

## LOGICS FOR DEFEASIBLE ARGUMENTATION

## 1 INTRODUCTION

Logic is the science that deals with the formal principles and criteria of validity of patterns of inference. This chapter surveys logics for a particular group of patterns of inference, namely those where arguments for and against a certain claim are produced and evaluated, to test the tenability of the claim. Such reasoning processes are usually analysed under the common term ‘defeasible argumentation’. We shall illustrate this form of reasoning with a dispute between two persons, *A* and *B*. They disagree on whether it is morally acceptable for a newspaper to publish a certain piece of information concerning a politician’s private life.<sup>1</sup> Let us assume that the two parties have reached agreement on the following points.

- (1) The piece of information *I* concerns the health of person *P*;
- (2) *P* does not agree with publication of *I*;
- (3) Information concerning a person’s health is information concerning that person’s private life

*A* now states the moral principle that

- (4) Information concerning a person’s private life may not be published if that person does not agree with publication.

and *A* says “So the newspapers may not publish *I*” (Fig. 1, page 220). Although *B* accepts principle (4) and is therefore now committed to (1-4), *B* still refuses to accept the conclusion that the newspapers may not publish *I*. *B* motivates his refusal by replying that:

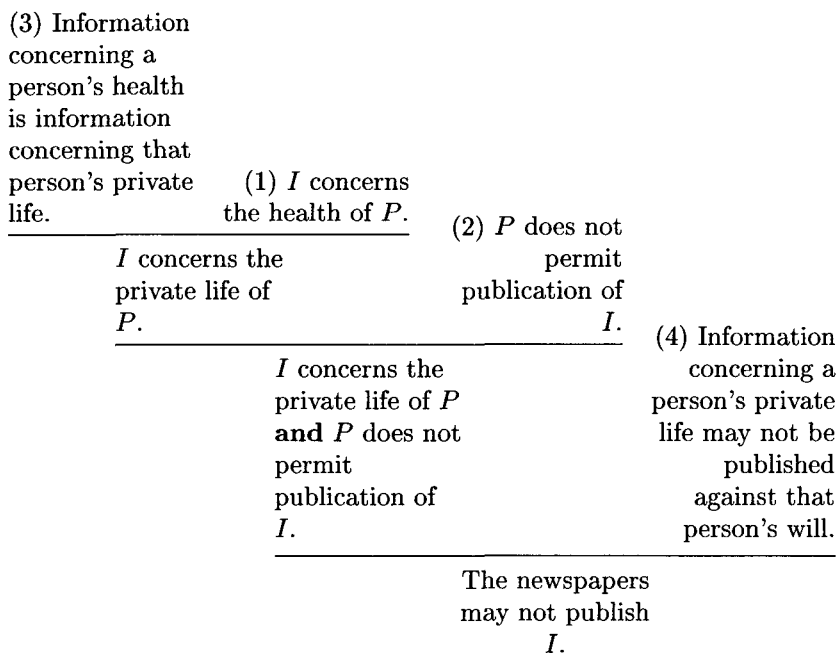
- (5) *P* is a cabinet minister
- (6) *I* is about a disease that might affect *P*’s political functioning
- (7) Information about things that might affect a cabinet minister’s political functioning has public significance

Furthermore, *B* maintains that there is also the moral principle that

- (8) Newspapers may publish any information that has public significance

---

<sup>1</sup>Adapted from [Sartor, 1994].

Figure 1. *A*'s argument.

*B* concludes by saying that therefore the newspapers may write about *P*'s disease (Fig. 2, page 221). *A* agrees with (5–7) and even accepts (8) as a moral principle, but *A* does not give up his initial claim. (It is assumed that *A* and *B* are both male.) Instead he tries to defend it by arguing that he has the stronger argument: he does so by arguing that in this case

(9) The likelihood that the disease mentioned in *I* affects *P*'s functioning is small.

(10) If the likelihood that the disease mentioned in *I* affects *P*'s functioning is small, then principle (4) has priority over principle (8).

Thus it can be derived that the principle used in *A*'s first argument is stronger than the principle used by *B* (Fig. 3, page 222), which makes *A*'s first argument stronger than *B*'s, so that it follows after all that the newspapers should be silent about *P*'s disease.

Let us examine the various stages of this dispute in some detail. Intuitively, it seems obvious that the accepted basis for discussion after *A* has stated (4) and *B* has accepted it, viz. (1,2,3,4), warrants the conclusion that the piece of information *I* may not be published. However, after *B*'s counterargument and *A*'s acceptance of its premises (5–8) things have changed.

(5) <i>P</i> is a cabinet minister.	(6) <i>I</i> is about a disease that might affect <i>P</i> 's political functioning.	(7) Information about things that might affect a cabinet minister's political functioning has public significance.	(8) Newspapers may publish any information that has public significance.
<hr/>		<i>I</i> has public significance.	<hr/>
		The newspapers may publish <i>I</i> .	

Figure 2. *B*'s argument.

At this stage the joint basis for discussion is (1-8), which gives rise to two conflicting arguments. Moreover, (1-8) does not yield reasons to prefer one argument over the other: so at this point *A*'s conclusion has ceased to be warranted. But then *A*'s second argument, which states a preference between the two conflicting moral principles, tips the balance in favour of his first argument: so after the basis for discussion has been extended to (1-10), we must again accept *A*'s moral claim as warranted.

This chapter is about logical systems that formalise this kind of reasoning. We shall call them 'logics for defeasible argumentation', or 'argumentation systems'. As the example shows, these systems lack one feature of 'standard', deductive logic (say, first-order predicate logic, FOL). The notion of 'warrant' that we used in explaining the example is clearly not the same as first-order logical consequence, which has the property of monotonicity: in FOL any conclusion that can be drawn from a given set of premises, remains valid if we add new premises to this set. So according to FOL, if *A*'s claim is implied by (1-4), it is surely also implied by (1-8). From the point of view of FOL it is pointless for *B* to accept (1-4) and yet state a counterargument; *B* should also have refused to accept one of the premises, for instance, (4).

- |  |  |
|--|--|
| (9) The likelihood that<br>the disease mentioned in<br><i>I</i> affects <i>P</i> 's functioning<br>is small. | (10) If the likelihood that the disease<br>mentioned in <i>I</i> affects <i>P</i> 's functioning<br>is small, then principle (4) has priority<br>over principle (8). |
|--|--|
- 

Principle (4) has priority  
over principle (8).

Figure 3. *A*'s priority argument.

Does this mean that our informal account of the example is misleading, that it conceals a subtle change in the interpretation of, say, (4) as the dispute progresses? This is not so easy to answer in general. Although in some cases it might indeed be best to analyse an argument move like *B*'s as a reinterpretation of a premise, in other cases this is different. In actual reasoning, rules are not always neatly labelled with an exhaustive list of possible exceptions; rather, people are often forced to apply 'rules of thumb' or 'default rules', in the absence of evidence to the contrary, and it seems natural to analyse an argument like *B*'s as an attempt to provide such evidence to the contrary. When the example is thus analysed, the force of the conclusions drawn in it can only be captured by a consequence notion that is nonmonotonic: although *A*'s claim is warranted on the basis of (1–4), it is not warranted on the basis of (1–8).

Such nonmonotonic consequence notions have been studied over the last twenty years in an area of artificial intelligence called 'nonmonotonic reasoning' (recently the term 'defeasible reasoning' has also become popular), and logics for defeasible argumentation are largely a result of this development. Some might say that the lack of the property of monotonicity disqualifies these notions from being notions of logical consequence: isn't the very idea of calling an inference 'logical' that it is (given the premises) beyond any doubt? We are not so sure. Our view on logic is that it studies criteria of warrant, that is, criteria that determine the degree according to which it is reasonable to accept logical conclusions, even though some of these conclusions are established non-deductively: sometimes it is reasonable to accept a conclusion of an argument even though this argument is not strong enough to establish its conclusion with absolute certainty.

Several ways to formalise nonmonotonic, or defeasible reasoning have been studied. This chapter is not meant to survey all of them but only discusses the argument-based approach, which defines notions like argument, counterargument, attack and defeat, and defines consequence notions in terms of the interaction of arguments for and against certain conclusions. This approach was initiated by the philosopher John Pollock [1987], based

on his earlier work in epistemology, e.g. [1974], and the computer scientist Ronald Loui [1987]. As we shall see, argumentation systems are able to incorporate the traditional, monotonic notions of logical consequence as a special case, for instance, in their definition of what an argument is.

The field of defeasible argumentation is relatively young, and researchers disagree on many issues, while the formal meta-theory is still in its early stages. Yet we think that the field has sufficiently matured to devote a handbook survey to it.<sup>2</sup> We aim to show that there are also many similarities and connections between the various systems, and that many differences are variations on a few basic notions, or are caused by different focus or different levels of abstraction. Moreover, we shall show that some recent developments pave the way for a more elaborate meta-theory of defeasible argumentation.

Although when discussing individual systems we aim to be as formal as possible, when comparing them we shall mostly use conceptual or quasi-formal terms. We shall also report on some formal results on this comparison, but it is not our aim to present new technical results; this we regard as a task for further research in the field.

The structure of this chapter is as follows. In Section 2 we give an overview of the main approaches in nonmonotonic reasoning, and argue why the study of this kind of reasoning is relevant not only for artificial intelligence but also for philosophy. In Section 3 we give a brief conceptual sketch of logics for defeasible argumentation, and we argue that it is not obvious that they need a model-theoretic semantics. In Section 4 we become formal, studying how semantic consequence notions for argumentation systems can be defined given a set of arguments ordered by a defeat relation. This discussion is still abstract, leaving the structure of arguments and the origin of the defeat relation largely unspecified. In Section 5 we become more concrete, in discussing particular logics for defeasible argumentation. Then in Section 6 we discuss one way in which argumentation systems can be formulated, viz. in the form of rules for dispute. We end this chapter in Section 7 with some concluding remarks, and with a list of the main open issues in the field.

## 2 NONMONOTONIC LOGICS: OVERVIEW AND PHILOSOPHICAL RELEVANCE

Before discussing argumentation systems, we place them in the context of the study of nonmonotonic reasoning, and discuss why this study deserves a place in philosophical logic.

---

<sup>2</sup>For a survey of this topic from a computer science perspective, see [Chesñevar *et al.*, 1999].

Handbook of Philosophical Logic

Gabbay, D.; Guenther, F. (Eds.)

2002, XIII, 431 p. 1 illus., Hardcover

ISBN: 978-1-4020-0139-0