Chapter 11
The Toulmin Argument Model in Artificial Intelligence
Or: how semi-formal, defeasible argumentation schemes creep into logic

Bart Verheij

1 Toulmin's 'The Uses of Argument'

In 1958, Toulmin published *The Uses of Argument*. Although this anti-formalistic monograph initially received mixed reviews (see section 2 of [20] for Toulmin's own recounting of the reception of his book), it has become a classical text on argumentation, and the number of references to the book (when writing these words[1] — by a nice numerological coincidence — 1958) continues to grow (see [7] and the special issue of *Argumentation* 2005; Vol. 19, No. 3). Also the field of Artificial Intelligence has discovered Toulmin's work. Especially four of Toulmin's themes have found follow-up in Artificial Intelligence. First, argument analysis involves half a dozen distinct elements, not just two. Second, many, if not most, arguments are substantial, even defeasible. Third, standards of good reasoning and argument assessment are non-universal. Fourth, logic is to be regarded as generalised jurisprudence. Using these central themes as a starting point, this chapter provides an introduction to Toulmin's argument model and its connections with Artificial Intelligence research. No attempt is made to give a comprehensive history of the reception of Toulmin's ideas in Artificial Intelligence; instead a personal choice is made of representative steps in AI-oriented argumentation research.

When Toulmin wrote his book, he was worried. He saw the influence of the successes of formal logic on the philosophical academia of the time, and was afraid that as a consequence seeing formal logic's limitations would be inhibited. He wrote *The Uses of Argument* to fight the — in his opinion mistaken — idea of formal logic as a universal science of good reasoning. In the updated edition of *The Uses of Argument* [19], he describes his original aim as follows:

> to criticize the assumption, made by most Anglo-American academic philosophers, that any significant argument can be put in formal terms: not just as a syllogism, since for Aristotle himself any inference can be called a 'syllogism' or 'linking of statements', but a rigidly demonstrative deduction of the kind to be found in Euclidean geometry. ([19], vii)

Bart Verheij
Artificial Intelligence, University of Groningen

[1] Source: Google Scholar citation count, April 1, 2008.

I. Rahwan, G. R. Simari (eds.), *Argumentation in Artificial Intelligence,*
DOI 10.1007/978-0-387-98197-0_11,

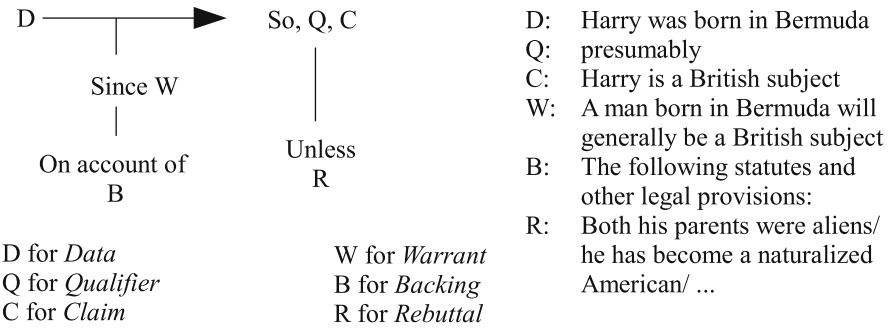

Fig. 11.1 Toulmin's layout of arguments with an example ([18], 104–5)

In short: Toulmin wanted to argue that there are other arguments than formal ones. It is also clear from this quote that Toulmin's goals were first and foremost aimed at his fellow philosophers. In the Preface to the 2003 edition, Toulmin says it thus:

> In no way had I set out to expound a theory of rhetoric or argumentation: my concern was with twentieth-century epistemology, not informal logic. ([19], vii)

Let us look closer at some of Toulmin's points.

1.1 Argument analysis involves half a dozen distinct elements, not just two

Toulmin is perhaps most often read because of his argument diagram (Figure 1). Whereas a formal logical analysis uses the dichotomy of premises and conclusions when analyzing arguments, Toulmin distinguishes six different kinds of elements: Data, Claim, Qualifier, Warrant, Backing and Rebuttal. Before explaining the roles of these elements, let us look at Toulmin's famous example of Harry, who may or may not be a British subject.

When someone claims that Harry is a British subject, it is natural to ask, so says Toulmin: What have you got to go on? An answer to that question can provide the data on which the claim rests, here: Harry was born in Bermuda. But having datum and claim is not enough. A further important question needs to be answered. Toulmin phrases it thus: How do you get there? In other words, why do you think that the datum gives support for your claim? An answer to this question must take the form of a rule-like general statement, the warrant underlying the step from datum to claim. In the example, the warrant is that a man born in Bermuda will generally be a British subject. As the example shows, warrants need not express universal generalizations. Here the warrant is not that *each* man born in Bermuda is a British subject, but merely that a man born in Bermuda will *generally* be a British subject. As a result, on the basis of datum and warrant a claim needs to be qualified. Here the claim becomes that *presumably* Harry is a British subject. When datum, qualified claim and warrant have been made explicit, a further question needs to be asked: Why do you think that the warrant holds? An answer will be provided by the backing of the warrant. In the example, Toulmin refers to the existence of statutes and other legal provisions (without specifying them) that can provide the backing

for the warrant that who is born in Bermuda will generally be British subjects. The sixth and final kind of element to be distinguished is that of conditions of exception or rebuttal (101/93)[2]. Conditions of rebuttal indicate 'circumstances in which the general authority of the warrant would have to be set aside' or 'exceptional circumstances which might be capable of defeating or rebutting the warranted conclusion' (101/94). In the example, Harry's parents could be aliens or he could have become a naturalized American. Toulmin refers to Hart and Ross as predecessors for his discussion of rebuttal. Hart coined the term 'defeasibility' (see also [9]) and used it in legal and philosophical settings (contract, free will, responsibility), while Ross emphasized that moral rules must have exceptions (142/131–2).

Here is Toulmin's defence of the difference between a datum and the negation of a rebuttal, which predates discussions about the relation between rule conditions and exceptions:

> [T]he fact that Harry was born in Bermuda and the fact that his parents were not aliens are both of them directly relevant to the question of his present nationality; but they are relevant in different ways. The one fact is a datum, which by itself establishes a presumption of British nationality; the other fact, by setting aside one possible rebuttal, tends to confirm the presumption thereby created. (102/95)

Summarizing, Toulmin distinguishes six kinds of elements in arguments:

> *Claim*: The Claim is the original assertion that we are committed to and must justify when challenged (97/90). It is the starting point of the argument.
> *Datum*: The Datum provides the basis of the claim in response to the question: What have you got to go on? (97–8/90)
> *Warrant*: The Warrant provides the connection between datum and claim. A warrant expresses that '[d]ata such as D entitle[s] one to draw conclusions, or make claims, such as C'. Warrants are 'general, hypothetical statements, which can act as bridges, and authorise the sort of step to which our particular argument commits'. They are 'rules, principles, inference-licences or what you will, instead of additional items of information'. (98/91)
> *Qualifier*: The Qualifier indicates the strength of the step from datum to claim, as conferred by the warrant (101/94)
> *Backing*: The Backing shows why a warrant holds. Backing occurs when not a particular claim is challenged, but the range of arguments legitimized by a warrant (103–4/95–6).
> *Rebuttal*: A Rebuttal can indicate 'circumstances in which the general authority of the warrant would have to be set aside' or 'exceptional circumstances which might be capable of defeating or rebutting the warranted conclusion' (101/94).

1.2 Many, if not most, arguments are substantial, even defeasible

Let us consider another of Toulmin's example arguments:

> (1) Anne is one of Jack's sisters;
> All Jack's sisters have red hair;
> So, Anne has red hair. (123/115)

This example is a variant of the paradigmatic example of a syllogism: 'Socrates is a man. All men are mortal. So, Socrates is mortal'. Anyone accustomed to the

[2] Page numbers before the slash refer to the original 1958 edition of *The Uses of Argument* [18], those after the slash to the updated 2003 edition [19]

standard logical treatment of syllogisms will recognize the following logical form underlying this type of syllogistic argument:

(2) $P(t)$
$(\forall x)\ (P(x) \rightarrow Q(x))$
$\overline{\qquad\qquad\qquad}$
$Q(t)$

In this logical analysis, argument (1) has two premises and one conclusion. Moreover, the two premises have clearly distinctive roles, one often referred to as the minor premise ($P(t)$), the other the major premise ($(\forall x)\ (P(x) \rightarrow Q(x))$).

If we look at example (1) and its logical analysis (2) — an analysis to which Toulmin does not object for *this* type of argument[3] — one may ask: Why the richness of primitives in his scheme? Shouldn't we apply Occam's razor and be satisfied with the good-old dichotomy of premises and conclusions instead of Toulmin's six-fold scheme?

Toulmin's answer is: no, we shouldn't be satisfied. A central place in the defence of his position is the claim that syllogistic arguments of the logical form in (2) are atypical, even rare (147–150/136–139; also: 125–6/116–7). They have special characteristics that do not hold for other kinds of arguments. Toulmin discusses the following five characteristics of (2)-fitting arguments:

1. They are *unequivocal in their consequences.*
However, there are also arguments (e.g., the Harry example; Figure 1) that only allow drawing a conclusion tentatively. Hence the need for qualifiers.
2. They are *formally valid*.
Toulmin speaks of a formally valid argument when the argument's conclusion can be achieved by 'shuffling' the premises and their constituent parts (118/110). Arguments of the form 'D; W. So C' can in this way be phrased as formally valid, but arguments of the form 'D; B. So C' normally cannot. Toulmin refers to the Harry example (Figure 1) to make his point (123/114).
3. They are expressed in terms of *'logical words'*.
Toulmin says it thus: 'The acceptable, logical words include 'all', 'some', 'or', and a few others: these are firmly herded away from the non-logical goats, i.e. the generality of nouns, adjectives and the like, and unruly connectives and quantifiers such as 'most', 'few', 'but'.' (149/138)
4. They are *warrant-using*.
But, says Toulmin, there are also warrant-establishing arguments, as they for instance occur in scientific papers (120–1/112–3). Toulmin refers to Ryle, who contrasted warrant-using and warrant-establishing arguments by the analogy of taking a journey along a railway already built and the building of a fresh railway. Toulmin connects warrant-using arguments to the term 'deduction', and warrant-establishing arguments to 'induction'.
5. They are *analytic*.
Toulmin calls an argument analytic if and only if the backing for the warrant authorising it includes, explicitly or implicitly, the information conveyed in the conclusion itself. For instance, the universal statement that all of Jack's sisters have red hair, in a way includes that Anne, who is one of Jack's sisters, has red hair (123–127/114–118). Arguments that are not analytic are substantial.

[3] The class of analytic arguments, for which both 'D. W. So, C' and 'D. B. So, C' can be expressed in the formally valid way (2) (123/114).

It is because of the accidental concurrence of these five properties in (2)-arguments that the idea has come about that *all* arguments have them, and must have them; and this is an unfortunate fact of history, says Toulmin.

In the connection of analytic versus substantial arguments, Toulmin distinguishes two variants of the Anne argument (1) (124/115):

(1, backing version)
Anne is one of Jack's sisters;
Each one of Jack's sisters has (been checked individually to have) red hair;
So, Anne has red hair.
(1, warrant version)
Anne is one of Jack's sisters;
Any sister of Jack's will (i.e. may be taken to) have red hair;
So, Anne has red hair.

Note the different phrasing of the general statement used in the first and the second variant. In the former, it is formulated as a backing, here taking the form of an empirical fact about Jack's sisters, thereby encompassing the instance of sister Anne having red hair (at a certain moment). In the latter, it is formulated as a warrant, i.e., an inference-licensing general statement ('may be taken to'). The former can be used as a backing for the second.

Under which circumstances is (1, backing version) a genuinely analytic argument defending the claim that Anne has red hair? Well, says Toulmin, 'only if at this very moment I have all of Jack's sisters in sight. .... The thing to do now is use one's eyes, not hunt up a chain of reasoning' (126/117). The 'So' in the argument could be just as well replaced by 'In other words' or 'That is to say'.

In all other situations (which is: most), the conclusion will not be given with datum and backing, hence the argument will be a substantial one. Toulmin continues: 'If the purpose of an argument is to establish conclusions about which we are not entirely confident by relating them back to other information about which we have greater assurance, it begins to be a little doubtful whether any genuine, practical argument could ever be properly analytic.'

Here is how Toulmin extends the Anne example, making it fit his own format:

Datum:	Anne is one of Jack's sisters.
Claim:	Anne has red hair.
Warrant:	Any sister of Jack's will (i.e. may be taken to) have red hair.
Backing:	All his sisters have previously been observed to have red hair.
Qualifier:	Presumably
Rebuttal:	Anne has dyed/gone white/lost her hair ...

Note how Toulmin has added a qualifier and rebuttals, even though the backing assumes that all sisters have been checked. But Toulmin says rightly that normally an argument like this occurs later than at the time of establishing the warrant by the backing; hence making it non-demonstrative/subject to exceptions/defeasible/..., hence substantial, and not analytic (in Toulmin's sense). Checking hair colour today is not a guarantee for hair colour tomorrow.

Toulmin mentions one field in which arguments seem to be safe: mathematics. But then again: 'As a model argument for formal logicians to analyse, it [i.e., a

solution to a mathematical problem] may be seducingly elegant, but it could hardly be less representative' (127/118).

1.3 Standards of good reasoning and argument assessment are not universal, but context-dependent

According to Toulmin, our standards for the assessment of real arguments are not universal, but depend on a context. In a section, where he discusses this issue, he uses the term 'possibility' as an illustration (36/34): whereas in mathematics 'possibility' has to do with the absence of demonstrable contradiction, in most cases 'possibility' is based on a stronger standard. His example statement is 'Dwight D. Eisenhower will be selected to represent the U.S.A. in the Davis Cup match against Australia'. This statement involves no contradiction, while still (now former, then actual) President Eisenhower will not be considered a possible team member. In other words, 'possibility' is judged using different standards, some more formal ('absence of contradiction'), others more substantial ('being a top-level tennis player').

The example is however an example of different standards for the possible, not of different standards for the assessment of arguments, his ultimate aim. Here is a succinct phrasing of his position:

> It is unnecessary, we argued, to freeze statements into timeless propositions before admitting them into logic: utterances are made at particular times and in particular situations, and they have to be understood and assessed with one eye on this context. The same, we can now argue, is true of the relations holding between statements, at any rate in the majority of practical arguments. The exercise of the rational judgement is itself an activity carried out in a particular context and essentially dependent on it: the arguments we encounter are set out at a given time and in a given situation, and when we come to assess them they have to be judged against this background. So the practical critic of arguments, as of morals, is in no position to adopt the mathematician's Olympian posture. (182–3/168–9; the quote appears in a section entitled "Logic as a System of Eternal Truths")

According to Toulmin, the differences between standards of reasoning are reflected in the backings that are accepted to establish warrants. For instance, he considers the following three warrants (103–4/96):

> A whale will be a mammal.
> A Bermudan will be a Briton.
> A Saudi Arabian will be a Muslim.

Each of these warrants gives in a similar way the inferential connection between certain types of data and certain kinds of claims. The first allows inferring that a *particular* whale is a mammal, the second that a *particular* Bermudan is a Briton, the third that a *particular* Saudi Arabian is a Muslim (all these inferences, of course, subject to qualification and rebuttal). The different standards become visible when information about the corresponding backings is inserted:

> A whale will be (i.e. *is classifiable as*) a mammal
> A Bermudan will be (*in the eyes of the law*) a Briton
> A Saudi Arabian will be (*found to be*) a Muslim

Toulmin explains (104/96):

> One warrant is defended by relating it to a system of taxonomical classification, another by appealing to the statutes governing the nationality of people born in the British colonies, the third by referring to the statistics which record how religious beliefs are distributed among people of different nationalities.

For Toulmin, the establishment of standards of argument assessment, hence of good reasoning, is an empirical question, cf. the following excerpt:

> Accepting the need to begin by collecting for study the actual forms of argument current in any field, our starting-point will be confessedly empirical: we shall study ray-tracing techniques because they are used to make optical inferences, presumptive conclusions and 'defeasibility' as an essential feature of many legal arguments, axiomatic systems because they reflect the pattern of our arguments in geometry, dynamics and elsewhere. (257/237)[4]

Toulmin goes one step further. Our standards of good reasoning are not only to be established empirically, they are also to be considered historically: they change over time and can be improved upon:

> To think up new and better methods of arguing in any field is to make a major advance, not just in logic, but in the substantive field itself: great logical innovations are part and parcel of great scientific, moral, political or legal innovations. [...] We must study the ways of arguing which have established themselves in any sphere, accepting them as historical facts; knowing that they may be superseded, but only as the result of a revolutionary advance in our methods of thought. (257/237)

Because of his views that standards of good reasoning and argument assessment are non-universal and depend on field, even context, Toulmin has been said to revive Aristotle's Topics (Toulmin 2003, viii).

1.4 Logic is generalised jurisprudence

Toulmin discusses the relation of logic with a number of research areas (3–8/3–8). When logic is regarded as *psychology*, it deals with the laws of thought, distinguishing between what is normal and abnormal, thereby perhaps even allowing a kind of "psychopathology of cognition" (5/5). In logic as psychology, the goal is at heart descriptive: to formulate generalisations about thinkers thinking. But logic can also be seen as a kind of *sociology*. Then it is not individual thinkers that are at issue, but the focus is on general habits and practices. Here Toulmin refers to Dewey, who explains the passage from the customary to the mandatory: inferential habits can turn into inferential norms. Logic can also be regarded as a kind of *technology*, i.e., as providing a set of recipes for rationality or the rules of a craft. Here he speaks of logic as an art, like medicine. In this analogy, logic aims at the formulation of maxims, 'tips', that remind thinkers how they should think. And then there is logic as *mathematics*. There the goal of logic becomes to find truths about logical relations. There is no connection with thinking and logic becomes an objective science.

[4] Toulmin here considers the study of defeasibility an empirical question, to be performed by looking at the law! Toulmin has predicted history, by foreseeing what actually has happened and still is happening in the field of AI & law.

Finally Toulmin comes to the metaphor that he prefers and uses as the basis for his work: to view logic as *jurisprudence*:

> Logic is concerned with the soundness of the claims we make-with the solidity of the grounds we produce to support them, the firmness of the backing we provide for them-or, to change the metaphor, with the sort of case we present in defence of our claims. (7/7)

The jurisprudence metaphor emphasises the critical, procedural function of logic, thereby fundamentally changing the perspective on logic. It helps to change logic from an 'idealised logic' to a 'working logic' (cf. the title of the fourth essay in *The Uses of Argument*). At the end of his book he says that jurisprudence should not be seen as merely an analogy, but, more strongly, as providing an example to follow, as being a kind of 'best practice':

> Jurisprudence is one subject which has always embraced a part of logic within its scope, and what we called to begin with 'the jurisprudential analogy' can be seen in retrospect to amount to something more than a mere analogy. If the same as has long been done for legal arguments were done for arguments of other types, logic would make great strides forward. (255/235)

2 The reception and refinement of Toulmin's ideas in AI

The reception of Toulmin's ideas is marked by historical happenstance. It was already mentioned that his original audience, primarily the positivist, logic-oriented philosophers of knowledge of the time, was on the whole critical. For Toulmin's main messages to be appreciated a fresh crowd was needed. It was found in a radical movement in academic research and education refocusing on the analysis and assessment of real-life argument. This movement, referred to by names such as speech communication, informal logic and argumentation theory, started to blossom from the 1970s, continuing so to the present day (see [22]). One thing that Toulmin and this movement shared was the relativising, at times antagonistic, attitude towards logic as a formal science. The swing had swung back to exploring the possibilities of more formal approaches in the 1990s, when Toulmin's project of treating logic as a generalised jurisprudence was almost literally taken up in the field of Artificial Intelligence and Law (see Feteris' [4] for a related development in argumentation theory). Successful attempts were made to formalize styles of legal reasoning in a way that respected actual legal reasoning. The approach taken in this field was rooted in an independent development in Artificial Intelligence, where so-called nonmonotonic logics were studied from the 1980s. In that line of research, formal logical systems were studied that allowed for the retraction of conclusions when new information, indicating exceptional or contradictory circumstances, became available. Also in the 1990s, the study of nonmonotonic logics evolved towards what might be called argumentation logics. More generally, attention was reallocated to implemented systems and an agent-oriented perspective.

The following does not give a fully representative, historical account of AI work taking up Toulmin's ideas. A personal choice of relevant research has been made in

order to highlight how Toulmin's points of view have been adopted and refined in Artificial Intelligence.

2.1 Reiter's default rules

An early strand of research in Artificial Intelligence, in which a number of Toulmin's key positions are visible, is Reiter's work on the logic of default reasoning [15]. Reiter's formalism is built around the concept of a default: an expression $\alpha : M\beta_1, ..., M\beta_n / \gamma$, in which α, β_1, ..., β_n, and γ are sentences of first-order logic. Defaults are a kind of generalized rules of inference. The sentence α is the default's prerequisite, playing the role of what Toulmin refers to as the datum. The sentence γ is the default's consequent, comparable to Toulmin's notion of a claim. The sentences β_i are called the default's justifications. A default expresses that its consequent follows given its prerequisite, but only when its justifications can consistently be assumed.

Reiter does not refer to Toulmin in his highly influential 1980 paper, nor in his other work. Being thoroughly embedded in the fertile logic-based AI community of the time, Reiter does not refer to less formal work. Still, in Reiter's work two important ideas defended by Toulmin recur in a formal version. The first is the idea of defeasibility. As said, Reiter's defaults are a kind of generalized rules of inference, but of a defeasible kind. For instance, the default $p : M\neg e / q$ expresses that q follows from p unless $\neg e$ cannot be assumed consistently. Reiter's formal definitions are such that, given only p, it follows that q, while if both p and e are given q does not follow. Reiter's justifications are hence closely related to Toulmin's rebuttals, but as opposites: in our example the opposite e of the default's justification $\neg e$ is a kind of rebuttal in Toulmin's sense. This holds more generally: opposites of justifications can be thought of as formal versions of Toulmin's rebuttals.

There is a second way in which Reiter's work formally explicates one of Toulmin's prime concerns: defaults are contingent rules of inference, in the sense that they are not fixed in the logical system, as is the case for the natural deduction rules of first-order logic. Concretely, in Reiter's approach, defaults are part of the theory from which consequences can be drawn, side by side with the other, factual, information. One can therefore say that Toulmin's creed that standards of reasoning are field-dependent has found a place in Reiter's work. There is one important limitation however. Although Reiter's defaults can be used to construct arguments — in what Toulmin refers to as warrant-using arguments —, they cannot be argued about. In other words, there is no counterpart of warrant-establishing arguments. A default can for instance not have a default as its conclusion. Since Reiter's defaults are givens, it is not possible to give reasons for why they hold. Whereas in Toulmin's model warrants do not stand by themselves, but can be given support by backings, this has no counterpart for Reiter's defaults. (See section 2.7 for an approach to warrant-establishing arguments.)

How did Reiter extend or refine Toulmin? The first way is obvious: Reiter has given a precise explication of a part of Toulmin's notions, which is a direct consequence of the fact that Reiter's approach is formally specified, whereas Toulmin's only exists in the form of an informal philosophical essay. Reiter has shown that it is possible to give a formal elaboration of rebuttals and of warrants.

The other way is perhaps more important, as it concerns a genuine extension of what Toulmin had in mind: Reiter's logical formalism proposes a way of determining which consequences follow from given information. The key formal notion is that of an extension of a default theory (consisting of a set of factual assumptions and a set of defaults), which can be thought of as a possible set of consequences of the theory. Essentially, a set of sentences S is an extension of a default theory if S is equal to the set of consequences of the factual information that one obtains by applying a subset of the defaults, namely those defaults the justifications of which are consistent with S. (Note that S occurs in the definiens and in the definiendum.) Let me show how and to what extent Reiter's system can formalize Toulmin's Harry-example. We will leave out the qualifier and the backing as these have no obvious counterpart in Reiter's work. The core of a formalization of the Harry example is the default $d(x) : \mathrm{M}\neg r_1(x), \mathrm{M}\neg r_2(x) \,/\, c(x)$ and its instance $d(t) : \mathrm{M}\neg r_1(t), \mathrm{M}\neg r_2(t) \,/\, c(t)$. The following code is used:

t	Harry
$d(t)$	Harry was born in Bermuda
$c(t)$	Harry is a British subject
$r_1(t)$	Both his parents were aliens
$r_2(t)$	He has become a naturalized American

The default expresses that it follows that Harry is a British subject given that he is born in Bermuda, as long as it can be consistently assumed that his parents are *not* aliens and he has *not* become a naturalized American. Note that the default is a kind of hybrid of the example's warrant and rebuttals and that the default's list of justifications is not open-ended (in contrast with Toulmin's list of rebuttals).

Now consider two sets of sentences: S_1, the first-order closure of $d(t)$, $c(t)$ and S_2, the closure of $d(t)$, $r_1(t)$. Then S_1 is the unique extension of the theory consisting of the default and the factual information $d(t)$, while S_2 is the unique extension of the theory consisting of the default and the factual information $d(t)$ and $r_1(t)$. (Analogous facts hold when the other rebuttal $r_2(t)$ is used.) These facts can be interpreted as saying that, given the warrant encoded by the default, the claim follows from the data, but only when there is no rebuttal.

Three ways in which this formal version refines Toulmin's treatment seem noteworthy. First, here there is a distinction between 'generic' warrants and 'specific' warrants. The former is for Toulmin a pleonasm, while he does not consider the latter. Here the distinction is clear: on the one hand there is the generic inference license that a man born in Bermuda will generally be a British subject, on the other the specific inference license that if Harry was born in Bermuda, he is a British subject. Second, whereas Toulmin only treats single, unstructured sentences, Reiter's formal system inherits the elegant additional structuring of first-order sentences. For instance, disjunction and conjunction are directly inherited. Third and finally, Reiter's version specifies what happens when there is more than one default. Especially, his version incorporates naturally the situations that a sentential element (datum, rebuttal, ...) of one instance of Toulmin's model can be the claim of another. It has sometimes been charged against Toulmin that his model does not allow such recursiveness.

2.2 Pollock's undercutting and rebutting defeaters

Pollock's work on the philosophy and AI of argumentation has rightly achieved recognition in today's argumentation research. He can be regarded as being the first who combined theoretical, computational and practical considerations in his design of an 'artificial person', OSCAR (see, e.g., [12]). In this high ambition, he has had no followers. Pollock's work started with roots close to Toulmin's original audience, namely philosophers of knowledge. Gradually he began using methods from the field of Artificial Intelligence, where his ideas have gained most attention. Pollock does not seem to have been directly influenced by Toulmin. In his [11], where Pollock connects philosophical approaches to defeasibility with AI approaches, he cites work on defeasibility by Chisholm (going back to 1957, hence a year before Toulmin's *The Uses of Argument*) and himself (going back to 1967).

Here are some of Pollock's definitions [11]:

> P is a *prima facie reason* for S to believe Q if and only if P is a reason for S to believe Q and there is an R such that R is logically consistent with P but (P & R) is not a reason for S to believe Q. R is a *defeater* for P as a prima facie reason for Q if and only if P is a reason for S to believe Q and R is logically consistent with P but (P & R) is not a reason for S to believe Q.

So prima facie reasons are reasons that sometimes lead to their conclusion, but not always, namely not when there is a defeater. There is a close connection with non-monotonic consequence relations: when P is a prima facie reason for Q, and R is a defeater for P as a reason, then Q follows from P, but not from P & R. Pollock goes on to distinguish between two kinds of defeaters:

> R is a *rebutting defeater* for P as a prima facie reason for Q if and only if R is a defeater and R is a reason for believing $\sim$Q. R is an *undercutting defeater* for P as a prima facie reason for S to believe Q if and only if R is a defeater and R is a reason for denying that P wouldn't be true unless Q were true.

Undercutting defeaters only attack the inferential connection between reason and conclusion, whereas a defeater is rebutting if it is also a reason for the opposite of the conclusion. Pollock remarks that 'P wouldn't be true unless Q were true' is a kind of conditional, different from the material conditional of logic, but having learnt from an initial analysis, which he no longer finds convincing, he maintains that it is otherwise not clear how to analyze this conditional ([11], 485).[5]

Pollock's finding that there are different kinds of defeaters has been recognized as an important contribution both for the theory and for the practical analysis of arguments. Nothing of the sort can be found in Toulmin's *The Uses of Argument*.[6] (See section 2.7 for more on different conceptions of a rebuttal.)

---

[5] As far as I know, Pollock's later work (e.g., his [12]) does not contain a new analysis of this conditional. See section 2.7 for an approach addressing this.

[6] Pollock's work contributes significantly to several other aspects of argumentation (e.g., argument evaluation, semi-formal rules of inference and software implementation). See also Verheij's discussion [26], 104–110.

2.3 Prakken, Sartor & Hage on reasoning with legal rules

Toulmin's idea that logic should be regarded as a generalised form of jurisprudence (section 1.4), was taken up seriously in the 1990s in the field of Artificial Intelligence and Law. The work by Prakken, Sartor and Hage on reasoning with legal rules [13, 6] is representative.[7]

Influenced by logic-based knowledge representation (see, e.g., chapter 10 of [16]), Prakken & Sartor and Hage use an adapted first-order language as the basis of their formalism. For instance, here is a formal version of the rule that someone has legal capacity unless he can be shown to be a minor ([13], 340):

r_1: $\sim x$ is a minor $\Rightarrow$ x has legal capacity

Here r_1 is the name of the rule, which can be used to refer to it, and 'x is a minor' and 'x has legal capacity' are unary predicates. The tilde represents so-called weak negation, which here means that the rule's antecedent is fulfilled when it cannot be shown that x is a minor. If ordinary negation were used, the fulfilment of the antecedent would require something stronger, namely that it can be shown that x is not a minor.

In the system of Prakken & Sartor, arguments are built by applying Modus ponens to rules. There are two ways in which arguments can attack each other. First, an argument can attack a weakly negated assumption in the antecedent of a rule used in the attacked argument. Second, two arguments can have opposite conclusions. Information about rule priorities (expressed using the rules names) is then used to compare the arguments. Argument evaluation is defined in terms of winning strategies in dialogue games: an argument is called justified when it can be successfully defended against an opponent's counterarguments.

Hage's approach [6], in several ways similar to Prakken & Sartor's, is more ambitious and philosophically radical.[8] For Hage, rules are first-and-foremost to be thought of as things with properties. As a result, a rule is formalized as a structured term. A rule's properties are then formalized using predicates. For instance, the fact that the rule that thieves are punishable, is valid is formalized as

Valid(rule(theftl, thief(x), punishable(x))).

Here 'theft1' is the name of the rule, 'thief(x)' the rule's antecedent and 'punishable(x)' its consequent. Hage's work takes the possibilities of a knowledge representation approach to the modelling of legal reasoning to its limits. For instance, there are dedicated predicates to express reasons, rule validity, rule applicability and the weighing of reasons.

How does the work by Prakken, Sartor & Hage relate to Toulmin's views? First, they have provided an operationalisation of Toulmin's idea of law-inspired logic, by formalizing aspects of legal reasoning. Second, they have refined Toulmin's treatment of argument. Notably, Prakken & Sartor have modelled specific kinds of rebuttal, namely by the attack of weakly negated assumptions and on the basis of

[7] Some other important AI & Law work concerning argumentation is for instance [1, 2, 5, 10].

[8] Hage's philosophical and formal theory of rules and reasons Reason-Based Logic was initiated by Hage and further developed in cooperation with Verheij.

rule priorities, and embedded them in an argumentative dialogue. Hage has added a further kind of rebuttal, namely by the weighing of reasons. Also, Hage has distinguished the validity of a rule from its applicability. The former can be regarded as an expression of a warrant in Toulmin's sense, and since in Hage's system rule validity can depend on other information, it is natural to model Toulmin's backings as reasons for the validity of a rule. And perhaps most importantly: Prakken & Sartor and Hage (and other AI & law researchers) have worked on the embedding of defeasible argumentation in a genuine procedural, dialogical setting (see also section 2.5). A further refinement of Toulmin's view is given by Verheij and colleagues [28], who show how two kinds of warrants (viz. legal rules and legal principles) with apparent logical differences, can be seen as extremes of a spectrum.

2.4 Dung's admissible sets

Dung's paper [3] has supplied an abstract mathematical foundation for formal work on argumentation. Following earlier mathematically flavoured work (e.g., [17, 29, 30]), his abstraction of only looking at the attack relation between arguments has helped organize the field, e.g., by showing how several formal systems of nonmonotonic reasoning can be viewed from the perspective of argument attack. A set of (unstructured) arguments with an attack relation is called an argumentation framework.

Dung has studied the mathematics of three types of subsets of the set of arguments of an argumentation framework: stable, preferred and grounded extensions. A set of arguments is a *stable extension* if it attacks all arguments not in the set. A set of arguments is a *preferred extension* if it is a maximal set of arguments without internal conflicts and attacking all arguments attacking the set. The *grounded extension* (there is only one) is the result of an inductive process: starting from the empty set, consecutively arguments are added that are only attacked by arguments already defended against.

Stable extensions can be regarded as an 'ideal' interpretation of an argumentation framework. When an extension is stable, all conflicts between arguments can be regarded as solved. It turns out that sometimes there are distinct ways of resolving the conflicts (e.g., when two arguments attack each other, each argument by itself is a stable extension) and that sometimes there is no way (e.g., when an argument is self-attacking). Preferred extensions are a generalization of stable extensions, as all stable extensions are also preferred. However, an argumentation framework always has a preferred extension (perhaps several). Preferred extensions can be regarded as showing how as many conflicts as possible can be resolved by counterattack. The grounded extension exists always and is a subset of all preferred and stable extensions.

Dung's work shows that the mathematics of argument attack is non-trivial and interesting. Thereby he has significantly extended our understanding of Toulmin's concept of rebuttal.

2.5 Walton's argumentation schemes

Toulmin's proposal that the maxims provided by a standard formal logical system (such as first order predicate logic) are not the only criteria for good reasoning and argument assessment, posed a new problem: if there are other, more field- and context-dependent standards of reasoning, what are they? For, though many recognized the shortcomings of formal logic for practical argument assessment, few were happy with the possible relativistic implication that anything goes. A good way to avoid the trap of uncontrolled relativism is to provide a systematic specification of standards of good reasoning.

One approach in this direction, which is especially close to Toulmin's conception of warrants, can be found in Walton's work on argumentation schemes (e.g., [31]). Argumentation schemes can be thought of as a semi-formal generalization of the rules of inference found in formal logic. Argument from expert opinion is an example ([31], 65):

> *E* is an expert in domain *D*.
> *E* asserts that *A* is known to be true.
> *A* is within *D*.
> Therefore, *A* may (plausibly) be taken to be true.

As Walton's argumentation schemes are context-dependent, not universal; defeasible, not strict; and concrete, not abstract, there is a strong analogy with Toulmin's warrants. There are two important differences though. First, Walton's argumentation schemes are structured, whereas Toulmin's warrants are not. Walton's argumentation schemes have premises, consisting of one or more sentences (often with informal variables), and a conclusion;[9] Toulmin's warrants are expressed as rule-like statements, such as 'A man born in Bermuda will generally be a British subject'.[10] By giving generic inference licenses more structure, as in Walton's work, the question arises whether they become formal enough to give rise to a kind of 'concrete logic'. Verheij [24] argues that it is a matter of choice, perhaps: taste, whether one draws the border between form and content on either side of argumentation schemes. To indicate the somewhat ambiguous status of argumentation schemes, the term 'semi-formal' may be most appropriate.

Second, Walton's argumentation schemes have associated critical questions. Critical questions help evaluating applications of an argumentation scheme. As a result, they play an important role in the evaluation of practical arguments. For instance, Walton lists the following critical questions for the scheme 'Argument from expert opinion' ([31], 65):

1. Is *E* a genuine expert in *D*?
2. Did *E* really assert *A*?

[9] Sometimes Walton's schemes take another form, e.g., small chains of argument steps or small dialogues; see Verheij's [24] for a format for the systematic specification of argumentation schemes inspired by knowledge engineering technology.

[10] Occasionally, a bit more structure is made explicit. For instance, when Toulmin phrases a warrant in an 'if ... then ...' form (e.g., 'If anything is red, it will not also be black', 98/91), thereby making an antecedent and consequent recognizable.

3. Is *A* relevant to domain *D*?
4. Is *A* consistent with what other experts in *D* say?
5. Is *A* consistent with known evidence in *D*?

Critical questions are related to argument attack, as they point to circumstances in which application of the scheme is problematic (e.g., [24]). For instance, the question 'Is *E* a genuine expert in *D*?' questions whether the element '*E* is an expert in domain *D*' in the premises of the scheme really holds. Some critical questions are like Toulmin's notion of rebuttal. For instance, the question 'Is *A* consistent with what other experts in *D* say?' points to a rebuttal '*A* is not consistent with what other experts in *D* say', which, if accepted, can raise doubt whether the conclusion can justifiably be drawn. In general, four types of critical questions can be distinguished [24]:

1. Critical questions concerning the *conclusion* of an argumentation scheme. Are there other reasons, based on other argumentation schemes for or against the scheme's conclusion?
2. Critical questions concerning the elements of the *premises* of an argumentation scheme. Is *E* an expert in domain *D*? Did *E* assert that *A* is known to be true? Is *A* within *D*?
3. Critical questions based on the *exceptions* of an argumentation scheme. Is *A* consistent with what other experts in *D* say? Is *A* consistent with known evidence in *D*?
4. Critical questions based on the *conditions of use* of an argumentation scheme. Do experts with respect to facts like *A* provide reliable information concerning the truth of *A*?

The critical questions associated with an argumentation scheme point to the dialogical setting of argumentation. Toulmin mentions the dialogical and procedural setting of argumentation (as, e.g., when discussing the jurisprudence metaphor for logic), but the discussion is not elaborate. Much work on the relation between argumentation and dialogue has been done. There is for instance the pragma-dialectical school (e.g., [21]), but also Walton's conception of argumentation is embedded in a procedural, dialogical setting. For instance, Walton [32] expresses a view on how to determine the relevance of an argument in a dialogue. There are six issues to take into account: the dialogue type[11], the stage the dialogue is in, the dialogue's goal, the type of argument, which is determined by the argumentation scheme underlying the argument, the prior sequence of argumentation, and the institutional and social setting.

In conclusion, Walton's work has played a significant role in two developments in AI with respect to Toulmin's main themes. First, the study of argumentation schemes by Walton and others has made a start with the systematic specification of context-dependent, defeasible, concrete standards of argument assessment, as sought for by Toulmin. And, second, the idea of considering argumentation from a procedural, dialogue perspective has been elaborated upon.

2.6 Reed & Rowe's argument analysis software

Further steps towards the realization of Toulmin's goals have been made by the recent advent of software-support of argumentative tasks, often using argument diagrams [8, 26]. In this connection, Reed & Rowe's work on the Araucaria tool [14] is

[11] See also Walton & Krabbe's [33], a treatment of dialogue types that is especially influential in research in AI and multi-agent systems.

especially relevant for the achievement of Toulmin's goals, as they have presented Araucaria specifically as a software tool for argument analysis. Araucaria uses an argument diagramming format, in which the recursive tree-structure of reasons supporting conclusions is depicted. It is also possible to indicate statements that are in conflict. Araucaria's standard diagramming format[12] is different from Toulmin's in several ways, but especially by not graphically distinguishing warrants from data. In an interestingly different way, however, Araucaria's standard format does include the idea of context-dependent types of reasoning as argued for by Toulmin, namely by its incorporation of Walton-style argumentation schemes (cf. section 2.5). Argumentation schemes can be used in Araucaria to label argumentative steps. For instance, a concrete argument 'There is smoke. Therefore, there is fire' could be labelled as an instance of the scheme 'Argument from sign', thereby giving access two critical questions, such as 'Are there other events that would more reliably account for the sign?'. By this possibility, Reed & Rowe's Araucaria is a useful step towards software-supported argument assessment. The tool provides a significant extension of Toulmin's aim to change logic from an 'idealised logic' to a 'working logic'.

2.7 Verheij's formal reconstruction of Toulmin's scheme

Already the examples in this chapter show a wide variety of approaches to — what might be called — semi-formal defeasible argumentation; and this is just the tip of the iceberg. By this embarrassment of riches, the question arises whether there are fundamental differences, e.g., between explicitly Toulmin-oriented approaches and other; or is the similarity of subject matter strong enough to allow for a synthesis of approaches? Looking for answers, I have attempted to reconstruct Toulmin's scheme using modern formal tools [25]. I used the abstract argumentation logic DefLog [23]. DefLog uses two connectives $\times$ and $\rightsquigarrow$: the first for expressing the defeat of a prima facie justified statement ('negation-as-defeat', the semantics of which falls outside the scope of this chapter), the second for expressing a conditional relation between statements ('primitive implication', validating Modus Ponens, but lacking a so-called introduction rule).[13] Toulmin's notion of a qualifier has been left out of the reconstruction.

The key to the translation of Toulmin's scheme into DefLog is to explicitly express that a datum leads to a claim; in DefLog: $D \rightsquigarrow C$. DefLog's primitive implication can be thought of as expressing a specific inference license. It is an explicit expression of what Toulmin refers to as a 'logical gulf' (9/9) that seems to exist between a reason and the state of affairs it supports. In this way, the licensing of concrete argument steps is removed from the logic, i.e., the fixed formalized background specifying general argument validity, and shifted to the contingent information. In this way, it becomes possible to express substantial arguments about concrete inferential bridges.

---

[12] In later versions, two alternative formats are provided: Toulmin's and Wigmore's.

[13] DefLog is formally an extension of Dung's abstract argumentation framework (section 2.4), as Dung's attack between two arguments A and B can be expressed as $A \rightsquigarrow \times B$. DefLog analogues of Dung's stable and preferred semantics are defined and proven to coincide with Dung's when DefLog's language is restricted to Dung's.

In particular, the role of a warrant can now be expressed as a reason for such a conditional statement: $W \rightsquigarrow (D \rightsquigarrow C)$. To formally show that a datum and claim are specific, whereas a warrant is to be thought of as a generic inference license, we can use variables and their instances: $W \rightsquigarrow (D(\mathrm{t}) \rightsquigarrow C(\mathrm{t}))$. This clarifies the distinctions between the following three:

(1) A man born in Bermuda will generally be a British subject.
(2) If *Person* was born in Bermuda, then generally *Person* is a British subject.
(3) If Harry was born in Bermuda, then generally he is a British subject.

The first is the ordinary language expression of a warrant as a rule-like statement (formally: W). The third is a conditional sentence (formally: $D(\mathrm{t}) \rightsquigarrow C(\mathrm{t})$), instantiating the second, which is a scheme of conditional sentences (formally: $D(x) \rightsquigarrow C(x)$, where x is a variable, that can be instantiated by the concrete term t). Note that only (1) and (3) can occur in actual texts, whereas (2) is — by its use of a variable *Person* — an abstraction. One could say however that (1) and (2) imply each other: a warrant corresponds to a scheme of argumentative steps from datum to claim. Given this analysis of warrants and their relation to datum and claim, backings are simply reasons for warrants: $B \rightsquigarrow W$.

Rebuttals are an ambiguous concept in Toulmin's treatment. He associates rebuttals with 'circumstances in which the general authority of the warrant would have to be set aside' (101/94), 'exceptional circumstances which might be capable of defeating or rebutting the warranted conclusion' (101/94) and with the (non)applicability of a warrant (102/95). It turns out that these three can be distinguished, and, given the present analysis of the warrant-datum-claim part of Toulmin's scheme, even extended to five kinds of rebuttals, as there are five different statements that can be argued against: the datum D, the claim C, the warrant W, the conditional $D(\mathrm{t}) \rightsquigarrow C(\mathrm{t})$, expressing the inferential bridge from datum to claim, and the conditional $W \rightsquigarrow (D(\mathrm{t}) \rightsquigarrow C(\mathrm{t}))$, which expresses the application of the warrant in the concrete situation (see [25] for a more extensive explanation). A rebuttal of the latter conditional coincides conceptually with an undercutting defeater in the sense of Pollock's. Note that the analysis suggests an answer to Pollock's open issue of how to analyze the conditional 'P wouldn't be true unless Q were true' (section 2.2). If U undercuts P as a reason for Q, we would write $\mathrm{U} \rightsquigarrow \times(\mathrm{P} \rightsquigarrow \mathrm{Q})$.

In this analysis, there is a natural extension of Toulmin's concept of warrant: just as it is necessary to specify which data imply claims (by Toulmin's warrants), it is necessary to specify which rebuttals block the application of warrants. Informally: it is a matter of substance, not logic, which statements are rebuttals. It must be *shown* by argument whether some statement is a rebuttal. In the law, for instance, not only legal rules (a kind of warrants) find backing in statutes, but also exceptions to rules. In the present analysis, dealing with 'rebuttal warrants' is a matter of course, since there is an explicit expression that R is a rebuttal.[14]

A side effect of the reconstruction is that arguments modelled according to Toulmin's scheme can be formally evaluated. For instance, assuming that datum and

[14] When R is a rebuttal in the sense of a Pollockian undercutter, this requires a sentence of the form $\mathrm{R} \rightsquigarrow \times(W \rightsquigarrow (D \rightsquigarrow C))$, expressing that, if R holds, the warrant W is not applicable.

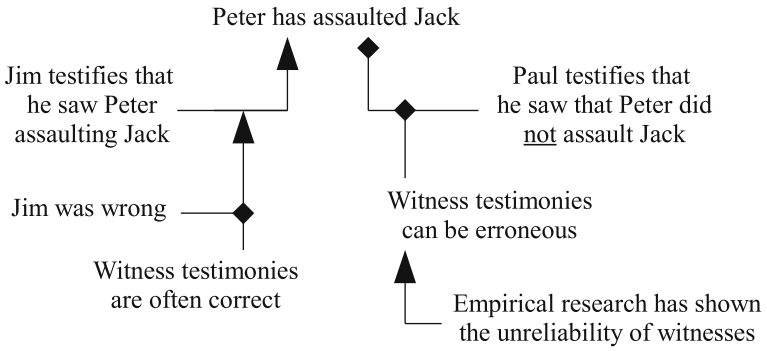

Fig. 11.2 An entangled dialectical argument

warrant hold, but not a rebuttal, the claim follows; when also a rebuttal is assumed, the claim does not follow.[15] A rebuttal of a rebuttal can be shown to reinstate a claim. Verheij [24] extends the approach to include Walton's argumentation schemes.

The result of the formal reconstruction of Toulmin's scheme showed some extensions, while retaining the original flavour. Figure 2 (using a diagramming format used in [27]) illustrates the basic relations between statements as distinguished here: Claims can have reasons for and against them (Jim's testimony supporting the assault by Jack, and Paul's attacking it), and the inferential bridges (the conditionals connecting reasons with their conclusions, here drawn as arrows) can be argued about just like other statements. The resulting argument structures are *dialectical*, by their incorporation of pros and cons, and *entangled* by their allowing the support and attack of inferential bridges.

3 Concluding remarks

It has been shown that central points of view argued for by Toulmin (1958), in particular the defeasibility of argumentation, the substantial, instead of formal, nature of standards of argument assessment, and the richer set of building blocks for argument analysis, are very much alive. Also Toulmin's 'research program' of treating logic as generalised jurisprudence has been taken up (with or without reference to him) and proven to be fertile.

There have also been refinements and extensions. It is now known that defeasible argumentation has interesting (and intricate) formal properties. There exist formal systems for the evaluation of defeasible arguments. Toulmin's argument diagram and its associated set of building blocks for argument analysis have been made precise and become refined, e.g., by distinguishing between kinds of argument attack. Not only is there now a wealth of studies of domain-bound, concrete forms of ar-

[15] Formally: using $W \rightsquigarrow (D \rightsquigarrow C)$, $\mathrm{R} \rightsquigarrow \times(W \rightsquigarrow (D \rightsquigarrow C))$, two sentences expressing that W is a warrant and R a rebuttal blocking its application, respectively, as background, and then assuming W, and D, one finds a unique dialectical interpretation, in which C holds, whereas adding R to the assumptions leads to a unique dialectical interpretation in which C does not hold.

gumentation, also methods for their systematic investigation have been proposed. Defeasible argumentation has been embedded in procedural models of dialogue. Toulmin's wish to develop logic into a practical tool has found a modern guise in the form of argumentation-support software, aiming at argument analysis and production.

Notwithstanding recent progress, there is ample room for innovative research. Some possible directions of future research are the continuing systematisation and specification of argumentation schemes; the further organisation of the wealth of evaluation paradigms for defeasible argumentation ('semantics'); the prolongation of research aiming at practically useful software tools, especially when supported by user studies or commercial success; the implementation of software agents capable of argumentative behaviour; and the coupling of empirical work on reasoning and argumentation to the findings in AI.

Toulmin ends his introduction in a modest, but hopeful mood:

> The studies which follow are, as I have said, only essays. If our analysis of arguments is to be really effective and true-to-life it will need, very likely, to make use of notions and distinctions that are not even hinted at here. But of one thing I am confident: that by treating logic as generalised jurisprudence and testing our ideas against our actual practice of argument-assessment, rather than against a philosopher's ideal, we shall eventually build up a picture very different from the traditional one. The most I can hope for is that some of the pieces whose shape I have here outlined will keep a place in the finished mosaic. (10/10)

As evidenced by the research discussed in this chapter, the present state of the art in AI-inspired argumentation research shows that Toulmin's hope has been fulfilled.

Acknowledgements The author would like to thank David Hitchcock, Douglas Walton and James Freeman for comments on a prepublication version of this text.

References

1. K. D. Ashley. *Modeling legal argument. Reasoning with cases and hypotheticals*. The MIT Press, Cambridge (Massachusetts), 1990.
2. T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.
3. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.
4. E. T. Feteris. *Fundamentals of legal argumentation. A Survey of Theories on the Justification of Judicial Decisions*. Kluwer Academic Publishers, Dordrecht, 1999.
5. T. F. Gordon. *The Pleadings Game. An Artificial Intelligence Model of Procedural Justice*. Kluwer Academic Publishers, Dordrecht, 1995.
6. J. C. Hage. A theory of legal reasoning and a logic to match. *Artificial Intelligence and Law*, 4:199–273, 1996.
7. D. L. Hitchcock and B. Verheij, editors. *Arguing on the Toulmin Model. New Essays in Argument Analysis and Evaluation (Argumentation Library, Volume 10).* Springer-Verlag, Dordrecht, 2006.
8. P. A. Kirschner, S. J. Buckingham Shum, and C. S. Carr. *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. Springer-Verlag, London, 2002.

9. R. P. Loui. Hart's critics on defeasible concepts and ascriptivism. In *The Fifth International Conference on Artificial Intelligence and Law. Proceedings of the Conference*, pages 21–30. ACM, New York (New York), 1995.
10. R. P. Loui and J. Norman. Rationales and argument moves. *Artificial Intelligence and Law*, 3:159–189, 1995.
11. J. L. Pollock. Defeasible reasoning. *Cognitive Science*, 11:481–518, 1987.
12. J. L. Pollock. *Cognitive Carpentry: A Blueprint for How to Build a Person*. The MIT Press, Cambridge (Massachusetts), 1995.
13. H. Prakken and G. Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 1996.
14. C. Reed and G. Rowe. Araucaria: Software for argument analysis, diagramming and representation. *International Journal of AI Tools*, 13(4):961–980, 2004.
15. R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.
16. S. J. Russell and P. Norvig. *Artificial Intelligence. A Modern Approach*. Prentice Hall, Upper Saddle River (New Jersey), 2003.
17. G. R. Simari and R. P. Loui. A mathematical treatment of defeasible reasoning and its applications. *Artificial Intelligence*, 53:125–157, 1992.
18. S. E. Toulmin. *The Uses of Argument*. Cambridge University Press, Cambridge, 1958.
19. S. E. Toulmin. *The Uses of Argument. Updated Edition*. Cambridge University Press, Cambridge, 2003.
20. S. E. Toulmin. Reasoning in theory and practice. In D. L. Hitchcock and B. Verheij, editors, *Arguing on the Toulmin Model. New Essays in Argument Analysis and Evaluation (Argumentation Library, Volume 10).*, pages 25–30. Springer-Verlag, Dordrecht, 2006.
21. F. H. van Eemeren, R. Grootendorst, S. Jackson, and S. Jacobs. *Reconstructing Argumentative Dialogue*. The University of Alabama Press, Tuscaloosa (Alabama), 1993.
22. F. H. van Eemeren, R. Grootendorst, and F. Snoeck Henkemans. *Fundamentals of Argumentation Theory. A Handbook of Historical Backgrounds and Contemporary Developments*. Lawrence Erlbaum Associates, Mahwah (New Jersey), 1996.
23. B. Verheij. DefLog: on the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation*, 13(3):319–346, 2003.
24. B. Verheij. Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial Intelligence and Law*, 11(1-2):167–195, 2003.
25. B. Verheij. Evaluating arguments based on toulmin's scheme. *Argumentation*, 19(3):347–371, 2005.
26. B. Verheij. *Virtual arguments. On the design of argument assistants for lawyers and other arguers.* TMC Asser Press, The Hague, 2005.
27. B. Verheij. Argumentation support software: Boxes-and-arrows and beyond. *Law, Probability and Risk*, 6:187–208, 2007.
28. B. Verheij, J. C. Hage, and H. J. van den Herik. An integrated view on rules and principles. *Artificial Intelligence and Law*, 6(1):3–26, 1998.
29. G. A. W. Vreeswijk. *Studies in defeasible argumentation*. 1993.
30. G. A. W. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence*, 90:225–279, 1997.
31. D. N. Walton. *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah (New Jersey), 1996.
32. D. N. Walton. *The New Dialectic: Conversational Contexts of Argument*. University of Toronto Press, Toronto, 1998.
33. D. N. Walton and E. Krabbe. *Commitment in Dialogue. Basic Concepts of Interpersonal Reasoning*. State University of New York Press, Albany (New York), 1995.