosophical notion of 'defeasible reasoning' coincides with what in AI is called 'non-monotonic reasoning.' As philosophical heritage for the study of defeasible reasoning, Pollock [1987] refers to works by Roderick Chisholm (going back to 1957) and himself (earliest reference in 1967). Ronald Loui [1995] places the origins of the notion of 'defeasibility' a decade earlier, namely in 1948 when the legal positivist H. L. A. Hart presented the paper 'The ascription of responsibility and rights' at the Aristotelian Society [Hart, 1951]. Although Toulmin [1958/2003] rarely uses the term defeasible in *The uses of argument*, he is obviously an early adopter of the idea of defeasible reasoning, but he is not mentioned by Pollock [1987]. Like Pollock, he mentions Hart, but also another philosopher, David Ross, who applied the idea to ethics, recognizing that moral rules may hold prima facie, but can have exceptions.

In Pollock's approach [1987], 'reasoning' is conceived as a process that proceeds in terms of reasons. Pollock's reasons correspond to the constellations of premises and a conclusion which argumentation theorists and logicians call (elementary) arguments. Pollock distinguishes two kinds of reasons:

1. A reason is *non-defeasible* when it logically implies its conclusion;

2. A reason P for Q is *prima facie* when there is a circumstance R such that $P \wedge R$ [where '$\wedge$' denotes logical conjunction] is not a reason for the reasoner to believe Q. R is then a *defeater* of P as a reason for Q.

Note how closely related the idea of a prima facie reason is to non-monotonic inference: Q can be concluded from P, but not when there is additional information R.

Pollock's standard example is about an object that looks red. 'X looks red to John' is a reason for John to believe that X is red, but there can be defeating circumstances, for instance, when there is a red light illuminating the object. See Figure 1.

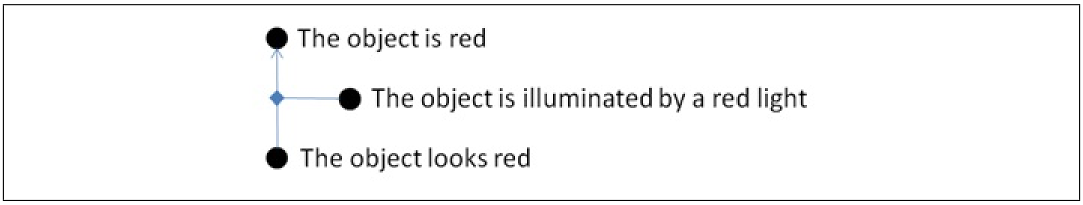

Figure 1: Pollock's red light example

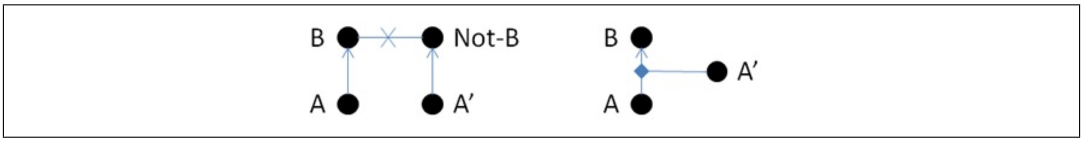

Figure 2: A rebutting defeater and an undercutting defeater

Pollock argues for the existence of two kinds of defeaters: 'rebutting' and 'undercutting defeaters.' A defeater is rebutting when it is a reason for the opposite conclusion (Figure 2, left). Undercutting defeaters attack the connection between the reason and the conclusion, and not the conclusion itself (Figure 2, right). The example about looking red concerns an undercutting defeater since when there is a red light it is not attacked that the object is red, but merely that the object's looking red is a reason for its being red.

A key element in Pollock's work on defeasible reasoning is the development of a theory of warrant. Pollock uses the term warrant as follows: a proposition is warranted in an epistemic situation if and only if an ideal reasoner starting in that situation would be justified in believing the proposition. Here justification is based on the existence of an undefeated argument with the proposition as conclusion. Pollock has developed his theory of warrant in a series of publications which formed the basis of his 1995 book *Cognitive Carpentry*. As a background for his approach to the structure of defeasible reasoning, Pollock provides a list of important classes of specific reasons: reasons based on logical deduction, perception, memory, statistics, or induction. Pollock's theory is embedded in what he called the OSCAR project [Pollock, 1995]. This project aims at the implementation of a rational agent. In the project Pollock addresses both theoretical (epistemic) and practical reasoning.[27]

In a theory of defeasible reasoning based on arguments that can defeat each other, such as Pollock's, the question needs to be considered which arguments can defeat which other arguments. Different forms of argument defeat can be distinguished:

[27]See Hitchcock [2001; 2002] for a survey and a discussion of the OSCAR project for those interested in argumentation. Hitchcock also gives further information about Pollock's work on practical reasoning, i.e., reasoning concerning what to do.

1. An argument can be *undermined*. In this form of defeat, the premises or assumptions of an argument are attacked.[28] Cf. the denial of the premises of an argument.

2. An argument can be *undercut*. In this form of defeat, the connection between a (set of) reason(s) and a conclusion in an argument is attacked. Cf. Pollock's undercutting defeaters.

3. An argument can be *rebutted*. In this form of defeat, an argument is attacked by giving an argument for an opposite conclusion. Cf. Pollock's rebutting defeaters.

4. An argument can be defeated by *sequential weakening*. Then each step in an argument is correct, but the argument breaks down when the steps are chained. An example is an argument based on the sorites paradox [Verheij 1996a, p. 122f.]:

   This body of grains of sand is a heap.
   So, this body of grains of sand minus 1 grain is a heap.
   So, this body of grains of sand minus 2 grains is a heap.
   ...
   So, this body of grains of sand minus n grains is a heap.

5. An argument can be defeated by parallel strengthening. This kind of defeat is associated with what has been called the 'accrual of reasons.' When reasons can accrue, it is possible that different reasons for a conclusion are together stronger than each reason separately. For instance, having robbed someone and having injured someone can be separate reasons for convicting someone. But when the suspect is a minor first offender, these reasons may each by itself be rebutted. On the other hand when a suspect has both robbed someone and also injured that person, the reasons may accrue and outweigh the fact that the suspect is a minor first offender. The argument for not punishing the suspect based on the reason that he is a minor first offender is defeated by the 'parallel strengthening' of the two arguments for punishing him.

[28]This form of defeat is the basis of Bondarenko *et al.* [1997]. We shall here not elaborate on the distinction between premises and assumptions. One way of thinking about assumptions is to see them as defeasible premises.

Building on experiences in the ASPIC project,[29] the recent state-of-the-art ASPIC+ system for the formal modelling of defeasible argumentation [Prakken, 2010][30] uses the first three kinds of defeat. The final two kinds of defeat are distinguished by Verheij [1996a, p. 122f.]. Pollock considered the accrual of reasons to be a natural idea, but argued against it [1995, p. 101f.]. More recent discussions of the accrual of reasons are to be found in Prakken [2005]; Gómez Lucero *et al.* [2009; 2013], and D'Avila Garcez *et al.* [2009, p. 155f.].

4 Argumentation and the structure of arguments in formal and computational perspective

4.1 Abstract argumentation

Phan Minh Dung's 1995 paper 'On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games' in the journal *Artificial Intelligence* [Dung, 1995] reformed the formal study of non-monotonic logic and defeasible reasoning. By his focus on argument attack as an abstract formal relation, Dung gave the field of study a mathematical basis that inspired many new insights. Dung's approach and the work inspired by it are generally referred to as abstract argumentation.

Dung's paper is strongly mathematically oriented, and has led to intricate formal studies. However, the mathematical tools used by Dung are elementary, hence various concepts studied by Dung can be explained without going into much formal detail.

The central innovation of Dung's 1995 paper is that he started the formal study of the attack relation between arguments, thereby separating the properties depending exclusively on argument attack from any concerns related to the structure of the arguments. Mathematically speaking, the argument attack relation is a directed graph, the nodes of which are the arguments, whereas the edges represent that one argument attacks another. Such a directed graph is called an argumentation framework. Figure 3 shows an example of an argumentation framework, with the dots representing arguments, and the arrows (ending in a cross to emphasize the

[29]The ASPIC project (full name: Argumentation Service Platform with Integrated Components) was supported by the EU 6th Framework Programme and ran from January 2004 to September 2007. In the project, academic and industry partners cooperated in developing argumentation-based software systems.

[30]Prakken [2010] speaks of ways of attack, where argument defeat is the result of argument attack.

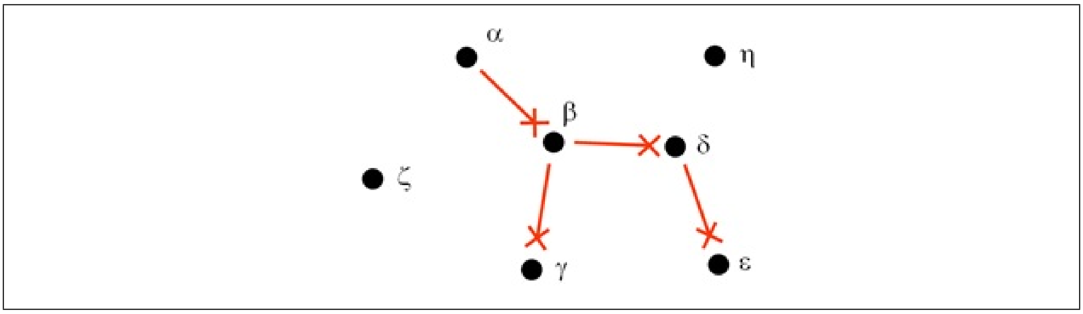

Figure 3: An argumentation framework representing attack between arguments

attacking nature of the connection[31]) representing argument attack.

In Figure 3, the argument α attacks the argument β, which in turn attacks both γ and δ, etc.

Dung's paper consists of two parts, corresponding to two steps in what he refers to as an 'analysis of the nature of human argumentation in its full generality' [Dung, 1995, p. 324]. In the first step, Dung develops the theory of argument attack and how argument attack determines argument acceptability. In the second part, he evaluates his theory by two applications, one consisting of a study of the logical structure of human economic and social problems, the other comprising a reconstruction of a number of approaches to non-monotonic reasoning, among them Reiter's and Pollock's. Notwithstanding the relevance of the second part of the paper, the paper's influence is largely based on the first part about argument attack and acceptability.

In Dung's approach, the notion of an 'admissible set of arguments' is central. A set of arguments is admissible if two conditions obtain:

1. The set of arguments is conflict-free, i.e., does not contain an argument that attacks another argument in the set (nor self-attacking arguments).

2. Each argument in the set is acceptable with respect to the set, i.e., when an argument in the set is attacked by an argument (which by (1) cannot be in the set itself), the set contains an argument that attacks the attacker.

In other words, a set of arguments is admissible if it contains no conflicts and if the set also can defend itself against all attacks. An example of an admissible set of arguments for the framework in Figure 3 is $\{\alpha, \gamma\}$. Since α and γ do not attack one another the set is conflict-free. The argument α is acceptable with respect to the set since it is not attacked, so that it needs no defence. The argument γ is

[31]This is especially helpful when also supporting connections are considered; see Section 4.2.

also acceptable with respect to $\{\alpha, \gamma\}$: the argument γ needs a defence against the attack by β, which defence is provided by the argument α, α being in the set. The set $\{\alpha, \beta\}$ is not admissible since it is not conflict-free. The set $\{\gamma\}$ is not admissible since it does not contain a defence against the argument β, which attacks argument γ.

Admissible sets of arguments can be used to define argumentation notions of what counts as a proof or a refutation.[32] An argument is '(admissibly) provable' when there is an admissible set of arguments that contains the argument. A minimal such set can be regarded as a kind of 'proof' of the argument, in the sense that the arguments in such a set are just enough to successfully defend the argument against counterarguments. An argument is '(admissibly) refutable' when there is an admissible set of arguments that contains an argument that attacks the former argument. A minimal such set can be regarded as a kind of 'refutation' of the attacked argument.

Dung speaks of the basic principle of argument acceptability using an informal slogan: the one who has the last word laughs best. The argumentative meaning of this slogan can be explained as follows. When someone makes a claim, and that is the end of the discussion, the claim stands. But when there is an opponent raising a counterargument attacking the claim, the claim is no longer accepted—unless the proponent of the claim provides a counterattack in the form of an argument attacking the counterargument raised by the opponent. Whoever has raised the last argument in a sequence of arguments, counterarguments, counter-counterarguments, etc., is the one who has won the argumentative discussion.

Formally, Dung's argumentation principle 'the one who has the last word laughs best' can be illustrated using the notion of an 'admissible set of arguments'. In Figure 3, a proponent of the argument γ has the last word and laughs best, since the only counterargument β is attacked by the counter-counterargument α. Formally, this is captured by the admissibility of the set $\{\alpha, \gamma\}$.

Although the principle of argument acceptability and the concept of an admissible set of arguments seem straightforward enough, it turns out that intricate formal puzzles loom. This has to do with two important formal facts:

1. It can happen that an argument is both admissibly provable and refutable.

2. It can happen that an argument is neither admissibly provable nor refutable.

The two argumentation frameworks shown in Figure 4 provide examples of these two facts. In the cycle of attacks on the left, consisting of two arguments α and β,

[32] In the following, we make use of terminology proposed by Verheij [2007].

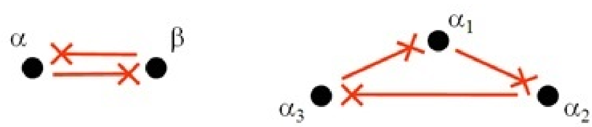

Figure 4: Arguments attacking each other in cycles

each of the arguments is both admissibly provable and admissibly refutable. This is a consequence of the fact that the two sets $\{\alpha\}$ and $\{\beta\}$ are each admissible. For instance, $\{\alpha\}$ is admissible since it is conflict-free and can defend itself against attacks: the argument α itself defends against its attacker α. By the admissibility of the set $\{\alpha\}$, the argument α is admissibly probable, and the argument β admissibly refutable.

The cycle of attacks on the right containing three arguments α_1, α_2 and α_3 is an example of the second fact above, the fact that it can happen that an argument is neither admissibly provable nor refutable. This follows from the fact that there is no admissible set that contains (at least) one of the arguments α_1, α_2 or α_3. Suppose that the argument α_3 is in an admissible set. Then the set should defend α_3 against the argument α_2, which attacks α_3. This means that α_1 should also be in the set, since it is the only argument that can defend α_3 against α_2. But this is not possible, because then α_1 and α_3 are both in the set, introducing a conflict in the set. As a result, there is only one admissible set: the empty set, which contains no arguments at all. We conclude that no argument is admissibly provable or admissibly refutable.

A related formal issue is that when two sets of arguments are admissible, it need not be the case that their union is admissible. The framework on the left in Figure 4 is an example. As we saw, the two sets $\{\alpha\}$ and $\{\beta\}$ are both admissible, but their union $\{\alpha, \beta\}$is not, since it contains a conflict. This has led Dung to propose the notion of a preferred extension of an argumentation framework, which is an admissible set that is as large as possible, in the sense that adding elements to the set makes it not admissible. The framework in Figure 3 has one preferred extension: the set $\{\alpha, \gamma, \delta, \zeta, \eta\}$. The framework in Figure 4 on the left has two preferred extensions $\{\alpha\}$ and $\{\beta\}$, the one on the right has one: the empty set.

Some preferred extensions have a special property, namely that each argument that is not in the set is attacked by an argument in the set. Such an extension is called a stable extension. Stable extensions are formally defined as conflict-free sets that attack each argument not in the set. It follows from this definition that a stable extension is also a preferred extension.

The preferred extension $\{\alpha, \gamma, \delta, \zeta, \eta\}$ of the framework in Figure 3, for instance, is stable, since the arguments β and ϵ, which are the only ones that are not in the set, are attacked by arguments in the set, α and δ, respectively. The preferred extensions $\{\alpha\}$ and $\{\beta\}$ of Figure 4 (left) are also stable. The preferred extension of Figure 4 (right), the empty set, is not stable, since none of the arguments α_1, α_2 and α_3 is attacked by an argument in the set. This example shows that there exist preferred extensions that are not stable. It also shows that there are argumentation frameworks that do not have a stable extension. In contrast, every argumentation framework has at least one preferred extension (which can be the empty set).

The concepts of preferred and stable extension of an argumentation framework can be regarded as different ways to interpret a framework, and therefore they are often referred to as 'preferred semantics' and 'stable semantics.' Dung [1995] proposed two other kinds of semantics: 'grounded semantics' and 'complete semantics,' and following his paper several additional kinds of semantics have been proposed (see Baroni *et al.* [2011], for an overview). By the abstract nature of argumentation frameworks, formal questions about the computational complexity of related algorithms and formal connections with other theoretical paradigms came within reach (see, e.g., [Dunne & Bench-Capon, 2003; Dunne, 2007; Egly *et al.*, 2010]).

Dung's original definitions are in terms of mathematical sets. An alternative way of studying argument attack is in terms of labelling. Arguments are marked with a label, such as 'Justified' or 'Defeated' (or IN/OUT, +/-, 1/0, 'Warranted'/ 'Unwarranted,' etc.), and the properties of different kinds of labelling are studied in the field. For instance, the notion of a stable extension corresponds to the following notion in terms of labelling:

> A *stable labelling* is a function that assigns one label 'Justified' or 'Defeated' to each argument in the argumentation framework such that the following property holds: an argument α is labelled 'Defeated' if and only if there is an argument β that attacks α and that is labelled 'Justified.'

A stable extension gives rise to a stable labelling by labelling all arguments in the extension 'Justified' and all other arguments 'Defeated.' A stable labelling gives rise to a stable extension by considering the set of arguments labelled 'Justified.'

The idea of labelling arguments can be thought of in analogy with the truth functions of propositional logic, where propositions are labelled with truth-values 'true' and 'false' (or 1/0, t/f, etc.). In the formal study of argumentation, labelling techniques predate Dung's abstract argumentation [1995]. Pollock [1994] uses labelling techniques in order to develop a new version of a criterion that determines warrant.

Verheij [1996b] applied the labelling approach to Dung's abstract argumentation frameworks. He uses argument labelling also as a technique to formally model which arguments are taken into account: in an interpretation of an abstract argumentation framework, the arguments that are assigned a label can be regarded as the ones taken into account, whereas the unlabelled arguments are not considered. Using this idea, Verheij defines two new kinds of semantics: the 'stage semantics' and the 'semi-stable semantics.'[33] Other authors using a labelling approach are Jakobovits and Vermeir [1999] and Caminada [2006]. The latter author translated each of Dung's extension types into a mode of labelling.

As an illustration of the labelling approach, we give a labelling treatment of the grounded extension of an argumentation framework as defined by Dung.[34] Consider the following procedure in which gradually labels are assigned to the arguments of an argumentation framework:

1. Apply the following to each unlabelled argument α in the framework: if the argument α is only attacked by arguments that have been labelled 'Defeated' (or perhaps is not attacked at all), label the argument α as 'Justified.'

2. Apply the following to each unlabelled argument α in the framework: if the argument α is attacked by an argument that has been labelled 'Justified,' label the argument α as 'Defeated.'

3. If step 1 and/or step 2 have led to new labelling, go back to step 1; otherwise stop.

When this procedure is completed (which always happens after a finite number of steps when the argumentation framework is finite), the arguments labelled 'Justified' constitute the grounded extension of the argumentation framework. Consider, for instance, the framework of Figure 3. In the first step, the arguments α, ζ and η are labelled 'Justified.' The condition that all arguments attacking them have been 'Defeated' is vacuously fulfilled, since there are no arguments attacking them. In the second step the argument β is labelled 'Defeated', since α has been labelled 'Justified.' Then a second pass of step 1 occurs and the arguments γ and δ are labelled 'Justified,' since their only attacker β has been labelled 'Defeated.' Finally, the argument ϵ is labelled 'Defeated,' since δ has been labelled 'Justified.' The arguments α, γ, δ, ζ and η (i.e., those labelled 'Justified') together form the grounded

[33] In establishing the concept Verheij [1996b] used the term admissible stage extensions. The now standard term semi-stable extension was proposed by Caminada [2006].

[34] Dung's own definition of grounded extension, which does not use labelling, is not discussed here.

extension of the framework. Every argumentation framework has a unique grounded extension. In the framework of Figure 3, the grounded extension coincides with the unique preferred extension that is also the unique stable extension. The framework in Figure 4 (left) shows that the grounded extension is not always a stable or preferred extension. Its grounded extension is here the empty set, but its two preferred and stable extensions are not empty.

4.2 Arguments with structure

Abstract argumentation, discussed in the previous subsection, focuses on the attack relation between arguments, abstracting from the structure of arguments. We now discuss various themes related to the structure of arguments for and against conclusions, and how it has been studied: arguments and specificity, the comparison of conclusive force, arguments with prima facie assumptions, arguments and classical logic, and the combination of support and attack.

Argument specificity An early theme in the formal study of argumentation was that of 'argument specificity' in relation to the resolution of a conflict between arguments. The key idea connecting arguments and specificity is that when two arguments are conflicting, with one of them being based on more specific information, the more specific argument wins the conflict, and defeats the more general argument.

Guillermo Simari and Ronald Loui [1992] have provided a mathematical formalization of this connection between arguments and specificity, taking inspiration from Poole's [1985] work in non-monotonic logic, and connecting to Pollock's work on argumentative warrant. In their proposal, an argument is a pair (T, h), with T being a set of defeasible rules that are applied to arrive at the argument's conclusion h given the argument's premises (formalized in the background knowledge). Arguments are assumed to be consistent, in the sense that no contradiction can be derived (not even defeasibly). Also arguments are assumed to be minimal, in the sense that all rules are needed to arrive at the conclusion. Formally, for an argument (T, h), it holds that when T' is the result of omitting one or more rules in T, the pair (T', h) is not an argument. Two arguments (T, h) and (T', h') disagree when h and h' are logically incompatible, given the background knowledge. An argument (T, h) counter-argues an argument (T', h') if (T, h) disagrees with an argument (T'', h'') that is a sub-argument of (T', h'), i.e., T'' is a subset of T'. An argument (T, h) defeats an argument (T', h') when (T, h) disagrees with a sub-argument of (T', h') that is strictly less specific. Simari and Loui's approach has been developed further—with applications in artificial intelligence, multi-agent systems, and logic by the Bahia Blanca group, led by Simari (e.g., [García & Simari, 2004; Chesñevar *et al.*, 2004; Falappa *et al.*, 2002]). García and Simari [2004] show the close connection between argumentation and logic programming that was also an inspiration for Dung [1995].

Conclusive force A second theme connected to arguments and their structure is conclusive force. Arguments that have more conclusive force will survive a conflict more easily than arguments with less conclusive force. One idea that connects conclusive force with argument defeat is the weakest link principle, which Pollock characterizes as follows:

The degree of support of the conclusion of a deductive argument is the minimum of the degrees of support of its premises [1995, p. 99].

Pollock presents the weakest link principle as an alternative to a Bayesian approach, which he rejects. Gerard Vreeswijk [1997] has proposed an abstract model of argumentation with defeasible arguments that focuses on the comparison of the conclusive force of arguments. In his model, conclusive force is not modelled directly but as an abstract comparison relation that expresses which arguments have more conclusive force than which other arguments. Vreeswijk defines an abstract argumentation system as a triple $(L, R, \leq)$, where L is a set of sentences expressing the claims made in an argument, R is a set of defeasible rules allowing the construction of arguments, and $\leq$ represents the conclusive force relation between arguments. The rules come in two flavours: strict and defeasible. Arguments are constructed by chaining rules. A set of arguments Σ is a defeater of an argument α if Σ and α are incompatible (i.e., imply an inconsistency), and α is not an underminer of Σ. An argument α is an underminer of a set of arguments Σ if Σ contains an argument β that has strictly lower conclusive force than α. Whereas Dung's [1995] system is abstract by its focus on argument attack, Vreeswijk's proposal is abstract in particular also because the conclusive force relation is left unspecified. Vreeswijk gives the following examples of conclusive force relations:

1. *Basic order.* In this order, a strict argument has more conclusive force than a defeasible argument. In a strict argument, no defeasible rule is used.

2. *Number of defeasible steps.* An argument has more conclusive force than another argument if it uses less defeasible steps. Vreeswijk remarks that this is not a very natural criterion, but it can be used to give formal examples and counterexamples.

3. *Weakest link.* Here the conclusive force relation on arguments is derived from an ordering relation on the rules. An argument has more conclusive force than another if its weakest link is stronger than the weakest link of the other.

4. *Preferring the most specific argument.* Of two defeasible arguments, one has more conclusive force than the other if the first has the premises of the second among its conclusions.

Prima facie assumptions A third theme related to arguments and their structure is arguments with prima facie assumptions. In particular, the defeat of arguments can be the result of prima facie assumptions that are successfully attacked. In their abstract, argumentation-theoretic approach to default reasoning, Bondarenko, Dung, Kowalski, and Toni [1997] use such an approach. Using a given deductive system (L, R) that consists of a language L and a set of rules R, so-called 'deductions' are built by the application of rules. Given a deductive system (L, R), an assumption-based framework is then a triple $(T, Ab, \text{Contrary})$, where T is a set of sentences expressing the current beliefs, Ab expresses assumptions that can be used to extend T, and Contrary is a mapping from the language to itself that expresses which sentences are contraries of which other sentences. Bondarenko and colleagues define a number of semantics (similar to Dung's 1995 in the context of abstract argumentation). For instance, a stable extension is a set of assumptions Δ such that the following properties hold:

1. Δ is closed, meaning that Δ contains all assumptions that are logical consequences of the beliefs in T and Δ itself.

2. Δ does not attack itself, meaning that there is no deduction from the beliefs in T and Δ with a contrary of an element of Δ as conclusion.

3. Δ attacks each assumption not in Δ, meaning that, for every assumption outside Δ, there is a deduction from T and Δ with a contrary of that assumption as conclusion.

Verheij [2003a] has also developed an assumption-based model of defeasible argumentation. In contrast with Bondarenko *et al.* [1997], in Verheij's system, the rules from which arguments are constructed are part of the prima facie assumptions. Technically, the rules have become conditionals of the underlying language. As a result, it can be the issue of an argument whether some proposition supports another proposition. In this way, Pollock's undercutting defeaters can be modelled as an attack on a conditional. Pollock's example of an object that looks red (Section 3.4) is formalized using two conditional sentences:

$\texttt{looks\_red} \leadsto \texttt{is\_red}$
$\texttt{red\_light} \leadsto \times(\texttt{looks\_red} \leadsto \texttt{is\_red})$

The first expresses the conditional prima facie assumption that if something looks red, it is red. The second expresses an attack on this prima facie assumption: when there is a red light illuminating the object, it no longer holds that if the object looks red, it is red. The sentences illustrate the two connectives of the language: one to express the conditional ($\leadsto$), the other to express what is called dialectical negation ($\times$). The two conditional sentences correspond exactly to two graphical elements in Figure 1: the first to the arrow connecting the reason and the conclusion, the second, nested, conditional to the arrow (ending in a diamond) that expresses the attack on the first conditional. This isomorphism between formal structures of the language and graphical elements has been used for the diagrams supported by the argumentation software ArguMed [Verheij, 2005b; see Section 4.5]).

The use of assumptions raises the question how they are related to an argument's ordinary premises. Assumptions can be thought of as the defeasible premises of an argument, and as such they are akin to defeasible rules[35] with an empty antecedent. The Carneades framework [Gordon *et al.*, 2007] distinguishes three kinds of argument premises: ordinary premises, presumptions (much like the prima facie assumptions discussed here) and exceptions (which are like the contraries of assumptions).

Arguments and classical logic A fourth theme connected to arguments and their structure is how they are related to classical logic. In particular, the relation between classical logic and defeasible argumentation remains a puzzle. Above we already saw different attempts at combining elements of classical logic and defeasible argumentation. In Pollock's system, classical logic is one source of reasons. Often conditional sentences ('rules') are used to construct arguments by chaining them (e.g., [Vreeswijk, 1997]). Chaining rule applications is closely related to the inference rule modus ponens of classical logic. Verheij's [2003a] system gives conditionals which validate modus ponens a central place. Bondarenko *et al.* [1997] allow generalized rules of inference by their use of a contingent deductive system as starting point.

Besnard and Hunter [2008] have proposed to formalize arguments in classical logic entirely. For them, an argument is a pair (Φ, α), such that Φ is a set of sentences and α is a sentence, and such that Φ is logically consistent, Φ logically

[35] Some would object to the use of the term rules here. Rules are here thought of in analogy with the inference rules of classical logic. An issue is then that, as such, they are not expressed in the logical object language, but in a meta-language. In the context of defeasible reasoning and argumentation (and also in non-monotonic logic), this distinction becomes less clear. Often there is one logical language to express ordinary sentences, a second formal language (with less structure and/or less semantics, and therefore not usually referred to as 'logical') used to express the rules, and the actual meta-language that is used to define the formal system.

entails α (in the classical sense), and Φ is a minimal such set. (Note the analogy with the proposal by Simari and Loui [1992], discussed earlier.) Φ is the support of the argument, and α the claim. They define defeaters as arguments that refute the support of another argument. More formally, a defeater for an argument (Φ, α) is an argument (Ψ, β), such that β logically entails the negation of the conjunction of some of the elements of Φ. An undercut for an argument (Φ, α) is an argument (Ψ, β) where β is equal to (and not just entails) the negation of the conjunction of some of the elements of Φ. A rebuttal for an argument (Φ, α) is an argument (Ψ, β) such that $\beta \leftrightarrow \neg\alpha$ is a tautology. Besnard and Hunter give the following example [2008, p. 46]:

p	Simon Jones is a Member of Parliament.
p → ¬q	If Simon Jones is a Member of Parliament, then we need not keep quiet about details of his private life.
r	Simon Jones just resigned from the House of Commons.
r → ¬p	If Simon Jones just resigned from the House of Commons, then he is not a Member of Parliament.
¬p → q	If Simon Jones is not a Member of Parliament, then we need to keep quiet about details of his private life.

Then ({p, p → ¬q}, ¬q) is an argument with the argument (r, r → ¬p, ¬p) as an undercut and the argument (r, r → ¬p, ¬p → q, q) as a rebuttal.

Besnard and Hunter focus on structural properties of arguments, in part because of the diversity of proposals for semantics (see Section 4.1). For instance, when they discuss these systems, they note that the semantic conceptualization of such systems is not as clear as the semantics of classical logic, which is the basis of their framework [p. 221, also p. 226]. At the same time, they note that knowledge representation can be simpler in systems based on defeasible logic (see below) or inference rules.

Combining support and attack A fifth and final theme discussed here in connection with arguments and their structure is how support and attack are combined. In several proposals, support and attack are combined in separated steps. In the first step, argumentative support is established by constructing arguments for conclusions from a given set of possible reasons or rules (of inference). The second step determines argumentative attack. Attack is, for instance, based on defeaters or on

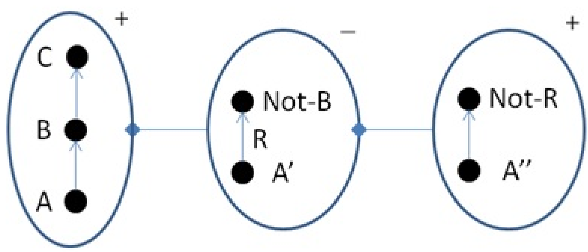

Figure 5: Supporting arguments that attack each other

the structure of the supporting arguments in combination with a preference relation on arguments. In the third and final step, it is determined which arguments are warranted or undefeated. We already saw that several criteria have been proposed (e.g., Pollock's gradual development of criteria for argumentative warrant, and Dung's abstract argumentation semantics).

An example of this modelling style is depicted in Figure 5. Three supporting arguments are shown. The first on the left shows that A supports B, which in turn supports C. In the middle of the figure, this argument is attacked by a second argument, which reasons from A' for Not-B (hence against B). This argument is in turn attacked by a third argument, which reasons from A'' against the support relation R between A' and Not-B. Using the terminology of Section 3.4, the first subargument of the first argument is rebutted by the second, which is undercut by the third. The arguments are marked with a + sign when they are warranted, and a – sign when they are defeated (which can be thought of as a variant of the labelling approaches of Section 4.1). The argument on the right is warranted, since it is not attacked. As a result, the middle argument is defeated, since it is attacked by a warranted argument. The left argument is then also warranted, since its only attacker is defeated. (See the procedure for computing the grounded extension of an argumentation framework discussed in Section 4.1.)

In this approach, the relation with Dung's abstract argumentation is that we can abstract from the structure of the supporting arguments resulting in an abstract argumentation framework. For the three arguments in Figure 5, we obtain the abstract framework shown in Figure 6. In this example, the argumentation semantics is unproblematic at the abstract argument attack level since the grounded extension coincides with the unique preferred extension that is also stable. Special care is needed to handle parts of arguments. For instance, the middle argument has the premise A', which is not attacked, and should therefore remain undefeated.

Figure 6: The abstract argumentation framework associated with the example of Figure 5

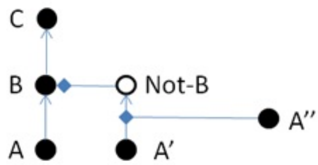

Figure 7: Arguments supporting and attacking conclusions

This type of combining support and attack is used in the ASPIC+ model [Prakken, 2010]. A second approach does not separate support and attack when combining them. Arguments are constructed from reasons for and against conclusions, which in turn determine whether a conclusion follows or not. Figure 7 models the same argumentative information as Figure 5, but now using this second approach.

Here the reason A'' undercuts the argument from A' to Not-B, so Not-B is not supported (indicated by the open circle). As a result, Not-B does not actually attack B, which is therefore justified by A and in turn justifies C.

In this approach, for instance, conditional sentences are used to express which reasons support or attack which conclusions. An example is Nute's defeasible logic [Nute, 1994; Antoniou *et al.*, 2001], which uses conditional sentences for the representation of strict rules and defeasible rules, and for defeater rules, which can block an inference based on a defeasible rule. Algorithms for defeasible logic have been designed with good computational properties. Another example of the approach is Verheij's DefLog [2003a], in which a conditional for the representation of support is combined with a negation operator for the representation of attack. A related proposal extending Dung's abstract argumentation frameworks by expressing both support and attack is bipolar argumentation [Cayrol & Lagasquie-Schiex, 2005; Amgoud *et al.*, 2008]. For DefLog and bipolar argumentation, generalisations of Dung's stable and preferred semantics are presented. DefLog has been used to formalize Toulmin's argument model [Verheij, 2005a].

A special case of the combination of support and attack occurs when the support and attack relations can themselves be supported or attacked. Indeed it can be at issue whether a reason supports or attacks a conclusion. The four ways of arguing

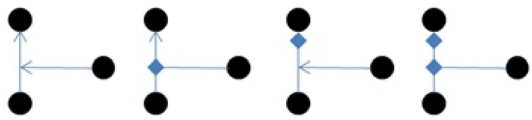

Figure 8: The four ways of arguing about support and attack

about support and attack are illustrated in Figure 8, from left to right: support of a support relation, attack of a support relation, support of an attack relation, and attack of an attack relation, respectively.

For instance, Pollock's undercutting defeaters can be thought of as attacks of a support relation (second from the left in Figure 8). In Verheij's DefLog [2003a; 2005b], the four ways are expressed using nested conditional sentences, in a way that extends the expressiveness of Dung's frameworks. Modgil [2009] has studied attacks of attacks (rightmost in 11) in a system that also extends Dung's expressiveness.

4.3 Formalizing argument schemes

Argumentation formalisms can only come to life when arguments are built from meaningful reasons. We already mentioned that Pollock made explicit which kinds of reasons he considered: deductive reasons, perception, memory, statistical syllogism, and induction.

An approach to the specification of meaningful kinds of reasons to construct arguments from is that of argument schemes, as they have been studied in argumentation theory. Argument schemes were already distinguished by Perelman and Olbrechts-Tyteca [1969].[36] In today's artificial intelligence research on argumentation, Douglas Walton's approach to argumentation schemes (his terminology) has been widely adopted (e.g., [Walton *et al.*, 2008].

Argument schemes can be thought of as analogues of the rules of inference of classical logic. An example of a rule of inference is, for instance, the following version of modus ponens:

> P
> If P, then Q
> Therefore: Q

[36] Although the term schème argumentative [argumentative scheme] was already used by Perelman and Olbrechts-Tyteca, according to Garssen [2001], van Eemeren *et al.* [1978; 1984] used the notion of argument(ation) scheme for the first time in its present sense. See also [van Eemeren and Grootendorst, 1992; Kienpointner, 1992; Walton, 1996; Walton *et al.*, 2008].

Whereas logical rules of inference, such as modus ponens, are abstract, strict, and (usually) considered to have universal validity, argumentation schemes are concrete, defeasible, and context-dependent. An example is the following scheme for witness testimony:

> Witness A has testified that P.
> Therefore: P

The use of this scheme is defeasible, as can be made explicit by asking critical questions, for instance:

> Wasn't A mistaken?
> Wasn't A lying?

A key reason why argument schemes have been taken up in artificial intelligence is that the critical questions associated with them correspond to defeating circumstances. For instance, the question whether A was mistaken gives rise to the defeater 'A was mistaken'.

Bex, Prakken, Reed and Walton [2003] applied the concept of 'argumentation schemes' to the formalization of legal reasoning from evidence. An example of a scheme in that paper (taken from [Walton, 1996]) is the following.

> Argument from expert opinion
> Source E is an expert in domain D.
> E asserts that proposition A is known to be true (false).
> A is within D.
> Therefore, A may plausibly be taken to be true (false).

This scheme has the following critical questions:

1. Expertise question: How credible is E as an expert source?
2. Field question: Is E an expert in D?
3. Opinion question: What did E assert that implies A?
4. Trustworthiness question: Is E personally reliable as a source?
5. Consistency question: Is A consistent with what other experts assert?
6. Backup evidence question: Is E's assertion based on evidence?

The authors elaborate on how these and other argumentation schemes related to evidential reasoning can be formalized.

From the perspective of artificial intelligence, the work on argumentation schemes of Walton and his colleagues can be regarded as contributions to the theory of knowledge representation. Gradually, a collection of argumentation schemes is being developed. When appropriate, a scheme is added, and existing schemes are adapted, for instance, by refining the scheme's premises or critical questions. This knowledge representation point of view is developed by Verheij [2003b], who like Bex *et al.* [2003] formalizes argumentation schemes as defeasible rules of inference. He notes that in Walton's work argumentation schemes sometimes take the form of small derivations, or sequences of argumentation schemes; or even of a small prototypical dialogue. To streamline the work on knowledge representation, Verheij proposes to treat argumentation schemes as consisting of four elements: Conclusion, Premises, Conditions of use, and Exceptions. The Exceptions correspond to answers to the critical questions of an argumentation scheme. By this representation format, it is also possible to consider different roles of critical questions: critical questions concerning a conclusion, a premise, a condition of use, or an exception.

Reed and Rowe [2004) have incorporated argumentation schemes in their Araucaria tool for the analysis of argumentative texts. Rahwan *et al.* [2007] have proposed formats for the integration of argumentation schemes in what is called the Semantic Web. The vision underlying the Semantic Web is that, when information on the Internet is properly tagged, it becomes possible to add meaning to such information that can be handled by a machine. For instance, when the Conclusion, Premises, Conditions of use, and Exceptions of an argumentation scheme are marked as such, software can be built that can handle these different elements of a scheme appropriately. Gordon, Prakken and Walton [2007] have integrated argumentation schemes in their Carneades model.

A fundamental issue concerning argumentation schemes is how to evaluate a scheme or set of schemes. When is a scheme good, under which circumstances? When is an adaptation appropriate? This issue is, for instance, discussed in Reed and Tindale [2010].

4.4 Formalizing argumentation dialogues

One reason why Toulmin's [2003] *The uses of argument* remains a thought-provoking study is his starting point that argument should be considered in its natural, critical, and procedural context. This starting point led him to propose that logic, in the sense of the theory of good argument, should be treated as 'generalized jurisprudence,' where a critical and procedural perspective on good argument is the norm. The critical and procedural sides of arguments come together in the study of argumentation dialogues.

The following is a fragment, taken from McBurney and Parsons [2002a], of an argumentation dialogue concerning the sale of a used car between a buyer (B) and seller (S), illustrating the study of argumentative dialogue in a computational setting:

> S: BEGIN(PERSUASION(Make);
> PERSUASION(Condition_of_Engine);
> PERSUASION(Number_of_Owners))
> S requests a sequence of three Persuasion dialogues over the purchase criteria Make, Condition of the Engine, and Number of Owners.
> B:
> AGREE(PERSUASION(Make);PERSUASION(Condition_of_Engine);
> PERSUASION(Number_of_Owners)) PERSUASION Dialogue 1 in the sequence of three opens.
> S: Argues that 'Make' is the most important purchase criterion, within any budget, because a typical car of one Make may remain in better condition than a typical car of another Make, even though older.
> B: Accepts this argument.
> PERSUASION Dialogue 1 closes upon acceptance of the proposition by B. PERSUASION Dialogue 2 opens.
> S: Argues that that 'Condition_of_Engine' is the next most important purchase criterion.
> B: Does not accept this. Argues that he cannot tell the engine condition of any car without pulling it apart. Only S, as the Seller, is able to tell this. Hence, B must use 'Mileage' as a surrogate for 'Condition_of_Engine.'
> PERSUASION Dialogue 2 closes with neither side changing its views: B does not accept 'Condition_of_Engine' as the second criterion, and S does not accept 'Mileage' as the second criterion. PERSUASION Dialogue 3 opens.

The fragment shows how dialogues about certain topics are opened and closed in relation to the arguments provided.

The formal and computational study of argumentation dialogues has primarily been performed in the fields of AI and law and of multi-agent systems, as addressed below.

In the field of AI and law, argumentation dialogues have been studied extensively (see [Bench-Capon *et al.*, 2004; 2009]). Ashley's [1990] HYPO, to be discussed more extensively in Section 5.2, takes a 3–ply dialogue model as starting point, in which a proponent makes a claim, which can be attacked by an opponent, and then defended by the proponent. An early AI and law conception of argumentation dialogue is

Thomas Gordon's [1993; 1995] *Pleadings game*. Gordon formalizes the pleading in a US-style civil law process, which is aimed at determining the legal and factual issues of a case. In the Pleadings Game, a proponent and opponent (in this setting referred to as 'plaintiff' and 'defendant') can concede, deny and defend claims, and also declare defeasible rules. Players can discuss the validity of a defeasible rule. Players are committed to the consequences of their claims, as prescribed by a non-monotonic logic underlying the Pleadings Game.

Other dialogue models of argumentation in AI and law have been proposed by Prakken and Sartor [1996; 1998], Hage *et al.* [1993], and Lodder [1999]. In Prakken and Sartor's approach [1996; 1998], dialogue models are presented as a kind of proof theory for their argumentation model. Prakken and Sartor interpret a proof as a dialogue between a proponent and opponent. An argument is justified when there is a winning strategy for the proponent of the argument. Hage *et al.* [1993] and Lodder [1999] propose a model of argumentation dialogues with the purpose of establishing the law in a concrete case. They are inspired by the idea of law as a pure procedure (though not endorsing it): when the law is purely procedural, there is no criterion for a good outcome of a legal procedure other than the procedure itself.

Some models emphasize that the rules of argumentative dialogue can themselves be the subject of debate. An actual example is a parliamentary discussion about the way in which legislation is to be discussed. In philosophy, Suber has taken the idea of self-amending games to its extreme by proposing the game of Nomic, in which the players can gradually change the rules.[37] Proposals to formalize such meta-argumentation include Vreeswijk [2000] and Brewka [2001], who have proposed formal models of argumentative dialogues allowing self-amendments.[38]

In an attempt to clarify how logic, defeasibility, dialogue and procedure are related, Henry Prakken [1997, p. 270f.] proposed to distinguish four layers of argumentation models. The first is the logical layer, which determines contradiction and support. The second layer is dialectical, which defines what counts as attack, counterargument, and also when an argument is defeated. The third layer is procedural and contains the rules constraining a dialogue, for instance, which moves parties can make, when parties can make a move, and when the dialogue is finished. The fourth and final layer is strategic. At this layer, one finds the strategies and heuristics used by a good, effective arguer.

Jaap Hage [2000] addresses the question of why dialogue models of argumentation became popular in the field of AI and law. He gives two reasons. The first is that legal reasoning is defeasible, and dialogue models are a good tool to study

[37] `http://en.wikipedia.org/wiki/Nomic`. See also Hofstadter [1996, chapter 4].

[38] See also the study of Nomic by Vreeswijk [1995a].

defeasibility. The second reason is that dialogue models are useful when investigating the process of establishing the law in a concrete case. Hage recalls the legal theoretic discussion about the law as an open system, in the sense that there can be disagreement about the starting points of legal arguments. As a result, the outcome of a legal procedure is indeterminate. A better understanding of this predicament can be achieved by considering the legal procedure as an argumentative dialogue.

Hage [2000] then discusses three functions of dialogue models of argumentation in AI and law. The first function is to define argument justification, in analogy with dialogical definitions of logical validity as can be found in the work by Lorenzen and Lorenz [1978]. In this connection, Hage refers to Barth and Krabbe's notion of the 'dialectical garb' of a logic as opposed to an axiomatic, inferential or model-theoretic garb [Barth & Krabbe, 1982, pp. 7–8]. Hage generalizes the idea of dialectical garb to what he refers to as battle of argument models of defeasible reasoning in which arguments attack each other, such as Loui's [1987], Pollock's [1987; 1994], Vreeswijk's [1993], Dung's [1995], and Prakken and Sartor's [1996]. Battle of argument models can or cannot be presented in a dialectical garb. In their dialectical garb, such models define the justification of an argument in terms of the existence of a winning strategy in an argumentative dialogue game.

The second function of dialogue models of argumentation that is distinguished by Hage is to establish shared premises. Proponent and opponent enter into a dialogue that leads to a shared set of premises. The conclusions that follow from these shared premises can be regarded as justified. In this category, Hage discusses Gordon's Pleadings Game, which we discussed above. Hage makes connections to legal theory, in particular to Alexy's [1978] procedural approach to legal justification, and the philosophy of truth and justification, in particular Habermas's [1973] consensus theory of truth, and Schwemmer's approach to justification, in which the basis of justification is only assumed as long as it is not actually questioned [Schwemmer & Lorenzen, 1973].

As a third and final function of dialogue models of argumentation in AI and law, Hage discusses the procedural establishment of law in a concrete case. In this connection, he discusses mediating systems, which are systems that support dialogues, instead of evaluating them. He uses Zeno [Gordon & Karacapilidis, 1997], Room 5 [Loui *et al.*, 1997] (see also Section 4.5) and DiaLaw [Lodder, 1999] as examples. Hage argues that regarding the law as purely procedural is somewhat counterintuitive, since there exist cases in which there is a clear answer, which can be known even without actually going through the whole procedure. Hage speaks therefore of the law as an imperfect procedure, in which the correctness of the outcome is not guaranteed.

Outside the field of AI and law, one further function of dialogue models of argumentation has been emphasized, namely that a dialogue perspective on argumentation can have computational advantages. For instance, argumentative dialogue can be used to optimize search, for instance, by cutting off dead ends or focusing on the most relevant issues. Vreeswijk [1995b] takes this assumption as the starting point of a paper:

> If dialectical concepts like argument, debate, and resolution of dispute are seemingly so important in practical reasoning, there must be some reason as to why these techniques survived as rulers of commonsense argument. Perhaps the reason is that they are just most suited for the job [Vreeswijk, 1995b, p. 307].

Vreeswijk takes inspiration from a paper by Loui [1998], which circulated in an earlier version since 1992. Loui emphasises the relevance of protocol, the assignment of burdens to parties, termination conditions, and strategy. A key idea is that argumentation dialogues are well-suited for reasoning in a setting of bounded resources (see also [Loui & Norman, 1995]).

Inspired by the computational perspective on argumentation, approaches to argumentative dialogue have been taken up in the field of multi-agent systems.[39] The focus in that field is on the interaction between autonomous software agents that pursue their own goals or goals shared with other agents. Since the actions of one agent can affect those of another, beyond control of an individual agent or the system as a whole, the kinds of problems when designing multi-agent software systems are of a different nature than those in the design of software where control can be assumed to be centralized. Computational models of argumentation have inspired the development of interaction protocols for the resolution of conflicts among agents and for belief formation. The typology of argumentative dialogue that has been proposed by Douglas Walton and Erik Krabbe [1995] has been especially influential.[40] In this typology, seven dialogue types are distinguished:

1. Persuasion, aimed at resolving or clarifying an issue;
2. Inquiry, aimed at proving (or disproving) a hypothesis;
3. Discovery, aimed at choosing the best hypothesis for testing;
4. Negotiation, aimed at a reasonable settlement all parties can live with;

[39] For an overview of the field of multi-agent systems see the textbook by Wooldridge [2009], which contains a chapter entitled 'Arguing.'

[40] The 2000 Symposium on Argument and Computation at Bonskeid House Perthshire, Scotland, organized by Reed and Norman, has been a causal factor. See Reed and Norman [2004].

5. Information-seeking, aimed at the exchange of information;
6. Deliberation, aimed at deciding the best available course of action;
7. Eristic, aimed at revealing a deeper basis of conflict.

In particular, the persuasion dialogue, starting with a conflict of opinion and aimed at resolving the issue by persuading a participant, has been extensively studied. An early persuasion system—focusing on persuasion in a negotition setting—is Sycara's Persuader system [1989]. Persuader, developed in the field of what was then called Distributed AI, uses the domain of labour negotiation as an illustration. An agent forms a model of another agent's beliefs and goals, and determines its actions in such a way that it influences the other agent. For instance, agents can choose a so-called 'threatening argument,' i.e., an argument that is aimed at persuading another agent to give up a goal. Here it is notable that in Walton and Krabbe's typology negotiation is a dialogue type different from persuasion.

Prakken [2006; 2009] gives an overview and analysis of dialogue models of persuasion. In a dialogue system, dialogues have a goal and participants. It is specified which kinds of moves participants can make, for instance, making claims or conceding. Participants can have specific roles, for instance, Proponent or Opponent. The actual flow of a dialogue is constrained by a protocol, consisting of rules for turn-taking and termination. Effect rules determine how the commitments of participants change after each dialogue move. Outcome rules define the outcome of the dialogue, by determining, for instance, in persuasion dialogues who wins the dialogue. These elements are common to all dialogue types. By specifying or constraining the elements, one generates a system of persuasion dialogue. In particular, the dialogue goal of persuasion dialogue consists of a set of propositions that are at issue and need to be resolved. Prakken formalizes these elements and then uses his analytic model to discuss several extant persuasion systems, among them Mackenzie's [1979] proposals, and Walton and Krabbe's [1995] model of what they call Permissive persuasion dialogue.

Sycara's Persuader system [1989] is a persuasion system applied to labour negotiation. Parsons, Sierra and Jennings [1998] also speak of negotiation as involving persuasion. Their model uses the Belief-Desire-Intention model of agents [Rao & Georgeff, 1995] and specifies logically how the beliefs, desires and intentions of the agents influence the process of negotiation.[41] Dignum, Dunin-Kęplicz and Verbrugge [2001] have studied the role of argumentative dialogue for the forming of coalitions

[41] A systematic overview of argumentation dialogue models of negotiation has been provided by Rahwan *et al.* [2003].

of agents that create collective intentions. Argumentation about what to do rather than about what is the case has been studied in a dialogue setting by Atkinson and colleagues [Atkinson *et al.*, 2005; 2006; Atkinson & Bench-Capon, 2007]. In this connection, it is noteworthy that Pollock's OSCAR model [1995] is an attempt to combine theoretical reasoning—about what to believe—with practical reasoning—about what to do—, though in a single agent, non-dialogical setting. Amgoud [2009] discusses the application of dialogical argumentation to decision making (see also [Girle *et al.*, 2004]). Deliberation has been studied by McBurney *et al.* [2007].

Several attempts have been made to systematize the extensive work on argumentation dialogue. Bench-Capon *et al.* [2000], for instance, propose a formal method for modelling argumentation dialogue. Prakken [2005b] provides a formal framework that can be used to study argumentation dialogue models with different choices of underlying argument model and reply structures. McBurney and Parsons [2002a; 2002b; 2009] have developed an abstract theory of argumentative dialogue in which syntactic, semantic, and pragmatic elements are considered.

4.5 Argumentation support software

When studying argumentation from an artificial intelligence perspective, it can be investigated how software tools can perform or support argumentative tasks. Some researchers in the field of argumentation in AI have openly addressed themselves to building an artificial arguer. The most prominent among them is John Pollock (see also Section 3.4), who titled one of his books about his OSCAR project ambitiously *How to build a person* [Pollock, 1989].[42] Most researchers however have not aimed at realizing the grand task of addressing the so-called 'strong AI' problem of building an intelligent artefact that can perform any intellectual task a human being can. Instead of building software mimicking human argumentative behaviour, the more modest aim of supporting humans performing argumentative tasks was chosen. A great deal of research has been aimed at the construction of argumentation support software. Here we discuss three recurring themes: argument diagramming in software, the integration of rules and argument schemes, and argument evaluation.[43]

Argument diagramming in software The first theme discussed is argument diagramming in software. In the literature on argumentation support software, much attention has been paid to argument diagramming. Different kinds of argument diagramming styles have been proposed, many inspired by non-computational research

[42]The book's subtitle adds modestly: A Prolegomenon.

[43]The reviews by Kirschner *et al.* [2003], Verheij [2005b], and Scheuer *et al.* [2010] provide further detail about argumentation support software.

on argument diagrams. We shall discuss three styles: boxes and arrows, boxes and lines, and nested boxes.

The first style of argument diagramming uses boxes and arrows. Argumentative statements are enclosed in boxes, and their relations indicated by arrows. A common use of arrows is to indicate the support relation between a reason and a conclusion. An example of a software tool that uses boxes and arrows diagrams is the Araucaria tool by Chris Reed and Glenn Rowe [2004] (Figure 9[44]). The Araucaria tool has been designed for the analysis of written arguments. Vertical arrows indicate reasons and their conclusions, and horizontal bi-directional arrows indicate conflicts between statements. The Araucaria software was one step in the development by the Dundee Argumentation Research Group, led by Reed, of open source argumentation software. For this purpose, a representation format, called the Argument mark-up language (AML), has been developed that allows for the exchange of arguments and their analyses using contemporary Internet technology. The format also allows for the exchange of sets of argument schemes (see Section 4.3) that can be used for argument analysis. Connected developments concerning machine-readable argument representation formats are the Argument interchange format [Chesñevar *et al.*, 2006] and ArgDF, a proposal for a language allowing for a World wide argument web [Rahwan *et al.*, 2007]. One aim of the latter work is to develop classification systems for arguments, using ontology development techniques in Artificial Intelligence. In AI, an 'ontology' is a systematic conceptualization of a domain, often taking the form of a hierarchical system of concepts and their relations.

Another example of a system using boxes and arrows is the Hermes system [Karacapilidis & Papadias, 2001], an extension of the Zeno system [Gordon & Karacapilidis, 1997]. Both Hermes and Zeno have been inspired by the IBIS approach. In IBIS, an abbreviation of Issue-Based Information Systems [Kunz & Rittel, 1970], problems are analysed in terms of issues, questions of fact, positions, and arguments. The focus is on what Rittel and Webber [1973] call wicked problems: problems with no definitive formulation, and no definitive solutions. Hence a goal of IBIS and systems such as Hermes and Zeno is to support the identification, structuring and settling of issues.

The second style of argument diagramming uses boxes and lines. In a boxes and lines style of argument diagramming, argumentative statements are depicted in boxes and their relations are indicated by (undirected) lines between them. This diagramming style abstracts from the directionality between statements, for instance, from a reason to a conclusion, or from a cause to an event. An example of a tool using the boxes and lines style is the Belvedere system [Suthers *et al.*, 1995; Suthers,

[44]Source: `http://staff.computing.dundee.ac.uk/creed/araucaria/`.

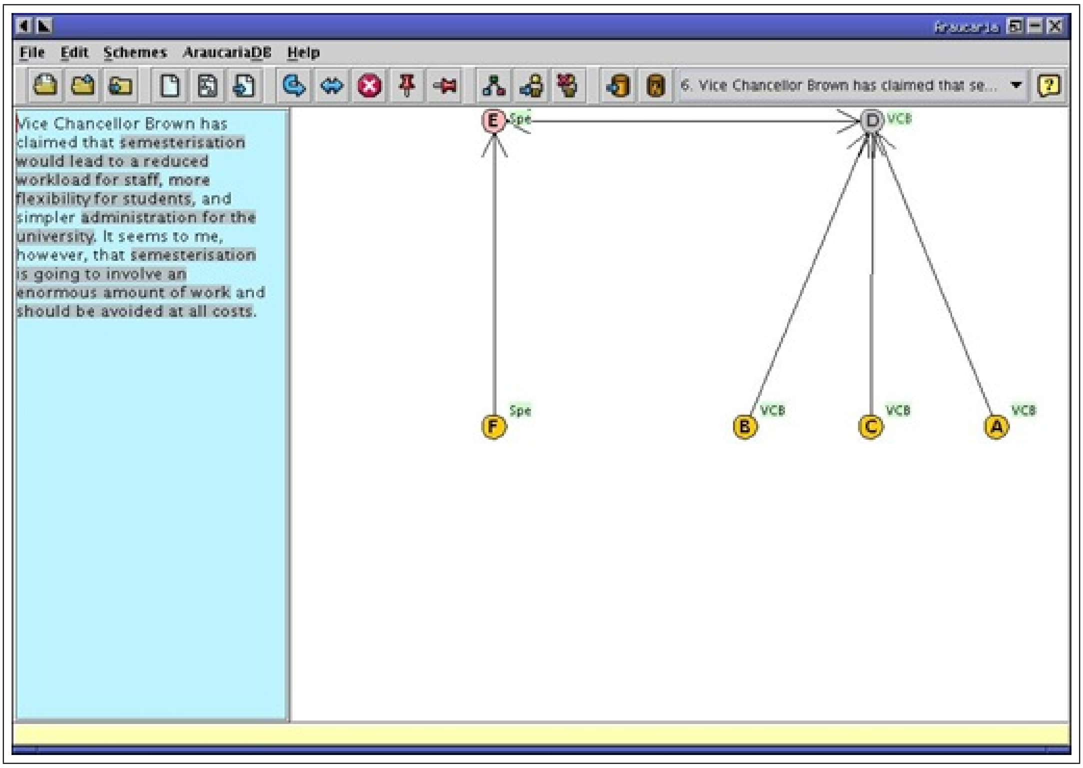

Figure 9: Boxes and arrows diagramming: The Araucaria system

1999]. A goal of the system was to stimulate the critical discussion of science and public policy issues by middle school and high-school students, taking the cognitive limitations of the intended users into account. Such limitations include difficulty in focusing attention, lack of domain knowledge, and lack of motivation. In early versions, the diagrams were richly structured: there were links for support, explanation, causation, conjunction, conflict, justification, and undercutting. Link types could be distinguished graphically and by label. To prevent unproductive discussions about which structure to use, the graphical representation was significantly simplified in later versions [Suthers, 1999]. Two types of statements were distinguished: data and hypotheses; and two link types: expressing a consistency and an inconsistency relation between statements. Figure 10[45] shows an example of a Belvedere screen using an even further simplified format with one statement type and one link type.

The third style of argument diagramming uses nested boxes. In this style, too, the argumentative statements are enclosed in boxes, but their relationships are indi-

[45]Source: `http://belvedere.sourceforge.net/`.

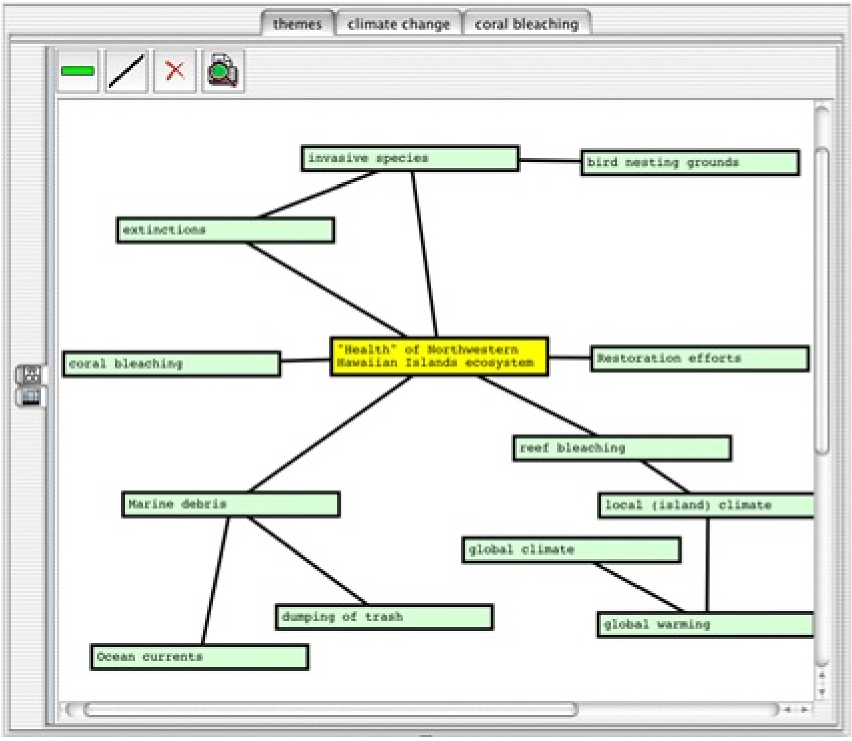

Figure 10: Boxes and lines diagramming: the Belvedere 4.1 system

cated by the use of nesting. An example of the use of nested boxes is the Room 5 tool designed by Loui, Norman and a group of students [Loui *et al.*, 1997]. The Room 5 system aimed at the collaborative public discussion of pending Supreme Court cases. It was web-based, which is noteworthy as the proposal predates Google and Wikipedia. In its argument-diagramming format, a box inside a box expresses support, and a box next to a box indicates attack. In the argument depicted in the Room 5 screen shown in Figure 11[46], for instance, the punishability of John is supported by the reason that he has stolen a CD, and attacked by the reason that he is a minor first offender.

The integration of rules and argument schemes A second theme concerning the design of argumentation support software is the integration of rules and argument schemes. The integration of rules and argument schemes in argument diagramming software has been addressed in different ways: by the use of schematic

[46]Screenshot of Room 5, as shown in Verheij [2005b]. See also Bench-Capon *et al.* [2012].

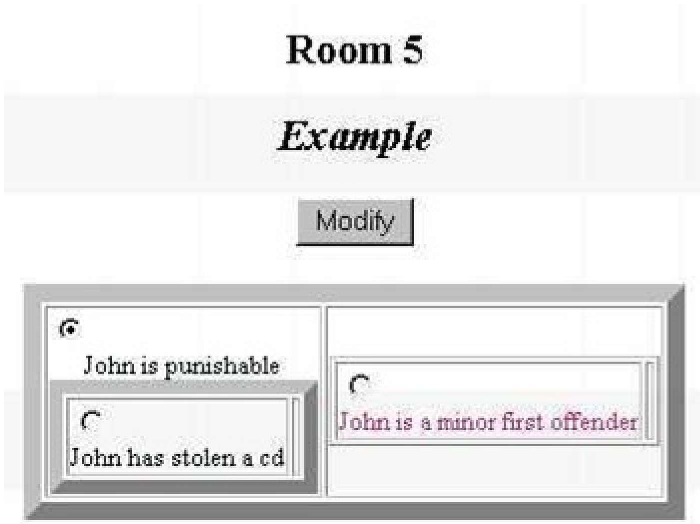

Figure 11: Nested boxes diagramming: the Room 5 system

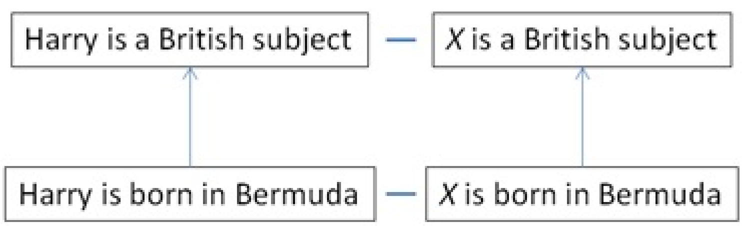

Figure 12: An elementary argument step as an instance of a schematic argument

arguments, conditional sentences, nested arrows and rule nodes. Consider, for instance, the elementary argument that Harry is a British subject because he is born in Bermuda (borrowed from Toulmin), and its underlying rule (or 'warrant' in Toulmin's terminology) that people born in Bermuda are British subjects.

A first approach is to consider such an argument as an instance of a scheme that abstracts from the person Harry in the argument. In Figure 12, an associated schematic argument is shown to the right of the argument about Harry. In the schematic argument, X appears as a variable that serves as the placeholder of someone's name. In software, the schematic argument is normally not shown graphically.

A second approach uses conditional sentences. The conditional sentence that expresses the connection between reason and conclusion is made explicit as an auxiliary

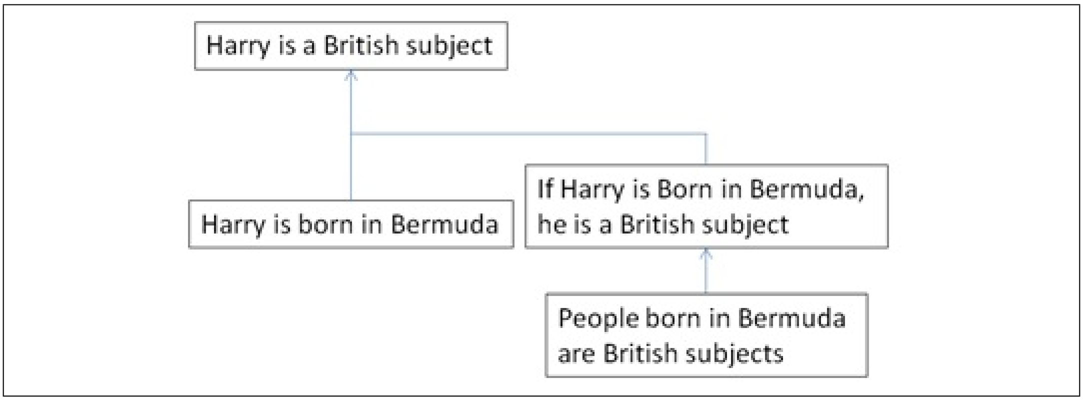

Figure 13: Using a conditional sentence

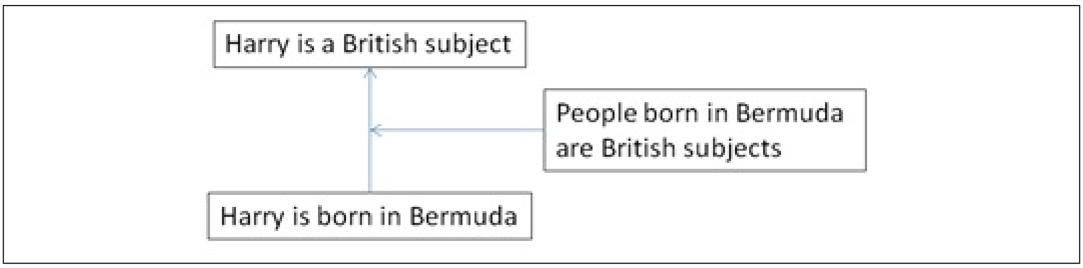

Figure 14: Nested arrows

premise. This conditional sentence can then be supported by further arguments, such as a warrant (as in Figure 13) or a backing. This approach is, for instance, proposed in the user-friendly Rationale[47] tool developed by van Gelder and his collaborators [van Gelder, 2007].

A third approach uses nested arrows. The arrows are treated as graphical expressions of the connection between the reason and conclusion, and can hence be argued about. In Figure 14, for instance, the warrant has been supplied as support for the connection between reason and conclusion. This approach has a straightforward generalisation when support and attack are combined (Section 4.2). The ArguMed tool developed by Verheij [2005b] uses this approach.

A variation of the nested arrows approach uses rule nodes (Figure 15), instead of nested arrows. The AVERs tool [van den Braak *et al.*, 2007] uses this approach.

[47] `http://rationale.austhink.com/`.

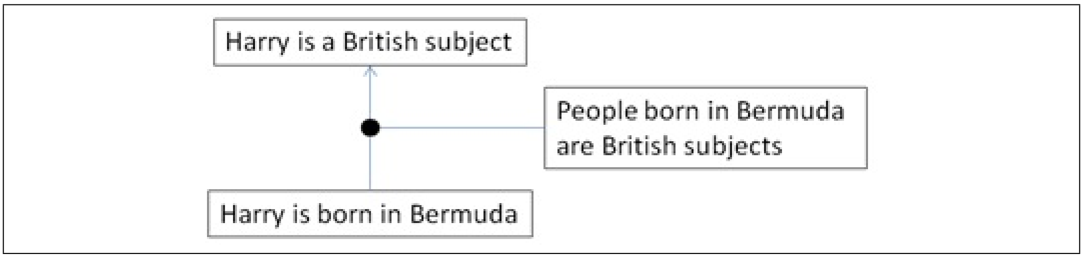

Figure 15: Rule nodes

Argument evaluation The third and final theme that we discuss in connection with the design of argumentation support software is argument evaluation. In argumentation software, different strategies for argument evaluation have been implemented. Some tools choose to leave argument evaluation as a task for the user of the system. For instance, in the Rationale system [van Gelder, 2007] a user can indicate which claims follow or do not follow given the reasons in the diagram. Specific graphical elements are used to show the user's evaluative actions.

In several other systems, some form of automatic evaluation has been implemented. Automatic evaluation algorithms can be logical, or numeric.

Logical evaluation algorithms in argumentation support tools have been grounded in versions of argumentation semantics (see Section 4.1). For instance, ArguMed [Verheij, 2005b] computes a version of stable semantics. Consider, for instance, Pollock's example of an undercutting defeater about red lights (see Section 3.4). ArguMed's evaluation algorithm behaves as expected: when the reason that the object looks red is assumed, the conclusion that the object is red will be justified, but that will no longer be the case when the defeater is added that the object is illuminated by a red light. A typical property of logical evaluation algorithms is reinstatement: when a defeating attacker of an initial argument is successfully attacked, the initial argument will no longer count as defeated and therefore be reinstated.

Numeric evaluation algorithms have been based on the numeric weights of the reasons supporting and attacking conclusions. A weight-based numeric evaluation algorithm has, for instance, been implemented in the Hermes system [Karacapilidis & Papadias, 2001]. In Hermes, positions can be assigned a numeric score by adding the weights of active pro-positions and subtracting the weights of active con-positions. A proof standard can be used to determine an activation label of a position. In the proof standard called Preponderance of evidence, for instance, a position is active when the active pro-positions outweigh the active con-positions.

A numeric evaluation algorithm of a different kind has been implemented in the so-called 'Convince me' system [Schank, 1995]. It uses ECHO, which is a connectionist version of Thagard's [1992] theory of explanatory coherence. In Convince me, statements are assigned numerical values by a step-wise constraint satisfaction algorithm. In the algorithm, incremental changes of the default weights of a statement are made by considering the excitatory and inhibitory links connected to a statement. When changes become too small to be taken into account (or computation is taking too long), the algorithm stops.

5 Specific kinds of argumentation in formal and computational perspective

In this section, we discuss specific kinds of argumentation using rules, cases, values and evidence. We end the section with applications and case studies.

5.1 Reasoning with rules

We already saw examples showing the close connections between argumentation research in artificial intelligence and legal applications. Since argumentation is an everyday task of professional lawyers this is not unexpected. An institutional reason however is that there exists an interdisciplinary research field, called artificial intelligence and law,[48] in which because of the nature of law the topic of argumentation has been given a great deal of attention. Early work in that field (e.g., [McCarty, 1977; Gardner, 1987]) already showed the intricacies and special characteristics of legal argumentation. Thorne McCarty [1977] attempted to formalize the detailed reasoning underlying a US Supreme Court case. Anne Gardner [1987] proposed a system aimed at what she called issue spotting. In a legal case, there is an issue when no rule applies or when conflicting rules apply and the conflict cannot be resolved. In this section, we pay special attention to the work inspired by developments in non-monotonic logic that has been carried out, mostly in the mid-1990s, regarding reasoning with (legal) rules.

Henry Prakken's [1997] book Logical tools for modelling legal argument provides an extensive and careful treatment of the contributions of techniques from non-monotonic logic to the formal modelling of legal reasoning.[49] The formal tools presented by Prakken have gradually evolved into the ASPIC+ model already mentioned [Prakken, 2010]. Parts of the material were developed in close collaboration

[48]The primary journal of the field of AI & law is *Artifical Intelligence and Law*, with the biennial ICAIL and annual JURIX as the main conferences.

[49]The book is based on Prakken's [1993] doctoral dissertation.

with Sartor (e.g., [Prakken & Sartor, 1996; 1998]; see also the excellent resource [Sartor, 2005]).

The following example shows how Prakken models a case in contract law [1997, p. 171]. The example concerns the defeasible rule that contracts only bind the contracting parties (d_1), and a defeasible, possibly contravening, rule specifically for contracts that concern the lease of a house, saying that such contracts also bind future owners of the house (d_2). Another exception is added by a defeasible rule saying that, even in the case of a house lease, when a tenant agrees to make such a stipulation only the contracting parties are bound (d_3). The factual statements f_1 and f_2 say respectively (1) that a house lease is a special kind of contract and (2) that binding only the contracting parties and binding also future owners of a house do not go together.

d_1 : x is a contract $\Rightarrow$ x only binds its parties
d_2 : x is a lease of house y $\Rightarrow$ x binds all owners of y
d_3 : x is a lease of house $y \wedge$ tenant has agreed in x that x only binds its parties $\Rightarrow$ x only binds its parties
f_1 : $\forall x \forall y$ (x is a lease of a house $y \rightarrow x$ is a contract)[50]
f_2 : $\forall x \forall y \neg$ (x only binds its parties $\wedge x$ binds all owners of y)

When there is a contract about the lease of a house, there is an apparent conflict, since both d_1 and d_2 seem to apply. In the system, the application of d_2 blocks the application of d_2, using a mechanism of specificity defeat (see Section 4.2). In a case where also the condition of d_3 is fulfilled, namely when the tenant has agreed that the lease contract only binds the contracting parties, the application of rule d_3 blocks the application of rule d_2, which in that case does no longer block the application of d_1.

Prakken uses elements from classical logic (for instance, classical connectives and quantifiers) and non-monotonic logic (defeasible rules and their names), and shows how they can be used to model rules with exceptions, as they occur prominently in the law. He treats, for instance, the handling of explicit exceptions, preferring the most specific argument, reasoning with inconsistent information, and reasoning about priority relations.

In the same period, Hage developed Reason-based logic ([Hage, 1997]; see also [Hage, 2005)].[51] Hage presents Reason-based logic as an extension of first-order predicate logic in which reasons play a central role. Reasons are the result of the

[50] '$\forall x$...' stands for 'for every entity x it holds that ...'. Similarly, for '$\forall y$...'

[51] Reason-based logic exists in a series of versions, some introduced in collaboration with Verheij (e.g., [Verheij, 1996a]).

application of rules.[52] Treating them as individuals allows the expression of properties of rules. Whether a rule applies depends on the rule's conditions being satisfied, but also on possible other reasons for or against applying the rule. Consider, for instance, the rule that thieves are punishable:

punishable: thief(x) $\Rightarrow$ punishable(x)

Here 'punishable' before the colon is the rule's name. When John is a thief (expressed as thief(john)), the rule's applicability can follow:

Applicable(thief(john) $\Rightarrow$ punishable(john))

This gives a reason that the rule ought to be applied. If there are no reasons against the rule's application, this leads to the obligation to apply the rule. From this it will follow that John is punishable.

A characteristic aspect of Reason-based logic is that it models the weighing of reasons. In this system, there is no numerical mechanism for weighing; rather it can be explicitly represented that certain reasons for a conclusion outweigh the reasons against the conclusion. When there is no weighing information the conflict remains unresolved and no conclusion follows.

Like Prakken, Hage uses elements from classical logic and non-monotonic logic. In his theory, because of the emphasis on philosophical and legal considerations, the flavour of Reason-based logic is less that of pure logic, but comes closer to representing the ways of reasoning in the domain of law. Where Prakken's book remains closer to the field of AI, Hage's book reads more like a theoretical essay in philosophy or law.

Reason-based logic has been applied, for instance, to a well-known distinction made by the legal theorist Dworkin [1978]: whereas legal rules seem to lead directly to their conclusion when they are applied, legal principles are not as direct, and merely give rise to a reason for their conclusion. Only a subsequent weighing of possibly competing reasons leads to a conclusion. Different models of the distinction between rules and principles in Reason-based logic have been proposed. Hage [1997] follows Dworkin and makes a strict formal distinction, whereas Verheij *et al.* [1998] show how the distinction can be softened by presenting a model in which rules and principles are the extremes of a spectrum.

Loui and Norman [1995] have argued that there is a calculus associated with what they call the compression of rationales, i.e., the combination and adaptation

[52]We shall simplify Hage's formalism a bit by omitting the explicit distinction between rules and principles.

of the rules underlying arguments which are akin to Toulmin's warrants. They give the following example of a compression of rules (rationales). When there is a rule 'vehicles used for private transportation are not allowed in the park' and also a rule 'vehicles are normally for private transportation,' then a two-step argument based on these two rules can be shortened when the so-called compression rationale 'no vehicles in the park,' based on these two rules, is used.

5.2 Case-based reasoning

Reasoning with rules (Section 5.1) is often contrasted with case-based reasoning. Whereas the former is about following rules that describe existing conditional patterns, the latter is about finding relevantly similar examples that, by analogy, can suggest possible conclusions in new situations. In the domain of law, rule-based reasoning is associated with the application of legal statutes, and case-based reasoning with the following of precedents. The contrast can be appreciated by looking at the following two examples.

> *Art. 300 of the Dutch Criminal Code*
> 1. Inflicting bodily harm is punishable with up to two years of imprisonment or a fine of the fourth category.
> 2. When the fact causes grievous bodily harm, the accused is punished with up to four years of imprisonment or a fine of the fourth category.
> 3. [...]

> *Dutch Supreme Court July 9, 2002, NJ 2002, 499*
> Theft requires the taking away of a good. Can one steal an already stolen car? The Supreme Court's answer is: yes.

The first example is an excerpt from a statutory article expressing a material rule of Dutch criminal law, stating the kinds of punishment associated with inflicting bodily harm. The levels of punishment depend on specific conditions, with more severe bodily harm being punishable with longer imprisonment. The second example is a (very) brief summary of a Supreme Court decision. In this case, an already stolen car was stolen from the thief. One of the statutory requirements of the crime theft is that a good is taken away, and here the car was already taken away from the original owner of the car. The new legal question was addressed whether stealing from the original thief can count as theft from the car's owner. In other words, can an already stolen car still be taken away from the original owner? Here the Supreme Court decided that stealing a stolen car can count as theft since the original ownership is the deciding criterion; it does not matter whether a good is actually in the control of the owner at the time of theft. When used as a precedent, this Supreme Court decision has the effect that similar cases are decided alike.

In case-based reasoning, the stare decisis doctrine is leading: when deciding a new case one should not depart from an earlier, relevantly similar decision, but decide analogously. In the field of AI and law, Kevin Ashley's HYPO system [1990] counts as a milestone in the study of case-based reasoning.[53] In HYPO, cases are treated as sets of factors, where factors are generalised facts pleading for or against a case. Consider the following example about an employee who has been dismissed by his employer, and aims to void (i.e., cancel) the dismissal.[54]

Issue:
    Can a dismissal be voided?

Precedent case:
+ The employee's behaviour was always good.
- There was a serious act of violence.
Outcome:
+ (voided)

Current case:
+ The employee's behaviour was always good.
- There was a serious act of violence.
+ The working atmosphere was not affected.
Outcome:
?

There is a precedent case with one factor pleading for voidance (the good behaviour), and one pleading against voidance (the violence). In this precedent case, it was decided that voidance was in place. In the current case, the same factors apply, but there is also one additional factor pleading for voidance, namely that the working atmosphere was not affected. One could say that the decision taken in the precedent case is even more strongly supported in the current case. As a result, in HYPO and similar systems the suggested conclusion is that also in the current case voidance of the dismissal would be called for.

[53]See also Rissland and Ashley [1987], Ashley [1989], and Rissland and Ashley [2002].
[54]The example is inspired by the case material used by Roth [2003].

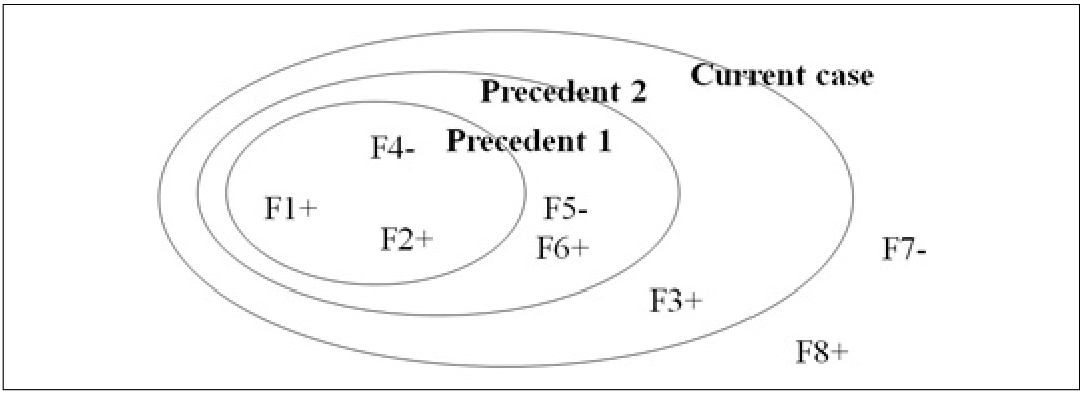

Figure 16: Factors in two precedent cases and the current case

The example in Figure 16 shows that factors can be handled formally without knowing what they are about. There is a first precedent with pro-factors F1 and F2 and a con-factor F4. The second precedent has as additional factors a con-factor F5 and a pro-factor F6. The current case has all these factors and one more pro-factor F3. The domain also contains con-factor F7 and pro-factor F8 which do not apply to these cases.

Assume now that the first precedent was decided negatively, and the second positively. The second precedent is more on point, in the sense that it shares more factors with the current case than the first precedent. Since the current case even has an additional pro-factor, it is suggested that the current case should be decided positively, in analogy with precedent 2. Precedents do not always determine the outcome of the current case. For instance, if the second precedent had been decided negatively, there would be no suggested outcome for the current case, since pro-factor F3 may be or may not be strong enough to turn the case.

Another formal example is shown in Figure 17. When both precedents have been decided positively, the suggested outcome for the current case is also positive. Precedent 1 can be followed because its support for a positive decision is weaker than that of the current case: the precedent has an additional con-factor, and the current case an additional pro-factor. Precedent 2 cannot be followed since F8 may be or may not be a stronger pro-factor than F3.

HYPO's aim is to form arguments about the current case, without determining a decision. This is made explicit in its model of 3–ply arguments. In HYPO's 3–ply model, the first argument move ('ply'), by the Proponent, is the citing of a precedent case in analogy with the current case. The analogy is based on the shared factors. The second argument move, by the Opponent, responds to the analogy, for instance,

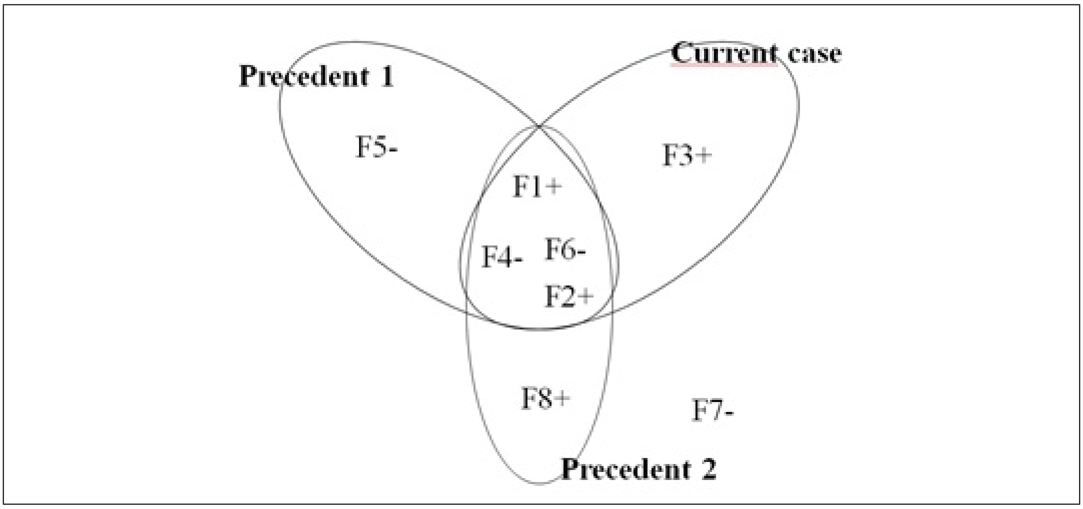

Figure 17: A different constellation of precedents

by distinguishing between the cited precedent case and the current case, pointing out differences in relevant factors, or by citing counterexamples. The third argument move, again by the Proponent, responds to the counterexamples, for instance, by making further distinctions.

HYPO's factors not only have a side (pro or con) associated with them, but can also come with a dimension pertaining in some way to the strength of the factor. This allows the citation of cases that share a certain factor, but have this factor with a different strength. For instance, by the use of dimensions, the good behaviour of the employee (of the first informal example) can come in gradations, say from good, via very good to excellent.

Vincent Aleven extended the HYPO model by the use of a factor hierarchy that allowed modelling of factors with hierarchical dependencies [Aleven, 1997; Aleven & Ashley, 1997a; 1997b]. For instance, the factor that one has a family to maintain is a special case of the factor that one has a substantial interest in keeping one's job. Inspired by Verheij's DefLog model [2003a], which allowed for reasoning about support and attack (Section 4.2), Roth [2003] developed case-based reasoning based on what he referred to as an entangled factor hierarchy, in order to expand the possible argumentative moves (Figure 18). For instance, the relevance of the factor that one has a family to maintain is strengthened by one's having children that go to university and weakened by one's having a wife with a good income. A factor hierarchy allows new kinds of argument moves by making it possible to downplay or emphasize a distinction. For instance, the factor of having a family to maintain

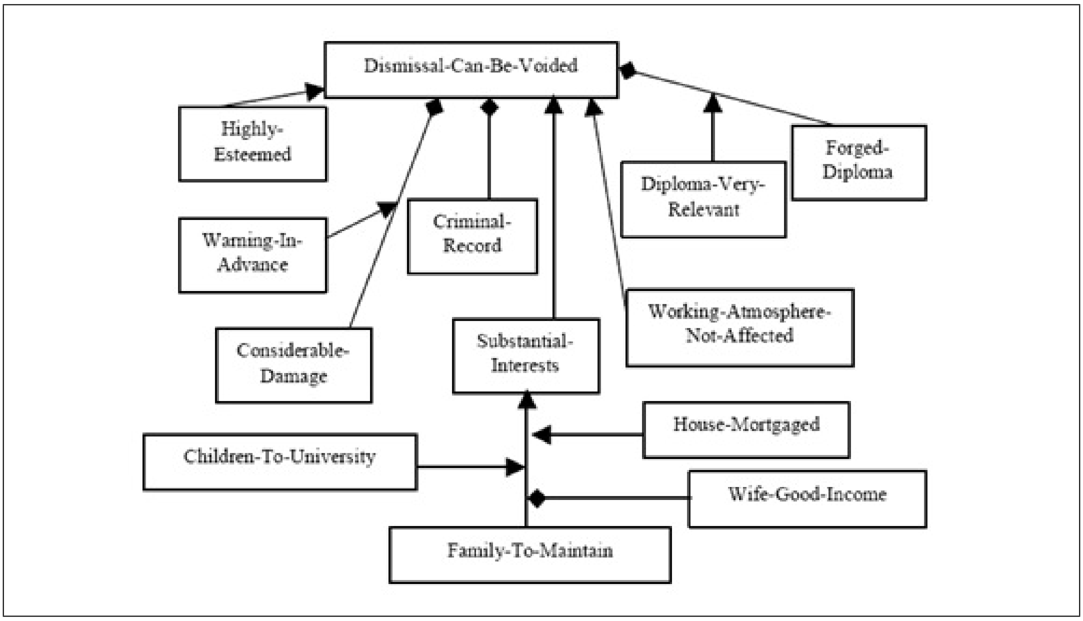

Figure 18: An entangled factor hierarchy [Roth, 2003]

can be downplayed by pointing out that one has a partner with a good income, or emphasized by mentioning that one has children going to university.

Proposals have been made to combine case-based and rule-based reasoning. For instance, Branting's GREBE model [1991; 2000] aims to generate explanations of decisions in terms of rules and cases. Both rules and cases can serve as warrants for a decision. Branting extends Toulmin's approach to warrants by using a so-called warrant reduction graph, in which warrants can be special cases of other warrants. Prakken and Sartor [1998] have applied their model of rule-based reasoning ([Prakken & Sartor, 1996]; see also Section 5.1) to the setting of case-based reasoning. Analogizing and distinguishing are connected to the deletion and addition of rule conditions that describe past decisions.

5.3 Values and audiences

Trevor Bench-Capon [2003] has developed a model of the values underlying arguments.[55] In this endeavour he refers to Perelman and Olbrechts-Tyteca's new rhetoric:

[55]In AI and law, the importance of the modelling of the values and goals underlying legal decisions was already acknowledged by Berman and Hafner [1993].

> If men oppose each other concerning a decision to be taken, it is not because they commit some error of logic or calculation. They discuss apropos the applicable rule, the ends to be considered, the meaning to be given to values, the interpretation and characterisation of facts [Perelman & Olbrechts-Tyteca, 1969, p. 150].

Because of the character of real-life argumentation, it is not to be expected that cases will be conclusively decided. Bench-Capon therefore aims to extend formal argumentation models by the inclusion of the values of the audiences addressed. This allows him to model the persuasion of an audience by means of argument.

Bench-Capon [2003] uses Dung's [1995] abstract argumentation frameworks as a starting point. He defines a value-based argumentation framework as a framework in which each argument has an associated (abstract) value. The idea is that values associated with an argument are promoted by accepting the argument. For instance, in a parliamentary debate about a tax raise it can be argued that accepting the raise will promote the value of social equality, while the value of enterprise is demoted. In an audience-specific argumentation framework, the preference ordering of the values can depend on an audience. For instance, the Labour Party may prefer the value of social equality, and the Conservative Party that of enterprise.

Bench-Capon continues to model defeat for an audience: an argument A defeats an argument B for audience *a* if A attacks B and the value associated with B is not preferred to the value associated with A for audience *a*. In his model, an attack succeeds, for instance, when the arguments promote the same value, or when there is no preference between the values. Dung's notions of argument acceptability, admissibility and preferred extension are then redefined relative to audience attack.

Bench-Capon uses a value-based argumentation framework with two values 'red' and 'blue' as an example (Figure 19). The underlying abstract argumentation framework is the same as that in Figure 6. In its unique preferred extension (which is also grounded and stable), A and C are accepted and B is rejected. For an audience preferring 'red,' defeat for the audience coincides with the underlying attack relation. In the preferred extension for an audience preferring 'red,' therefore, A and C are accepted and B is rejected. However, for an audience preferring 'blue,' A does not defeat B. But for such an audience B still defeats C. For a 'blue'-preferring audience, A and B are accepted and C is not.

Bench-Capon illustrates value-based argumentation by considering the case of a diabetic who almost collapses into a coma by lack of insulin, and therefore takes another diabetic's insulin after entering her house. He analyses the case by discussing the roles of the value of property right infringement as opposed to that of saving one's life.

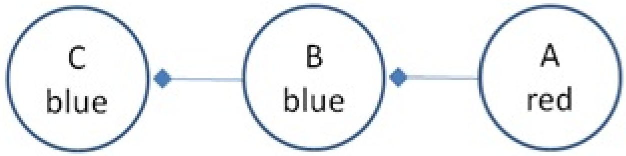

Figure 19: A value-based argumentation framework with two values (adapted from [Bench-Capon, 2003])

Bench-Capon and Sartor [2003] have used the value-based perspective in a treatment of legal reasoning that combines rule-based and case-based reasoning (see Sections 5.1 and 5.2). Legal reasoning takes the form of constructing and using a theory that explains a decision in terms of the values promoted and demoted by the decision. Precedent decisions have the role of revealing preferences holding between factors. This is similar to the role of precedents in HYPO that reveal how the factors in a precedent case are weighed. In Bench-Capon and Sartor's approach, the factor preferences in turn reveal preferences between values. The resulting preferences can then be used to decide new cases.

5.4 Burden of proof, evidence, and argument strength

Some arguments are more successful than others. An argument can meet or not meet the burden of proof fitting the circumstances of the debate. An argument can be founded on better evidence than another. An argument can also be stronger than another. In this section, we address the topics of burden of proof, evidence and argument strength.

Burden of proof and evidence The topic of burden of proof is strongly connected to the dialogical setting of argumentation. A burden of proof is assigned to a party in an argumentative dialogue when the quality of the arguments produced in the dialogue depends in part on whether the arguments produced by that party during the dialogue meet certain constraints. Such constraints can be procedural, for instance, requiring that a counterargument is met by a counterattack, or material, for instance, requiring that an argument is sufficiently strong in the light of the other arguments. Constraints of the latter, material, non-procedural type are also referred to as proof standards.

The topic of burden of proof is especially relevant in the law, as argumentation in court is often constrained by burden of proof constraints. As a result, in legal theory the topic has been studied extensively. The topic has also been addressed in

AI approaches to argumentation, in particular by researchers connected to the field of AI and law (see also Section 4.4). In the Carneades argumentation model [Gordon *et al.*, 2007], for instance, statements are categorized using three proof standards:

> *SE (Scintilla of Evidence).* A statement meets this standard if and only if it is supported by at least one defensible pro argument.
> *BA (Best Argument).* A statement meets this standard if and only if it is supported by some defensible pro argument with priority over all defensible con arguments.
> *DV (Dialectical Validity).* A statement meets this standard if and only if it is supported by at least one defensible pro argument and none of its con arguments are defensible.

A theme related to proof standards is argument accrual. What happens when there are several arguments for a conclusion? See Section 4.2, where research addressing the relation between argument defeat and accrual is discussed.

AI models of argumentation have been helpful in clarifying distinctions made in legal theory. Prakken and Sartor in particular have in a series of articles [Prakken & Sartor, 2007; 2009] contributed to the explication of different forms of burden of proof. They distinguish a burden of persuasion, a burden of production, and a tactical burden. A burden of persuasion requires that a party proves a statement to a specified degree (the standard of proof) or runs the risk of losing on the issue at the end of the debate. A burden of production has been assigned to a party when the party is required by law to provide evidence for a certain claim. Burdens of persuasion and burdens of production are assigned by the applicable law. The tactical burden of proof depends on a party's own assessment of whether sufficient grounds have been adduced about a claim made by the party. Prakken and Sartor connect these different notions to a formal dialogue model of argumentation.

Probability and other quantitative approaches to argument strength Argument strength can be considered by using quantitative approaches. For instance, a conditional probability $p(H|E)$, expressing the probability of a hypothesis H given the evidence E, can be interpreted as a measure of the strength of the argument for the hypothesis based on the evidence. The idea is that higher values of $p(H|E)$ make H more strongly supported when given E. This interpretation of argument strength is associated with what is called Bayesian epistemology [Talbott, 2011]. Bayesian epistemology provides in the following way an interpretation of the relevance of additional evidence, say E': additional evidence E' strengthens the argument E for H when $p(H|E \wedge E') > p(H|E)$. In this interpretation, Bayes' theorem:

$$p(H|E) = p(E|H) \times p(H)/p(E)$$

connects the strength of the argument from E to H and that of the argument from H to E, thereby reversing the direction of the arrow. This relation is helpful, when the values of $p(E|H), p(H)$ and $p(E)$ are available, or when they are more easily established than $p(H|E)$ itself. Bayesian epistemology also provides a perspective on the comparison of hypotheses given additional evidence. When there are two hypotheses H and H', the odds form of Bayes' theorem can be used to update the odds of the hypotheses in light of new evidence E. The following relation shows how the prior odds $p(H)/p(H')$ is connected to the posterior odds $p(H|E)/p(H'|E)$:

$$p(H|E)/p(H'|E) = (p(H)/p(H')) \times (p(E|H)/p(E|H'))$$

This formal relation is helpful when the prior odds $p(H)/p(H')$, and the values of $p(E|H)$ and $p(E|H')$ are available.

Pollock has argued against a probabilistic account of argument strength (e.g., [Pollock, 1995; 2006; 2010]), referring to this position as 'generic Bayesianism' or 'probabilism.' Pollock argues that in a probabilistic account we would be justified in believing a mathematical theorem even before it is proven. This is especially absurd in cases such as Fermat's last theorem, which remained a conjecture for centuries before Wiles finally could complete a proof in the 1990s. Fitelson [2010] defends a probabilistic account against this and other criticisms advanced by Pollock.

Zukerman, McConachy and Korb [1998] have discussed the possibility of generating arguments from Bayesian networks, which are a widely studied tool for the representation of probabilistic information. Riveret *et al.* [2007] consider success in argument games in connection with probability. Dung and Thang [2010] have presented an approach to probabilistic argumentation in the setting of dispute resolution. Verheij [2012; 2017] has proposed a formal theory of defeasible argumentation in which logical and probabilistic properties are connected. Hunter [2013] discusses a model of deductive argumentation with uncertain premises. Verheij *et al.* [2016] discuss connections between arguments, scenarios and probabilities as normative tools in forensic reasoning with evidence.

Evidence and inference to the best explanation When an argument is aimed at establishing the truth, empirical evidence can be used to support alleged facts. For instance, a witness's testimony can provide evidence for the claim that the suspect was at the scene of a crime, a clinical test can provide evidence against a medical diagnosis, and the outcome of a laboratory experiment can be evidence confirming (or falsifying) a psychological phenomenon. The conclusions based on the available evidence can be regarded as hypothetical explanations for the occurrence of the evidence. As a result, reasoning on the basis of evidence is a specimen of what Peirce referred to as abductive reasoning, or inference to the best explanation: reasoning that goes from data describing something to a hypothesis that best explains or accounts for the data [Josephson & Josephson, 1996, p. 5]. Josephson and Josephson conceive of inference to the best explanation as a kind of argument scheme (see Section 4.3):

> D is a collection of data (facts, observations, givens).
> H explains D (would, if true, explain D).
> No other hypothesis can explain D as well as H does.
> Therefore, H is probably true.
> [Josephson & Josephson, 1996, p. 5]

The explanatory connection between D and H is often regarded as going against the causal direction. For instance, a causal, expectation-evoking rule 'If there is a fire, then there is smoke' can be used to infer, or argue for, the effect 'there is smoke' after observing the cause 'there is fire.' The causal rule has an evidential, explanation-evoking counterpart, 'If there is smoke, then there is a fire,' that can be used to infer (argue for) the explanation 'there is a fire' after observing 'there is smoke.' Arguments based on causal or evidential rules are typically defeasible: not all fires generate smoke, and not all smoke stems from a fire.

In artificial intelligence, the distinction between causal and evidential rules has been emphasized by Pearl [1988, p. 499f.]. He argues that special care is needed when mixing causal and evidential reasoning. To make his point, Pearl uses the following examples:

> Bill showed slight difficulties standing up, so I believed he was injured.
> Harry seemed injured, so I believed he would be unable to stand up.

The former uses the evidential pathway from the observation of Bill's difficulties in standing up to the explanation that he is injured, and the latter the reverse causal pathway from the observation of Harry's injuries to the effect that he is unable to stand up. The question is then addressed whether it is likely that Bill or Harry are likely to be drunk, drunkenness being a second cause for difficulties in standing up, independent from injury. Both Bill's and Harry's intoxicated state could be argued for using the evidential rule 'If someone has difficulties standing up, then he may be drunk.' However, for Bill the conclusion that he may be drunk seems more likely than for Harry, since for Bill both explanations for his difficulties in standing up, namely injury or being drunk, seem to be reasonable, whereas for Harry drunkenness is a less likely hypothesis now that an injury has been observed. The distinction between causal and evidential rules has played a central role in Pearl's thinking about causality [Pearl, 2000/2009], which relates to the probabilistic modelling tool of Bayesian Networks (see [Jensen & Nielsen, 2007; Kjaerulff & Madsen, 2008]). Bayesian Networks have been connected to the modelling of argumentation with legal evidence by Hepler *et al.* [2007] and by Fenton *et al.* [2012] (see also [Taroni *et al.*, 2006]). Vlek *et al.* [2014; 2016] discuss the design and understanding of Bayesian Networks for evidential reasoning using scenarios. Timmer *et al.* [2017] discuss an algorithm to extract argumentative information from a Bayesian Network modeling hypotheses and evidence. Verheij [2017] investigates connections between arguments, scenarios and probabilities in one formal model.

The distinction between causal and evidential rules has also been used in the formalized hybrid argumentative-narrative model of reasoning with evidence developed by Bex and his colleagues [Bex *et al.*, 2010; Bex, 2011]. In this model, the elements of a scenario, or narrative, describing how a crime may have been committed, can be supported by arguments grounded in the available evidence. Causal connections between the elements of a scenario contribute to its coherence. It is possible that more than one scenario is available, each scenario with different evidential support and a different kind of coherence. Bex and Verheij [2012] have developed the argumentative-narrative model in terms of argument schemes and their associated critical questions (see Section 4.3).

5.5 Applications and case studies

A first reason for the popularity of argumentation research in the field of artificial intelligence is that it has led to theoretical advances. A second reason is that the theoretical advances have been corroborated by a variety of interesting applications and case studies, including advances in natural language processing. We give some examples.

Fox and Das [2000] provided a book-length study of AI technology in medical diagnosis and decision making, with much emphasis on the argumentative aspects (see also Fox and Modgil, 2006, where argumentation-based decision making is used to extend the Toulmin model). Aleven and Ashley [1997a; 1997b] developed a case-based argumentation tool that was empirically tested for its effects on learning. Buckingham Shum and Hammond [1994] approached the design of artefacts such as software as an argumentation problem. Grasso *et al.* [2000] worked on argumentative conflict resolution in the context of health promotion. Teufel [1999] has worked on the problem of automatically estimating a sentence's role in argumentation, using a model of seven text categories called argumentative zones. Mochales Palau and

Moens [2009] developed software for the mining of argumentative elements in legal texts. Hunter and Williams [2010] investigated the aggregation of evidence in a healthcare setting. Grasso [2002] and Crosswhite *et al.* [2004] have worked on the computational modelling of rhetorical aspects of argument. Reed and Grasso [2007] have collected argumentation-oriented research using natural language techniques. They discuss, for instance, the generation of argumentative texts as studied by Elhadad [1995], Reed [1999], Zukerman *et al.* [1998], and Green [2007].

Rahwan and McBurney [2007] edited a special issue on argumentation technology of the journal *IEEE Intelligent Systems.* Application areas addressed in the issue are medical decision-making, emotional strategies to persuade people to follow a healthy diet, ontology engineering, discussion mediation, and web services. In the 2012 edition of the COMMA conference proceedings series on the computational modelling of argument, a separate section was devoted to innovative applications. The topics included: automatic mining of arguments in opinions, a learning environment for scientific argumentation, semi-automatic analysis of online product reviews, argumentation with preferences in the setting of eco-efficient biodegradable packaging, hypothesis generation from cancer databases, sense making in policy deliberation, music recommendation, and argumentation about firewall policy. For applications focusing on argumentation support and facilitation, the reader is referred to Section 4.5.

In the domain of AI and law theories and systems were developed and tested by the use of case studies. For instance, McCarty [1977; 1995] analysed a seminal case in US tax law (Eisner v. Macomber, 252 U.S. 189 [1920]). In that case, the US Supreme Court decided that a federal rule of tax law was invalid. McCarty's aims were set high, namely to build a software implementation that could handle a number of elusive, argumentative aspects of legal reasoning, illustrated in the majority opinion and dissenting opinions concerning the issues in this case. Quoting McCarty [1995]:

1. Legal concepts cannot be adequately represented by definitions that state necessary and sufficient conditions. Instead, legal concepts are incurably 'open-textured'.

2. Legal rules are not static, but dynamic. As they are applied to new situations, they are constantly modified to 'fit' the new 'facts'. Thus the important process in legal reasoning is not theory application, but theory construction.

3. In this process of theory construction, there is no single 'right answer'. However, there are plausible arguments, of varying degrees of persuasiveness, for each alternative version of the rule in each new factual situation.

Berman and Hafner [1993] studied the 1805 Pierson v. Post case concerning the ownership of a dead fox chased by Post, but killed and taken by Pierson. They emphasize the teleological aspects of legal argumentation, in which the goals of legal rules and decisions are taken into account. Bex [2011] used the Anjum case, a Dutch high media profile murder case, to test his proposal for a hybrid argumentative-narrative model of reasoning with evidence. Atkinson [2012] edited an issue of the journal *Artificial Intelligence and Law* on the modelling of a 2002 case about the ownership of a baseball, representing possibly value in the order of a million dollars, being the one that Barry Bonds hit when he broke the record of home-runs in one season (Popov v. Hayashi).

6 Conclusion

In the previous sections, we have introduced argumentation and argumentation theory as a field of study that goes back to classical times, passing through a neo-classical and anti-formal period in the second half of the 20$^{\text{nd}}$ century, and since the final decade of the 2$^{\text{nd}}$ millenium going through a formal and computational turn.

In Section 2, we discussed crucial concepts that have been indispensable in the study of argumentation before the recent formal and computational turn: standpoints, unexpressed premises, argument schemes, argumentation structures, and fallacies. All of these also played—and still play—a significant role in current formal and computational approaches to argumentation.

Standpoints occur in formal and computational work as the conclusions of arguments—possibly intermediate—and as the commitments of the players in a computational dialogue game. Recently we see a move towards standpoints with a complex structure, in work that allows a complex hypothesis (such as a plan or a scenario) as the conclusion of an argument.

Unexpressed premises have been studied in the context of manually analyzing argumentative texts in software tools. In today's research on argument mining, attempts are made to automatically understand argumentative texts, and we see that the ubiquity of unexpressed elements in argumentative discourse provides a significant hurdle.

Argument schemes have been the source of much interaction between the non-formal and formal/computational research communities. This is not a coincidence as argument schemes can be regarded as being intermediate between non-formal and the formal: argument schemes are formal in the sense that they have a well-organized structure, including elements such as premises, conclusions and critical questions; and argument schemes are non-formal in the sense that they handle just about every

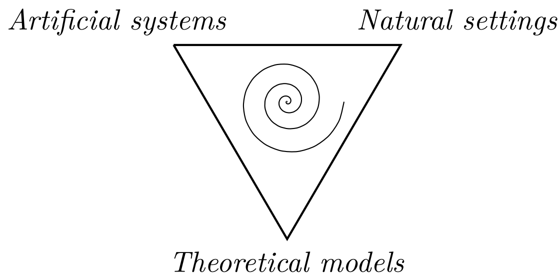

Figure 20: Perspectives on argumentation

area of human reasoning, whether legal, medical, or common-sense. Because of their intermediate position, argument schemes have been referred to as semi-formal.

Argumentation structures have been extensively studied both in non-formal and in formal research into argumentation theory. Today's argumentation logics and argumentation diagramming tools provide carefully designed structuring tools that fit the non-formal theory well, and that have been applied to argument analysis and design. In the study of argumentation structures, we see perhaps most convincingly that the anti-logical period in argumentation theory of the second half of the 20$^{\text{nd}}$ century is now superseded by a fruitful interaction between formal and non-formal methods.

Fallacies have received mostly indirect attention in the formal and computational study of argumentation, in particular because the mirror image of fallacies—correct argumentation—is and always has been in the center of formal attention. Much progress has been made in the characterization of typically argumentative versions of validity, initially distancing from classical formal theories, and nowadays gradually returning to an integration with classical logic and standard probability theory, this time while engaging with the needs of actual human argumentation as uncovered in argumentation theory.

We hope that it has become clear that there are a great many issues that can be fruitfully researched if argumentation and artificial intelligence scholars cooperate (cf. the research programme initiated by Reed & Norman [2004]). The distinction between non-formal and formal argumentation theory becomes ever more blurred, and argumentation theory is ever further turning into an interdisciplinary enterprise, integrating insights from different perspectives (see Figure 20).

In the *theoretical models* perspective, the focus is on theoretical (possibly non-formal) and formal models of argumentation, for instance, extending the long tradition of philosophical and formal logic. In the *artificial systems* perspective, the aim is to build computer programmes that model or support argumentative tasks, for instance, in online dialogue games or in knowledge-based systems (computer programmes that reproduce the reasoning of an expert, for instance, in the law or in medicine). The *natural settings* perspective helps to ground research by concentrating on argumentation in its natural form, for instance, in the human mind or in an actual debate. We are curious where the continuing synergy between these perspectives will bring our understanding of argumentation, this utterly human characteristic of civilized coexistence.

References

[Aleven, 1997] Aleven, V. (1997). *Teaching case-based reasoning through a model and examples.* Doctoral dissertation University of Pittsburgh.

[Aleven & Ashley, 1997a] Aleven, V., & Ashley, K. D. (1997a). Evaluating a learning environment for case-based argumentation skills. In *Proceedings of the sixth international conference on artificial intelligence and law* (pp. 170–179). New York: ACM Press.

[Aleven & Ashley, 1997b] Aleven, V., & Ashley, K. D. (1997b). Teaching case-based argumentation through a model and examples. Empirical evaluation of an intelligent learning environment. In B. du Boulay & R. Mizoguchi (Eds.), *Artificial intelligence in education. Proceedings of AI-ED 97 world conference* (pp. 87–94). Amsterdam: IOS Press.

[Alexy, 1978] Alexy, R. (1978). *Theorie der juristischen Argumentation. Die Theorie des rationale Diskurses as Theorie der juristischen Begründung* (A theory of legal argumentation. The theory of rational discourse as theory of juridical justification). Frankfurt am Main: Suhrkamp. (Spanish transl. by M. Atienza & I. Espejo as *Teoría de la argumentación jurídica.* Madrid: Centro de Estudios Constitucionales, 1989).

[Amgoud, 2009] Amgoud, L. (2009). Argumentation for decision making. In I. Rahwan & G. R. Simari (Eds.), *Argumentation in artificial intelligence* (pp. 301–320). Dordrecht: Springer.

[Amgoud *et al.*, 2008] Amgoud, L., Cayrol, C., Lagasquie-Schiex, M. C., & Livet, P. (2008). On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, 23(10), 1062–1093.

[Anscombre, 1994] Anscombre, J. C. (1994). La nature des topoï [The nature of the topoi]. In J. C. Anscombre (Ed.), *La théorie des topoï* [The theory of the topoi] (pp. 49–84). Paris: Kimé.

[Anscombre & Ducrot, 1983] Anscombre, J. C., & Ducrot, O. (1983). *L'argumentation dans la langue* [Argumentation in language]. Brussels: Pierre Mardaga.

[Antonelli, 2010] Antonelli, G. A. (2010). Non-monotonic logic. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Summer 2010 ed. `http://plato.stanford.edu/archives/sum2010/entries/logic-non-monotonic/`.

[Antoniou *et al.*, 2001] Antoniou, G., Billington, D., Governatori, G., & Maher, M. (2001). Representation results for defeasible logic. *ACM Transactions on Computational Logic*, 2(2), 255–287.

[Aristotle, 1984] Aristotle (1984). *The complete works of Aristotle. The revised Oxford translation.* 2 volumes. J. Barnes (Ed.). Transl. a.o. by W. A. Pickard-Cambridge (*Topics* and *Sophistical refutations*, 1928), J. L. Ackrill (*Categories* and *De interpretatione*, 1963), A. J. Jenkinson (*Prior analytics*), and W. Rhys Roberts (*Rhetoric*, 1924). Princeton, NJ: Princeton University Press.

[Ashley, 1989] Ashley, K. D. (1989). Toward a computational theory of arguing with precedents. Accommodating multiple interpretations of cases. In *Proceedings of the second international conference on artificial intelligence and law* (pp. 93–102). New York: ACM Press.

[Ashley, 1990] Ashley, K. D. (1990). *Modeling legal argument. Reasoning with cases and hypotheticals*. Cambridge, MA: The MIT Press.

[Atkinson, 2012] Atkinson, K. (2012). Introduction to special issue on modelling Popov v. Hayashi. *Artificial Intelligence and Law*, 20, 1–14.

[Atkinson & Bench-Capon, 2007] Atkinson, K., & Bench-Capon, T. J. M. (2007). Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, 171, 855–874.

[Atkinson *et al.*, 2005] Atkinson, K., Bench-Capon, T. J. M., & McBurney, P. (2005). A dialogue game protocol for multi-agent argument over proposals for action. *Autonomous Agents and Multi-Agent Systems*, 11, 153–171.

[Atkinson *et al.*, 2006] Atkinson, K., Bench-Capon, T. J. M., & McBurney, P. (2006). Computational representation of practical argument. *Synthese*, 152, 157–206.

[Baroni *et al.*, 2011] Baroni, P., Caminada, M., & Giacomin, M. (2011). An introduction to argumentation semantics. *Knowledge Engineering Review*, 26(4), 365–410.

[Barth, 1972] Barth, E. M. (1972). *Evaluaties. Rede uitgesproken bij de aanvaarding van het ambt van gewoon lector in de logica met inbegrip van haar geschiedenis en de wijsbegeerte van de logica in haar relatie tot de wijsbegeerte in het algemeen aan de Rijksuniversiteit te Utrecht op vrijdag 2 juni 1972* (Evaluations. Address given at the assumption of duties as profssor of logic including its history and philosophy of logic in relation to philosophy in general at the University of Utrecht on Friday, 2 June 1972). Assen: van Gorcum.

[Barth & Krabbe, 1982] Barth, E. M., & Krabbe, E. C. W. (1982). *From axiom to dialogue. A philosophical study of logics and argumentation.* Berlin/New York: Walter de Gruyter.

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

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

[Bench-Capon & Sartor, 2003] Bench-Capon, T. J. M., & Sartor, G. (2003). A model of legal reasoning with cases incorporating theories and values. *Artificial Intelligence*, 150, 97–143.

[Bench-Capon *et al.*, 2012] Bench-Capon, T. J. M., Araszkiewicz, M., Ashley, K., Atkinson, K., Bex, F., Borges, F., Bourcier, D., Bourgine, D., Conrad, J. G., Francesconi, E., Gordon, T. F., Governatori, G., Leidner, J. L., Lewis, D. D., Loui, R. P., McCarty, L. T., Prakken, H., Schilder, F., Schweighofer, E., Thompson, P., Tyrrell, A., Verheij, B., Walton, D. N., & Wyner, A. Z. (2012). A history of AI and Law in 50 papers. 25 years of the international conference on AI and Law. *Artificial Intelligence and Law*, 20(3), 215–319.

[Bench-Capon *et al.*, 2000] Bench-Capon, T. J. M., Geldard, T., & Leng, P. H. (2000). A method for the computational modelling of dialectical argument with dialogue games. *Artificial Intelligence and Law*, 8, 233–254.

[Bench-Capon *et al.*, 2004] Bench-Capon, T. J. M., Freeman, J. B., Hohmann, H., & Prakken, H. (2004). Computational models, argumentation theories and legal practice. In Chr. A. Reed & T. J. Norman (Eds.), *Argumentation machines. New frontiers in argument and computation* (pp. 85–120). Dordrecht: Kluwer.

[Bench-Capon *et al.*, 2009] Bench-Capon, T. J. M., Prakken, H., & Sartor, G. (2009). Argumentation in legal reasoning. In I. Rahwan & G. R. Simari (Eds.), *Argumentation in artificial intelligence* (pp. 363–382). Dordrecht: Springer.

[Berger, 1977] Berger, F. R. (1977). *Studying deductive logic.* London: Prentice-Hall.

[Berman & Hafner, 1993] Berman, D., & Hafner, C. (1993). Representing teleological structure in case-based legal reasoning. The missing link. In *Proceedings of the fourth international conference on artificial intelligence and law*, 50–59. New York: ACM Press.

[Besnard & Hunter, 2008] Besnard, P., & Hunter, A. (2008). *Elements of argumentation.* Cambridge, MA: The MIT Press.

[Bex, 2011] Bex, F. J. (2011). *Arguments, stories and criminal evidence. A formal hybrid theory.* Dordrecht: Springer.

[Bex & Verheij, 2012] Bex, F. J., & Verheij, B. (2012). Solving a murder case by asking critical questions. An approach to fact-finding in terms of argumentation and story schemes. *Argumentation*, 26(3), 325–353.

[Bex *et al.*, 2010] Bex, F. J., Koppen, P. van, Prakken, H., & Verheij, B. (2010). A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law*, 18(2), 123–152.

[Bex *et al.*, 2003] Bex, F. J., Prakken, H., Reed, Chr., & Walton, D. N. (2003). Towards a formal account of reasoning about evidence. Argumentation schemes and generalisations. *Artificial Intelligence and Law*, 11, 125–165.

[Bondarenko, *et al.*, 1997] Bondarenko, A., Dung, P. M., Kowalski, R. A., & Toni, F. (1997). An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93, pp. 63–101.

[Braak van den *et al.*, 2007] Braak, S. W. van den, Vreeswijk, G., & Prakken, H. (2007). AVERs. An argument visualization tool for representing stories about evidence. In *Proceedings of the 11th international conference on artificial intelligence and law* (pp.

11–15). New York: ACM Press.

[Branting, 1991] Branting, L. K. (1991). Building explanations from rules and structured cases. *International Journal of Man-Machine Studies*, 34, 797–837.

[Branting, 2000] Branting, L. K. (2000). *Reasoning with rules and precedents. A computational model of legal analysis.* Dordrecht: Kluwer Academic.

[Bratko, 2001] Bratko, I. (2001). *PROLOG. Programming for artificial intelligence.* 3rd ed. Harlow: Pearson. (1st ed. 1986).

[Brewka, 2001] Brewka, G. (2001). Dynamic argument systems. A formal model of argumentation processes based on situation calculus. *Journal of Logic and Computation*, 11, 257–282.

[Buckingham Shum & Hammond, 1994] Buckingham Shum, S., & Hammond, N. (1994). Argumentation-based design rationale. What use at what cost? *International Journal of Human-Computer Studies*, 40(4), 603–652.

[Caminada, 2006] Caminada, M. (2006). Semi-stable semantics. In Dunne, P. E., Bench-Capon, T. & J. M. (Eds.), *Computational models of argument. Proceedings of COMMA 2006, September 11–12, 2006, Liverpool, UK. Frontiers in artificial intelligence and applications*, 144 (pp. 121–130). Amsterdam: IOS Press.

[Cayrol & Lagasquie-Schiex, 2005] Cayrol, C., & Lagasquie-Schiex, M. C. (2005). On the acceptability of arguments in bipolar argumentation frameworks. In L. Godo (Ed.), *Symbolic and quantitative approaches to reasoning with uncertainty. 8th European conference, ECSQARU 2005* (pp. 378–389). Berlin: Springer.

[Chesñevar *et al.*, 2004] Chesñevar, C. I., Simari, G. R., Alsinet, T., & Godo, L. (2004). A logic programming framework for possibilistic argumentation with vague knowledge. In *Proceedings of the 20th conference on uncertainty in artificial intelligence* (pp. 76–84).

[Chesñevar *et al.*, 2006] Chesnevar, J. McGinnis, Modgil, S., Rahwan, I., Reed, Chr., Simari, G., South, M., Vreeswijk, G., & Willmott, S. (2006). Towards an argument interchange format. *The Knowledge Engineering Review*, 21(4), 293–316.

[Crawshay-Williams, 1957] Crawshay-Williams, R. (1957). *Methods and criteria of reasoning. An inquiry into the structure of controversy.* London: Routledge & Kegan Paul.

[Crosswhite *et al.*, 2004] Crosswhite, J., Fox, J., Reed, Chr. A., Scaltsas, T., & Stumpf, S. (2004). *Computational models of rhetorical argument.* In Chr. A. Reed & T. J. Norman (Eds.), *Argumentation machines. New frontiers in argument and computation* (pp.175–209). Dordrecht: Kluwer.

[d'Avila Garcez *et al.*, 2009] d'Avila Garcez, A. S., Lamb, L. C., & Gabbay, D. M. (2009). *Neural-symbolic cognitive reasoning.* Springer: Berlin.

[Dignum *et al.*, 2001] Dignum, F., Dunin-Kęplicz, B., & Verbrugge, R. (2001). Creating collective intention through dialogue. *Logic Journal of the IGPL* 9(2), 305–319.

[Doury, 1997] Doury, M. (1997). *Le débat immobile. L'Argumentation dans le débat médiatique surles parasciences.* [The immobile debate. Argumentation in the media debate on the parasciences]. Paris: Kimé.

[Dung, 1995] 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.

[Dung & Thang, 2010] Dung, P. M., & Thang, P. M. (2010). Towards (probabilistic) argumentation for jury-based dispute resolution. In P. Baroni, F. Cerutti, M. Giacomin & G. R. Simari (Eds.), *Computational models of argument – Proceedings of COMMA 2010* (pp. 171–182). Amsterdam: IOS Press.

[Dunne, 2007] Dunne, P. E. (2007). Computational properties of argument systems satisfying graph-theoretic constraints. *Artificial Intelligence*, 171(10), 701–729.

[Dunne & Bench-Capon, 2003] Dunne, P. E., & Bench-Capon, T. J. M. (2003). Two party immediate response disputes. Properties and efficiency. *Artificial Intelligence*, 149(2), 221–250.

[Dworkin, 1978] Dworkin, R. (1978). *Taking rights seriously. New impression with a reply to critics.* London: Duckworth.

[Eagly & Chaiken, 1993] Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt Brace Jovanovich.

[van Eemeren, 1987] Eemeren, F. H. van (1987). Argumentation studies' five estates. In J. W. Wenzel (Ed.), *Argument and critical practices. Proceedings of the fifth SCA/AFA conference on argumentation* (pp. 9–24). Annandale, Virginia: Speech Communication Association.

[van Eemeren, 1990] Eemeren, F. H. van (1990). The study of argumentation as normative pragmatics. *Text. An Interdisciplinary Journal for the Study of Discourse*, 10(1/2), 37–44.

[van Eemeren, 2001] Eemeren, F. H. van (2001). Fallacies. In F. H. van Eemeren (Ed.), *Crucial concepts in argumentation theory* (pp. 135–164). Amsterdam: Amsterdam University Press.

[van Eemeren, 2010] Eemeren, F. H. van (2010). *Strategic maneuvering in argumentative discourse. Extending the pragma-dialectical theory of argumentation.* Amsterdam/Philadelphia: John Benjamins.

[van Eemeren, 2017] Eemeren, F. H. van (Ed.) (2017). *Prototypical argumentative patterns. Exploring the relationship between argumentative discourse and institutional context.* Amsterdam/Philadelphia: John Benjamins.

[van Eemeren *et al.*, 2009] Eemeren, F. H. van, Garssen, B., & Meuffels, B. (2009). *Fallacies and judgments of reasonableness. Empirical research concerning the pragma-dialectical discussion rules.* Dordrecht etc.: Springer.

[van Eemeren & Grootendorst, 1984] Eemeren, F. H. van, & Grootendorst, R. (1984). *Speech acts in argumentative discussions. A theoretical model for the analysis of discussions directed towards solving conflicts of opinion.* Dordrecht/Cinnaminson: Foris & Berlin: de Gruyter. (transl. into Russian (1994c), Spanish (2013)).

[van Eemeren & Grootendorst, 1992] Eemeren, F. H. van, & Grootendorst, R. (1992). *Argumentation, communication, and fallacies. A pragma-dialectical perspective.* Hillsdale, NJ: Lawrence Erlbaum.

[van Eemeren & Grootendorst, 2004] Eemeren, F. H. van, & Grootendorst, R. (2004). *A systematic theory of argumentation. The pragma-dialectical approach.* Cambridge: Cambridge University Press.

[van Eemeren *et al.*, 1978] Eemeren, F. H. van, Grootendorst, R., & Kruiger, T. (1978). *Argumentatietheorie* (Argumentation theory). Utrecht: Het Spectrum. (2$^{\text{nd}}$ enlarged ed. 1981; 3rd ed. 1986).

[van Eemeren *et al.*, 1984] Eemeren, F. H. van, Grootendorst, R., & Kruiger, T. (1984). *The study of argumentation.* New York: Irvington. (Engl. transl. by H. Lake of F. H. van Eemeren, R. Grootendorst & T. Kruiger (1981). *Argumentatietheorie.* 2$^{\text{nd}}$ ed. Utrecht: Het Spectrum). (1 st ed. 1978).

[van Eemeren & Houtlosser, 2002] Eemeren, F. H. van, & Houtlosser, P. (2002a). Strategic maneuvering in argumentative discourse. Maintaining a delicate balance. In F. H. van Eemeren & P. Houtlosser (Eds.), *Dialectic and rhetoric. The warp and woof of argumentation analysis* (pp. 131–159). Dordrecht: Kluwer Academic.

[van Eemeren *et al.*, 2010] Eemeren, F. H. van, Grootendorst, R., Jackson, S., & Jacobs, S. (1993). *Reconstructing argumentative discourse.* Tuscaloosa, AL: University of Alabama Press.

[van Eemeren *et al.*, 2014] Eemeren, F.H. van, Garssen, B., Krabbe, E.C.W., Snoeck Henkemans, A.F., Verheij, B., & Wagemans, J.H.M. (2014). *Handbook of Argumentation Theory.* Dordrecht: Springer.

[Egly, Gaggl & Woltran, 2010] Egly, U., Gaggl, S. A., & Woltran, S. (2010). Answer-set programming encodings for argumentation frameworks. *Argument and Computation*, 1(2), 147–177.

[Elhadad, 1995] Elhadad, M. (1995). Using argumentation in text generation. *Journal of Pragmatics*, 24, 189–220.

[Fahnestock, 1999] Fahnestock, J. (1999). *Rhetorical figures in science.* New York, NY: Oxford University Press.

[Fahnestock, 2009] Fahnestock, J. (2009). Quid pro nobis. Rhetorical stylistics for argument analysis. In F. H. van Eemeren (Ed.), *Examining argumentation in context. Fifteen studies on strategic maneuvering* (pp. 131–152). Amsterdam: John Benjamins.

[Falappa *et al.*, 2002] Falappa, M. A., Kern-Isberner, G., & Simari, G. R. (2002). Explanations, belief revision and defeasible reasoning. *Artificial Intelligence*, 141(1–2), 1–28.

[Fenton *et al.*, 2012] Fenton, N. E., Neil, M., & Lagnado, D. A. (2012). A general structure for legal arguments using Bayesian networks. *Cognitive Science*, 37: 61–102.

[Finocchiaro, 1987] Finocchiaro, M. A. (1987). Six types of fallaciousness. Toward a realistic theory of logical criticism. *Argumentation*, 1, 263–282.

[Fitelson, 2010] Fitelson, B. (2010). Pollock on probability in epistemology. *Philosophical Studies*, 148, 455–465.

[Fox & Das, 2000] Fox, J., & Das, S. (2000). *Safe and sound. Artificial intelligence in hazardous applications.* Cambridge MA: The MIT Press.

[Fox & Modgil, 2006] Fox, J., & Modgil, S. (2006). From arguments to decisions. Extending the Toulmin view. In D. L. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin model. New essays in argument analysis and evaluation* (pp. 273–287). Dordrecht: Springer.

[Freeman, 2005] Freeman, J. B. (2005). *Acceptable premises. An epistemic approach to an informal logic problem.* Cambridge: Cambridge University Press.

[Gabbay, *et al.*, 1994] Gabbay, D. M., Hogger, C. J., & Robinson, J. A. (Eds., 1994). *Handbook of logic in artificial intelligence and logic programming, 3. Non-monotonic reasoning and uncertain reasoning.* Oxford: Clarendon Press.

[García, & Simari, 2004] García, A. J., & Simari, G. R. (2004). Defeasible logic programming. An argumentative approach. *Theory and practice of logic programming*, 4(2), 95–138.

[Gardner, 1987] Gardner, A. (1987). *An artificial intelligence approach to legal reasoning.* Cambridge.

[Garssen, 2001] Garssen, B. J. (2001). Argument schemes. In F. H. van Eemeren (Ed.), *Crucial concepts in argumentation theory* (pp. 81–99). Amsterdam: Amsterdam University Press.

[Gelder van, 2007] Gelder, T. van (2007). The rationale for Rationale. *Law, Probability and Risk*, 6, 23–42.

[Gelfond & Lifschitz, 1988] Gelfond, M., & Lifschitz, V. (1988). The stable model semantics for logic programming. In R. A. Kowalski & K. A. Bowen (Eds.), *Logic programming. Proceedings of the fifth international conference and symposium* (pp. 1070–1080). Cambridge MA: The MIT Press.

[Ginsberg, 1994] Ginsberg, M. L. (1994). AI and non-monotonic reasoning. In D. M. Gabbay, C. J. Hogger & J. A. Robinson (Eds.), *Handbook of logic in artificial intelligence and logic programming, 3. Non-monotonic reasoning and uncertain reasoning* (pp. 1–33). Oxford: Clarendon Press.

[Girle *et al.*, 2004] Girle, R., Hitchcock, D. L., McBurney, P., & Verheij, B. (2004). *Decision support for practical reasoning. A theoretical and computational perspective.* In Chr. A. Reed & T. J. Norman (Eds.), *Argumentation machines. New frontiers in argument and computation* (pp. 55–83). Dordrecht: Kluwer Academic.

[Gómez Lucero *et al.*, 2009] Gómez Lucero, M., Chesñevar, C., & Simari, G. (2009). Modelling argument accrual in possibilistic defeasible logic programming. In *ECSQARU Ś09 Proceedings of the 10th European conference on symbolic and quantitative approaches to reasoning with uncertainty* (pp. 131–143). Springer: Berlin.

[Goodnight, 1982] Goodnight, G. Th. (1982). The personal, technical, and public spheres of argument. A speculative inquiry into the art of public deliberation. *Journal of the American Forensic Association,* 18, 214–227.

[Goodnight, 2012] Goodnight, G. T. (2012). The personal, technical, and public spheres. A note on 21st century critical communication inquiry. *Argumentation and Advocacy,* 48(4), 258–267.

[Gordon, 1993] Gordon, T. F. (1993). The Pleadings Game. *Artificial Intelligence and Law,* 2(4), 239–292.

[Gordon, 1995] Gordon, T. F. (1995). *The Pleadings Game. An artificial intelligence model of procedural justice.* Dordrecht: Kluwer.

[Gordon & Karacapilidis, 1997] Gordon, T. F., & Karacapilidis, N. (1997). The Zeno argumentation framework. In *Proceedings of the ICAIL 1997 conference* (pp. 10–18). New York: ACM Press.

[Gordon *et al.*, 2007] Gordon, T. F., Prakken, H., & Walton, D. N. (2007). The Carneades model of argument and burden of proof. *Artificial Intelligence*, 171, 875–896.

[Govier, 1987] Govier, T. (1987). *Problems in argument analysis and evaluation.* Dordrecht/Providence, RI: Foris.

[Grasso, 2002] Grasso, F. (2002). Towards computational rhetoric. *Informal Logic*, 22, 195–229.

[Grasso, Cawsey & Jones, 2000] Grasso, F., Cawsey, A., & Jones, R. (2000). Dialectical argumentation to solve conflicts in advice giving. A case study in the promotion of healthy nutrition. *International Journal of Human-Computer Studies*, 53(6), 1077–1115.

[Greco Morasso, 2011] Greco Morasso, S. (2011). *Argumentation in dispute mediation. A reasonable way to handle conflict.* Amsterdam/Philadelphia: John Benjamins.

[Green, 2007] Green, N. (2007). A study of argumentation in a causal probabilistic humanistic domain. Genetic counseling. *International Journal of Intelligent Systems*, 22, 71–93.

[Habermas, 1973] Habermas, J. (1973). Wahrheitstheorien (Theories of truth). In H. Fahrenbach (Ed.), *Wirklichkeit und Reflexion. Festschrift für Walter Schulz zum 60. Geburtstag* (Reality and reflection. Festschrift for Walter Schulz in celebration of his 60th birthday) (pp. 211–265). Pfullingen: Günther Neske.

[Habermas, 1984] Habermas, J. (1984). *The theory of communicative action.* Vol. 1, *Reason and the rationalization of society.* Boston: Beacon. (English transl.; original work in German 1981).

[Hage, 1997] Hage, J. C. (1997). *Reasoning with rules. An essay on legal reasoning and its underlying logic.* Dordrecht: Kluwer Academic.

[Hage, 2000] Hage, J. C. (2000). Dialectical models in artificial intelligence and law. *Artificial Intelligence and Law*, 8, 137–172.

[Hage, 2005] Hage, J. C. (2005). *Studies in legal logic.* Berlin: Springer.

[Hage, Leenes & Lodder, 1993] Hage, J. C., Leenes, R., & Lodder, A. R. (1993). Hard cases. A procedural approach. *Artificial Intelligence and Law*, 2(2), 113–167.

[Hamblin, 1970] Hamblin, C. L. (1970). *Fallacies.* London: Methuen. Reprinted in 1986, with a preface by J. Plecnik & J. Hoaglund. Newport News, VA: Vale Press.

[Hart, 1951] Hart, H. L. A. (1951). The ascription of responsibility and rights. In A. Flew (Ed.), *Logic and language* (pp. 171–194). Oxford: Blackwell. (Originally Proceedings of the Aristotelian Society, 1948–1949).

[Hepler *et al.*, 2007] Hepler, A. B., Dawid, A. P., & Leucari, V. (2007). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability & Risk*, 6, 275- 293.

[Hitchcock, 2001] Hitchcock, D. L. (2001). John L. Pollock's theory of rationality. In Chr. W. Tindale, H. V. Hansen & E. Sveda (Eds.), *Argumentation at the Century's turn.* Ontario Society for the Study of Argumentation. (Proceedings of the 3rd OSSA conference, 1999). (CD ROM).

[Hitchcock, 2002] Hitchcock, D. L. (2002). Pollock on practical reasoning. *Informal Logic*, 22, 247–256.

[Hofstadter, 1996] Hofstadter, D. (1996). *Metamagical themas. Questing for the essence of mind and pattern.* New York: Basic Books.

[Houtlosser, 2001] Houtlosser, P. (2001). Points of view. In F. H. van Eemeren (Ed.), *Crucial concepts in argumentation theory* (pp. 27–50). Amsterdam: Amsterdam University Press.

[Hunter, 2013] Hunter, A. (2013). A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1), 47–81.

[Hunter & Williams, 2010] Hunter, A., & Williams, M. (2010). Qualitative evidence aggregation using argumentation. In P. Baroni, F. Cerutti, M. Giacomin, & G. R. Simari (Eds.), *Computational models of argument – Proceedings of COMMA 2010* (pp. 287–298). Amsterdam: IOS Press.

[Jackson & Jacobs, 1982] Jackson, S., & Jacobs, S. (1982). The collaborative production of proposals in conversational argument and persuasion. A study of disagreement regulation. *Journal of the American Forensic Association*, 18, 77–90.

[Jakobovits & Vermeir, 1999] Jakobovits, H., & Vermeir, D. (1999). Robust semantics for argumentation frameworks. *Journal of Logic and Computation*, 9(2), 215–261.

[Jensen & Nielsen, 2007] Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs.* New York: Springer.

[Johnson, 2000] Johnson, R. H. (2000). *Manifest rationality. A pragmatic theory of argument.* Mahwah, NJ: Lawrence Erlbaum.

[Johnson & Blair, 2006] Johnson, R. H., & Blair, J. A. (2006). *Logical self-defense* (reprint of Johnson & Blair, 1994). New York: International Debate Education Association. (1st ed. 1977).

[Josephson & Josephson, 1996] Josephson, J. R., & Josephson, S. G. (Eds., 1996). *Abductive inference. Computation, philosophy, technology.* Cambridge: Cambridge University Press.

[Karacapilidis & Papadias, 2001] Karacapilidis, N., & Papadias, D. (2001). Computer supported argumentation and collaborative decision making. The HERMES system. *Information Systems*, 26, 259–277.

[Kienpointner, 1992] Kienpointner, M. (1992). *Alltagslogik. Struktur and Funktion von Argumentationsmustern* (Everyday logic. Structure and functions of prototypes of argumentation). Stuttgart: Frommann-Holzboog.

[Kirschner, *et al.*, 2003] Kirschner, P. A., Buckingham Shum, S. J., & Carr, C. S. (Eds., 2003). *Visualizing argumentation. Software tools for collaborative and educational sense-*

making. London: Springer.

[Kjaerulff & Madsen, 2008] Kjaerulff, U. B., & Madsen, A. L. (2008). *Bayesian networks and influence diagrams.* New York: Springer.

[Kopperschmidt, 1989] Kopperschmidt, J. (1989). *Methodik der Argumentationsanalyse* (Methodology of argumentation analysis). Stuttgart: Frommann-Holzboog.

[Kowalski, 2011] Kowalski, R. A. (2011). *Computational logic and human thinking. How to be artificially intelligent.* Cambridge: Cambridge University Press.

[Kunz & Rittel, 1970] Kunz, W., & Rittel, H. (1970). Issues as elements of information systems, Technical Report 0131, Universität Stuttgart, Institut für Grundlagen der Planung.

[Kyburg, 1994] Kyburg, H. E. (1994). Uncertainty logics. In D. M. Gabbay, C. J. Hogger & J. A. Robinson (Eds.), *Handbook of logic in artificial intelligence and logic programming, 3. Non-monotonic reasoning and uncertain reasoning* (pp. 397–438). Oxford: Clarendon Press.

[Leff, 2003] Leff, M. (2003). Rhetoric and dialectic in Martin Luther King's 'Letter from Birmingham Jail'. In F. H. van Eemeren, J. A. Blair, Ch. A. Willard & A. F. Snoeck Henkemans (Eds.), *Anyone who has a view. Theoretical contributions to the study of argumentation* (pp. 255–268). Dordrecht: Kluwer Academic.

[Lodder, 1999] Lodder, A. R. (1999). *DiaLaw. On legal justification and dialogical models of argumentation.* Dordrecht: Kluwer.

[Lorenzen & Lorenz, 1978] Lorenzen, P., & Lorenz, K. (1978). *Dialogische Logik* (Dialogic logic). Darmstadt: Wissenschaftliche Buchgesellschaft.

[Loui, 1987] Loui, R. P. (1987). Defeat among arguments. A system of defeasible inference. *Computational Intelligence*, 2, 100–106.

[Loui, 1995] Loui, R. P. (1995). Hart's critics on defeasible concepts and ascriptivism. *The fifth international conference on artificial intelligence and law. Proceedings of the conference* (pp. 21–30). New York: ACM. Extended report available at `http://www1.cse.wustl.edu/~loui/ail2.pdf`.).

[Loui, 1998] Loui, R. P. (1998). Process and policy. Resource-bounded nondemonstrative reasoning. *Computational Intelligence*, 14, 1–38.

[Loui & Norman, 1995] Loui, R., & Norman, J. (1995). Rationales and argument moves. *Artificial Intelligence and Law*, 3, 159–189.

[Loui *et al.*, 1997] Loui, R., Norman, J., Altepeter, J., Pinkard, D., Craven, D., Linsday, J., & Foltz, M. A. (1997). Progress on room 5. A testbed for public interactive semi-formal legal argumentation. In: *Proceedings of the sixth international conference on artificial intelligence and law* (pp. 207–214). New York: ACM Press.

[Mackenzie, 1979] Mackenzie, J. D. (1979). Question-begging in non-cumulative systems. *Journal of Philosophical Logic*, 8, 117–133.

[Makinson, 1994] Makinson, D. (1994). General patterns in non-monotonic reasoning. In D. M. Gabbay, C. J. Hogger & J. A. Robinson (Eds.), *Handbook of logic in artificial intelligence and logic programming, 3. Non-monotonic reasoning and uncertain reasoning* (pp. 35–110). Oxford: Clarendon Press.

[McBurney & Parsons, 2002a] McBurney, P. & Parsons, S. (2002a). Games that agents play. A formal framework for dialogues between autonomous agents. *Journal for Logic, Language and Information*, 11, 315–334.

[McBurney & Parsons, 2002b] McBurney, P. & Parsons, S. (2002b). Dialogue games in multi-agent systems. *Informal Logic*, 22, 257–274.

[McBurney & Parsons, 2009] McBurney, P., & Parsons, S. (2009). Dialogue games for agent argumentation. In I. Rahwan & G. R. Simari (Eds.), *Argumentation in artificial intelligence* (pp. 261–280). Dordrecht: Springer.

[McBurney, Hitchcock & Parsons, 2007] McBurney, P., Hitchcock, D. L., & Parsons, S. (2007). The eightfold way of deliberation dialogue. *International Journal of Intelligent Systems*, 22, 95–132.

[McCarty, 1977] McCarty, L. (1977). Reflections on TAXMAN. An experiment in artificial intelligence and legal reasoning. *Harvard Law Review*, 90, 89–116.

[McCarty, 1995] McCarty, L. (1995). An implementation of Eisner v. Macomber. In *Proceedings of the fifth international conference on artificial intelligence and law* (pp. 276–286). New York: ACM Press.

[Mochales Palau & Moens, 2009] Mochales Palau, R., & Moens, S. (2009). Argumentation mining. The detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law (ICAIL 2009)* (pp. 98–107). New York: ACM Press.

[Modgil, 2009] Modgil, S. (2009). Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9–10), 901–934.

[Næss, 1966] Næss, A. (1966). *Communication and argument. Elements of applied semantics.* (A. Hannay, transl.). London: Allen and Unwin. (English transl. of *En del elementære logiske emner.* Oslo: Universitetsforlaget, 1947).

[Nute, 1994] Nute, D. (1994). Defeasible logic. In D. M. Gabbay, C. J. Hogger & J. A. Robinson (Eds.), *Handbook of logic in artificial intelligence and logic programming, 3. Non-monotonic reasoning and uncertain reasoning* (pp. 353–395). Oxford: Clarendon Press.

[O'Keefe, 2002] O'Keefe, D. J. (2002). *Persuasion. Theory and research.* 2nd ed. Thousand Oaks, CA: Sage. (1st ed. 1990).

[Parsons *et al.*, 1998] Parsons, S., Sierra, C., & Jennings, N. R. (1998). Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8, 261–292.

[Pearl, 1988] Pearl, J. (1988). *Probabilistic reasoning in intelligent systems. Networks of plausible inference.* San Francisco: Morgan Kaufmann Publishers.

[Pearl, 2009] Pearl, J. (2000/2009). *Causality. Models, reasoning, and inference.* 2nd ed. Cambridge: Cambridge University Press. (1st ed. 2000).

[Perelman & Olbrechts-Tyteca, 1969] Perelman, Ch., & Olbrechts-Tyteca, L. (1969). *The new rhetoric. A treatise on argumentation.* Notre Dame, IN: University of Notre Dame Press. (English transl. by J. Wilkinson and P. Weaver of Ch. Perelman & L. Olbrechts-

Tyteca (1958). *La nouvelle rhétorique. Traité de l'argumentation.* Paris: Presses Universitaires de France. (3rd ed. Brussels: Éditions de l'Université de Bruxelles)).

[Plantin, 1996] Plantin, Chr. (1996). *L'argumentation* [Argumentation]. Paris: Le Seuil.

[Pollock, 1987] Pollock, J. L. (1987). Defeasible reasoning. *Cognitive Science*, 11, 481–518.

[Pollock, 1989] Pollock, J. L. (1989). *How to build a person. A prolegomenon.* Cambridge, MA: The MIT Press.

[Pollock, 1994] Pollock, J. L. (1994). Justification and defeat. *Artificial Intelligence*, 67, 377–407.

[Pollock, 1995] Pollock, J. L. (1995). *Cognitive Carpentry. A blueprint for how to build a person.* Cambridge, MA: The MIT Press.

[Pollock, 2006] Pollock, J. L. (2006). *Thinking about acting. Logical foundations for rational decision making.* New York: Oxford University Press.

[Pollock, 2010] Pollock, J. L. (2010). Defeasible reasoning and degrees of justification. *Argument & Computation*, 1(1), 7–22.

[Poole, 1985] Poole, D. L. (1985). On the comparison of theories. Preferring the most specific explanation. In *Proceedings of the ninth international joint conference on artificial intelligence* (pp. 144–147). San Francisco: Morgan Kaufmann.

[Prakken, 1993] Prakken, H. (1993). *Logical tools for modelling legal argument.* Doctoral dissertation Free University Amsterdam.

[Prakken, 1997] Prakken, H. (1997). *Logical tools for modelling legal argument. A study of defeasible reasoning in law.* Dordrecht: Kluwer.

[Prakken, 2005a] Prakken, H. (2005a). A study of accrual of arguments, with applications to evidential reasoning. In *Proceedings of the tenth international conference on artificial intelligence and law* (pp. 85–94). New York: ACM Press.

[Prakken, 2005b] Prakken, H. (2005b). Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 15, 1009–1040.

[Prakken, 2006] Prakken, H. (2006). Formal systems for persuasion dialogue. *The Knowledge Engineering Review*, 21(2), 163–188.

[Prakken, 2009] Prakken, H. (2009). Models of persuasion dialogue. In I. Rahwan & G. R. Simari (Eds.), *Argumentation in artificial intelligence* (pp. 281–300). Dordrecht: Springer.

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

[Prakken & Sartor, 1996] Prakken, H., & Sartor, G. (1996). A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4, 331–368.

[Prakken & Sartor, 1998] Prakken, H., & Sartor, G. (1998). Modelling reasoning with precedents in a formal dialogue game. *Artificial Intelligence and Law*, 6, 231–287.

[Prakken & Sartor, 2007] Prakken, H., & Sartor, G. (2007). Formalising arguments about the burden of persuasion. In *Proceedings of the eleventh international conference on artificial intelligence and law* (pp. 97–106). New York: ACM Press.

[Prakken & Sartor, 2009] Prakken, H., & Sartor, G. (2009). A logical analysis of burdens of proof. In H. Kaptein, H. Prakken & B. Verheij (Eds.), *Legal evidence and proof. Statistics, stories, logic* (pp. 223–253). Farnham: Ashgate.

[Prakken & Vreeswijk, 2002] Prakken, H., & Vreeswijk, G. A. W. (2002). Logics for defeasible argumentation. In D. Gabbay & F. Guenthner (Eds.), *Handbook of philosophical logic.* 2nd ed., IV (pp. 219–318). Dordrecht: Kluwer.

[Rahwan & McBurney, 2007] Rahwan, I., & McBurney, P. (2007). Argumentation technology. Guest editors' introduction. *IEEE Intelligent Systems*, 22(6), 21–23.

[Rahwan & Simari, 2009] Rahwan, I., & Simari, G. R. (Eds., 2009). *Argumentation in artificial intelligence.* Dordrecht: Springer.

[Rahwan *et al.*, 2003] Rahwan, I., Ramchurn, S. D., Jennings, N. R., McBurney, P., Parsons, S., & Sonenberg, E. (2003). Argumentation-based negotiation. *Knowledge Engineering Review*, 18(4), 343–375.

[Rahwan *et al.*, 2007] Rahwan, I., Zablith, F., & Reed, Chr. (2007). Laying the foundations for a world wide argument web. *Artificial Intelligence*, 171(10–15), 897–921.

[Rao & Georgeff, 1995] Rao, A., & Georgeff, M. (1995). BDI agents. From theory to practice. In *Proceedings of the 1st international conference on multi-agent systems* (pp. 312–319).

[Reboul, 1990] Reboul, O. (1990). Rhétorique et dialectique chez Aristote [Aristotle's views on rhetoric and dialectic]. *Argumentation.* 4, 35–52.

[Reed, 1997] Reed, Chr. A. (1997). Representing and applying knowledge for argumentation in a social context. *AI and Society*, 11(3–4), 138–154.

[Reed, 1999] Reed, Chr. A. (1999). The role of saliency in generating natural language arguments. In: *Proceedings of the 16th international joint conference on AI (IJCAI'99)* (pp. 876–881). San Francisco: Morgan Kaufmann.

[Reed & Grasso, 2007] Reed, Chr. A., & Grasso, F. (2007). Recent advances in computational models of natural argument. *International Journal of Intelligent Systems*, 22, 1–15.

[Reed & Norman *et al.*, 2004] Reed, Chr. A., & Norman, T. J. (Eds., 2004). *Argumentation machines. New frontiers in argument and computation.* Dordrecht: Kluwer.

[Reed & Rowe, 2004] Reed, Chr. A., & Rowe, G. W. A. (2004). Araucaria. Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, 13(4), 961–979.

[Reed & Tindale, 2010] Reed, Chr. A., & Tindale, Chr. W. (Eds., 2010). *Dialectics, dialogue and argumentation. An examination of Douglas Walton's theories of reasoning.* London: College Publications.

[Reiter, 1980] Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13, 81–132.

[Rigotti, 2009] Rigotti, E. (2009). Whether and how classical topics can be revived within contemporary argumentation theory. In F. H. van Eemeren & B. Garssen (Eds.), *Pondering on problems of argumentation* (pp. 157–178). New York: Springer.

[Rissland & Ashley, 2002] Rissland, E. L., & Ashley, K. D. (1987). A case-based system for trade secrets law. In *Proceedings of the first international conference on artificial intelligence and law* (pp. 60–66). New York: ACM Press.

[Rissland & Ashley, 2002] Rissland, E. L., & Ashley, K. D. (2002). A note on dimensions and factors. *Artificial Intelligence and Law*, 10, 65–77.

[Rittel & Webber, 1973] Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 155–169.

[Riveret *et al.*, 2007] Riveret, R., Rotolo, A., Sartor, G., Prakken, H., & Roth, B. (2007). Success chances in argument games. A probabilistic approach to legal disputes. In A. R. Lodder & L. Mommers (Eds.), *Legal knowledge and information systems (JURIX 2007)* (pp. 99–108). Amsterdam: IOS Press.

[Rocci, 2008] Rocci, A. (2008). Modality and its conversational backgrounds in the reconstruction of argumentation. *Argumentation*, 22, 165–189.

[Roth, 2003] Roth, B. (2003). *Case-based reasoning in the law. A formal theory of reasoning by case comparison.* Doctoral dissertation University of Maastricht.

[Sartor, 2005] Sartor, G. (2005). *Legal reasoning. A cognitive approach to the law, 5. Treatise on legal philosophy and general jurisprudence.* Berlin: Springer.

[Schank, 1995] Schank, P. (1995). *Computational tools for modeling and aiding reasoning: Assessing and applying the Theory of Explanatory Coherence.* Doctoral dissertation University of California, Berkeley.

[Scheuer *et al.*, 2010] Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation. A review of the state of the art. *Computer-Supported Collaborative Learning*, 5, 43–102.

[Schiappa, 2002] Schiappa, E. (2002). Evaluating argumentative discourse from a rhetorical perspective. Defining 'person' and 'human life' in constitutional disputes over abortion. In F. H. van Eemeren & P. Houtlosser (Eds.), *Dialectic and rhetoric. The warp and woof of argumentation analysis* (pp. 65–80). Dordrecht etc.: Kluwer.

[Schwemmer & Lorenzen, 1973] Schwemmer, O., & Lorenzen, P. (1973). *Konstruktive Logik, Ethik und Wissenschaftstheorie* (Constructive logic, ethics and theory of science). Mannheim: Bibliographisches Institut.

[Simari & Loui, 1992] Simari, G. R., & Loui, R. P. (1992). A mathematical treatment of defeasible reasoning and its applications. *Artificial Intelligence*, 53, 125–157.

[Snoeck Henkemans, 2001] Snoeck Henkemans, A. F. (2001). Argumentation structures. In F. H. van Eemeren (Ed.), *Crucial concepts in argumentation theory* (pp. 101–134). Amsterdam: Amsterdam University Press.

[Suthers, 1999] Suthers, D. (1999). Representational support for collaborative inquiry. In *Proceedings of the 32nd Hawaii international conference on the system sciences (HICSS-32)* . Institute of Electrical and Electronics Engineers (IEEE).

[Suthers *et al.*, 1995] Suthers, D., Weiner, A., Connelly, J., & Paolucci, M. (1995). Belvedere. Engaging students in critical discussion of science and public policy issues. In *Proceedings of the 7th world conference on artificial intelligence in education (AIED*

Ś95) , 266–273. Washington.

[Swearngen & Schiappa, 2009] Swearingen, C. J., & Schiappa, E. (2009). Historical studies in rhetoric. Revisionist methods and new directions. In A. A. Lunsford, K. H. Wilson & R. A. Eberly (Eds.), *The Sage handbook of rhetorical studies* (pp. 1–12). Los Angeles, CA: Sage.

[Sycara, 1989] Sycara, K. (1989). Argumentation. Planning other agents' plans. In *Proceedings of the eleventh international joint conference on artificial intelligence* (pp. 517–523).

[Talbott, 2011] Talbott, W. (2011). Bayesian epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Summer 2011 ed. `https://plato.stanford.edu/entries/epistemology-bayesian/`.

[Taroni *et al.*, 2006] Taroni, F., Aitken, C., Garbolino, P., & Biedermann, A. (2006). *Bayesian networks and probabilistic inference in forensic science.* Chichester: Wiley.

[Teufel, 1999] Teufel, S. (1999). *Argumentative zoning. Information extraction from scientific articles.* Doctoral dissertation University of Edinburgh.

[Thagard, 1992] Thagard, P. (1992). *Conceptual revolutions.* Princeton, NJ: Princeton University Press.

[Timmer *et al.*, 2017] Timmer, S., Meyer, J.J., Prakken, H., Renooij, S., & Verheij, B. (2017). A two-phase method for extracting explanatory arguments from Bayesian Networks. *International Journal of Approximate Reasoning*, 80, 475–494.

[Tindale, 1999] Tindale, Chr. W. (1996). From syllogisms to audiences. The prospects for logic in a rhetorical model of argumentation. In D. M. Gabbay & H. J. Ohlbach, *Practical reasoning. Proceedings of FAPR 1996* (pp. 596–605). Berlin: Springer.

[Tindale, 2004] Tindale, Chr. W. (1999), *Acts of arguing. A rhetorical model of argument.* Albany, NY: State University of New York Press.

[Toulmin, 2003] Toulmin, S. E. (1958). *The uses of argument.* Cambridge, England: Cambridge University Press. (updated ed. 2003).

[Verheij, 1996a] Verheij, B. (1996a). *Rules, reasons, arguments. Formal studies of argumentation and defeat.* Doctoral dissertation University of Maastricht.

[Verheij, 1996b] Verheij, B. (1996b). Two approaches to dialectical argumentation. Admissible sets and argumentation stages. In J.-J. Ch. Meyer & L. C. van der Gaag (Eds.), *NAIC'96. Proceedings of the eighth Dutch conference on artificial intelligence* (pp. 357–368). Utrecht: Utrecht University.

[Verheij, 2003a] Verheij, B. (2003a). DefLog. On the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation*, 13(3), 319–346.

[Verheij, 2003b] Verheij, B. (2003b). Dialectical argumentation with argumentation schemes. An approach to legal logic. *Artificial Intelligence and Law*, 11(1–2), 167–195.

[Verheij, 2005a] Verheij, B. (2005a). Evaluating arguments based on Toulmin's scheme. *Argumentation*, 19, 347–371. (Reprinted in D. L. Hitchcock & B. Verheij (Eds.) (2006), *Arguing on the Toulmin model. New essays in argument analysis and evaluation* (pp. 181–202). Dordrecht: Springer).

[Verheij, 2005b] Verheij, B. (2005b). *Virtual arguments. On the design of argument assistants for lawyers and other arguers.* The Hague: T. M. C. Asser Press.

[Verheij, 2007] Verheij, B. (2007). A labeling approach to the computation of credulous acceptance in argumentation. In M. M. Veloso (Ed.), *IJCAI 2007, Proceedings of the 20th international joint conference on artificial intelligence* (pp. 623–628). Hyderabad, India.

[Verheij, 2012] Verheij, B. (2012). Jumping to conclusions. A logico-probabilistic foundation for defeasible rule-based arguments. In L. Fariñas del Cerro, A. Herzig & J. Mengin (Eds.), *Logics in artificial intelligence. 13th European conference, JELIA 2012. Toulouse, France, September 2012. Proceedings (LNAI 7519)* (pp. 411–423). Springer, Berlin.

[Verheij, 2017] 205 Verheij, B. (2017). Proof with and without probabilities. Correct evidential reasoning with presumptive arguments, coherent hypotheses and degrees of uncertainty. *Artificial Intelligence and Law*, 25 (1), 127–154.

[Verheij *et al.*, 2016] Verheij, B., Bex, F.J., Timmer, S., Vlek, C., Meyer, J.J., Renooij, S., & Prakken, H. (2016). Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning. *Law, Probability & Risk*, 15, 35–70.

[Verheij *et al.*, 1998] Verheij, B., Hage, J. C., & Herik, H. J. van den (1998). An integrated view on rules and principles. *Artificial Intelligence and Law*, 6(1), 3–26.

[Vlek *et al.*, 2014] Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2014). Building Bayesian Networks for legal evidence with narratives: a case study evaluation. *Artificial Intelligence and Law*, 22 (4), 375–421.

[Vlek *et al.*, 2016] Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2016). A method for explaining Bayesian Networks for legal evidence with scenarios. *Artificial Intelligence and Law*, 24 (3), 285–324.

[Vreeswijk, 1993] Vreeswijk, G. A. W. (1993). *Studies in defeasible argumentation.* Doctoral dissertation Free University Amsterdam.

[Vreeswijk, 1995a] Vreeswijk, G. A. W. (1995a). Formalizing nomic. Working on a theory of communication with modifiable rules of procedure, Technical Report CS 95–02, Vakgroep Informatica (FdAW), Rijksuniversiteit Limburg, Maastricht.

[Vreeswijk, 1995b] Vreeswijk, G. A. W. (1995b). The computational value of debate in defeasible reasoning. *Argumentation*, 9, 305–342.

[Vreeswijk, 1997] Vreeswijk, G. A. W. (1997). Abstract argumentation systems. *Artificial Intelligence*, 90, 225–279.

[Vreeswijk, 2000] Vreeswijk, G. A. W. (2000). Representation of formal dispute with a standing order. *Artificial Intelligence and Law*, 8, 205–231.

[Walton, 1987] Walton, D. N. (1987). *Informal fallacies. Towards a theory of argument criticisms.* Amsterdam: John Benjamins.

[Walton, 1989] Walton, D. N. (1989). *Informal logic. A handbook for critical argumentation.* Cambridge: Cambridge University Press.

[Walton, 1996] Walton, D. N. (1996). *Argumentation schemes for presumptive reasoning.*

Mahwah, NJ: Lawrence Erlbaum.

[Walton & Krabbe, 1995] Walton, D. N., & Krabbe, E. C. W. (1995). *Commitment in dialogue. Basic concepts of interpersonal reasoning.* Albany, NY: State University of New York Press.

[Walton *et al.*, 2008] Walton, D. N., Reed, Chr. A., & Macagno, F. (2008). *Argumentation schemes.* Cambridge: Cambridge University Press.

[Wenzel, 1987] Wenzel, J. W. (1987). The rhetorical perspective on argument. In F. H. van Eemeren, R. Grootendorst, J. A. Blair & Ch. A. Willard (Eds.), *Argumentation. Across the lines of discipline. Proceedings of the conference on argumentation 1986* (pp. 101–109). Dordrecht/Providence: Foris.

[Wigmore, 1931] Wigmore, J. H. (1931). *The principles of judicial proof.* 2nd ed. Boston: Little Brown & Company. (1st ed. 1913).

[Woods & Walton, 1989] Woods, J., & Walton, D. N. (1989). *Fallacies. Selected papers 1972–1982.* Berlin/Dordrecht/Providence: de Gruyter/Foris.

[Wooldridge, 2009] Wooldridge, M. (2009). *An introduction to multiagent systems.* Chichester: Wiley.

[Zarefsky, 2006] Zarefsky, D. (2006). Strategic maneuvering through persuasive definitions. Implications for dialectic and rhetoric. *Argumentation*, 20(4), 399–416.

[Zarefsky, 2009] Zarefsky, D. (2009). Strategic maneuvering in political argumentation. In F. H. van Eemeren (Ed., 2009), *Examining argumentation in context. Fifteen studies on strategic maneuvering* (pp. 115–130). Amsterdam: John Benjamins.

[Zukerman *et al.*, 1998] Zukerman, I., McConachy, R., & Korb, K. (1998). Bayesian reasoning in an abductive mechanism for argument generation and analysis. In *Proceedings of the fifteenth national conference on artificial intelligence (AAAI-98, Madison)* (pp. 833–838). Menlo Park: AAAI Press.

 Received 14 July 2017