Logics of Argumentation and the Law[1]

Henry Prakken
Faculty of Law, University of Groningen, The Netherlands
&
Department of Information and Computing Sciences, Faculty of Science, Utrecht University, The Netherlands.

1 Introduction

In his *Legal Traditions of the World* (Glenn 2010), Patrick Glenn observes that the world contains many different legal traditions, often inconsistent with each other, and that even a single tradition can contain different sub-traditions that may be inconsistent with each other. Moreover, he notes that these traditions may interact with each other in complex ways. In chapter 10 Glenn raises the question of how to account for this from the perspective of formal logic. In chapter 14 of his *The Cosmopolitan State* (Glenn 2013) he writes that new logics may be needed that are multivalent, paraconsistent or non-monotonic and do not adhere to the classic rules of non-contradiction and the excluded middle. In this chapter I will explore the use of one such new kind of logic, namely, logics of argumentation. It will turn out that such logics offer what Patrick Glenn is asking for without giving up classical two-valued logic. Instead, classical logic is in argumentation logics embedded in a larger formal framework, and it is this larger framework that has the desired nonstandard behavior. Thus argumentation logics provide a way to cope with inconsistent legal traditions without having to give up two-valued logic, a way that is moreover arguably close to the way lawyers think since notions like argument, counterargument and rebuttal are natural them.

Introductory textbooks to logic often portray logically valid inference as 'foolproof' reasoning: an argument is deductively valid if the truth of its premises guarantees the truth of its conclusion. In other words, if one accepts all premises of a deductively valid argument, then one also has to accept its conclusion, no matter what. However, we all construct arguments from time to time that are not foolproof in this sense but that merely make their conclusion plausible when their premises are true. For example, if we are told that John and Mary are married and that John lives in Amsterdam, we conclude that Mary will live in Amsterdam as well, since we know that usually married people live where their spouses live. Sometimes such arguments are overturned by counterarguments. For example, if we are told that Mary lives in Rome to work at the foreign offices of her company for two years, we have to retract our previous conclusion that she lives in Amsterdam. However, as long as such counterarguments are not available, we are happy to live with the conclusions of our fallible arguments. The question is: are we then reasoning fallaciously or is there still logic in our reasoning?

[1] Parts of this chapter are adapted from Prakken & Sartor (2009).

The answer to this question has been given in more than thirty years of research in Artificial Intelligence (AI) on so-called logics for defeasible reasoning, partly inspired by earlier developments in philosophy and argumentation theory. At first sight it might be thought that patterns of defeasible reasoning are a matter of applying probability theory. However, many patterns of defeasible reasoning cannot be analysed in a probabilistic way. In the legal domain this is particularly clear: while reasoning about the facts can (at least in principle) still be regarded as probabilistic, reasoning about normative issues clearly is of a different nature. Moreover, even in matters of evidence reliable numbers are usually not available so that the reasoning has to be qualitative.

In this chapter an account is sketched of legal reasoning that respects that arguments can be fallible for various reasons. In short, the account is that reasoning consists of constructing arguments, attacking these arguments with counterarguments, and adjudicating between conflicting arguments on grounds that are appropriate to the conflict at hand. Just as in deductive reasoning, arguments must instantiate inference schemes (now called 'argument schemes') but only some of these schemes capture fool-proof reasoning: in our account deductive logic turns out to be the special case of argument schemes that can only be attacked on their premises.

This chapter is organised as follows. In Section 2 the notions of argument, counterargument and the relations of attack and defeat between conflicting arguments are introduced. These ingredients are in Section 3 combined into the idea of dialectical argument evaluation, which completes the general architecture of an argumentation logic. Then in Section 4 the distinction between deductive and defeasible arguments is introduced, and in Section 5 some stereotypical patterns of defeasible arguments are presented, together with stereotypical ways to attack them. Section 6 delves deeper into the formalisation of argumentation logics; this section is primarily meant for more formally interested readers with some background in formal logic. Section 7 then concludes this chapter.

This chapter is intended to be a tutorial on argumentation logics and their relevance for legal reasoning. For this reason I will be sparse with references. A more formal introduction to argumentation logics with references to the literature can be found in Prakken (2011). The use of argumentation logics for modelling legal reasoning is reviewed in Prakken & Sartor (2015).

2 Arguments and Counterarguments

As just said, we assume that any argument instantiates some argument scheme. (More generally, arguments chain instantiations of argument schemes into trees, since the conclusion of one argument can be a premise of another.) Argument schemes are inference rules: they have a set of premises and a conclusion. What are the 'valid' argument schemes of defeasible reasoning? Much can be said on this and we will do so later on in Sections 4 and 5, but at least the deductively valid inference schemes of standard logic will be among them. In this section we examine how deductive arguments can be the subject of attack.

Consider the following example. According to Section 3:32 of the Dutch civil code (natural) persons have the capacity to perform legal acts (this means, for instance, that they can engage in contracts or sell their property), unless the law provides otherwise. Suppose John argues that he has legal capacity since he is a person and the law does not provide otherwise. Then in standard propositional logic we can write this argument as follows:

Argument A:
Somebody is a person & $\neg$ The law provides otherwise $\rightarrow$ S/he has legal capacity
John is a person
$\neg$ The law provides otherwise
Therefore, John has legal capacity

(Here & stands for 'and', $\neg$ for 'it is not the case that' and $\rightarrow$ for 'if … then'.) This argument is deductively valid, since it instantiates the deductively valid argument scheme of *modus ponens*:

Modus Ponens Scheme:
$P \rightarrow Q$
P
Therefore, Q

(where P and Q can be any statement). This scheme is deductively valid: it is impossible to accept all its premises but still deny its conclusion, since the truth of its premises guarantees the truth of its conclusion.

Now does the deductive validity of argument A mean that we have to accept its conclusion? Of course not: any first lesson in logic includes the advice: if you don't like the conclusion of a deductive argument, then challenge its premises. According to Section 1:234 of the Dutch civil code minors have the capacity to perform legal acts if and only if they have consent from their legal representative. Now suppose John's father claims that John is in fact a minor and does not have such consent. Then the following deductive argument against the premise '$\neg$The law provides otherwise' can be constructed, in two steps. First application of 1:234 results in the conclusion that John does not have the capacity to perform legal acts:

Argument B:
Somebody is a minor $\rightarrow$ (S/he has consent $\leftrightarrow$ S/he has legal capacity)
John is a minor
$\neg$ John has consent
Therefore, $\neg$ John has legal capacity

(Here $\leftrightarrow$ stands for 'if and only if'). The double arrow expresses that when a person is a minor, then having consent is not only a sufficient but also a necessary condition for having the capacity to perform legal acts. So, since John is a minor but does not have such consent, he does not have legal capacity. This conclusion can then be used to attack the third premise of argument A:

Argument B (continued):
$\neg$ John has legal capacity
$\neg$ John has legal capacity $\rightarrow$ The law provides otherwise
Therefore, the law provides otherwise

Now we must choose whether to accept the premise '¬ The law provides otherwise' of argument A or whether to give it up and accept the conclusion of counterargument B. Clearly the phrase ''unless the law provides otherwise'' of section 3:32 of the Dutch civil code is meant to express that any place where the law expresses otherwise is an exception to section 3:32. Since argument B is based on such a statutory exception, we must therefore give up the premise of A and accept the counterargument. In this case we say that argument B not just *attacks* but also *defeats* argument A.

However, not all attacks are a matter of statutory exceptions. In our example, John might attack his father's argument B by saying that he does have consent of his legal representative since his mother consented and she is his legal representative. This gives rise to an argument attacking the third premise of argument B (before the continuation):

Argument C:
Somebody's mother consented $\rightarrow$ S/he has consent
John's mother consented
Therefore, John has consent

This time we have a genuine conflict, namely, between John's father's claim that John acted without consent of his legal representative (the third premise of argument B) and John's claim that he acted with consent of his legal representative (the conclusion of argument C). Now note that if one accepts all premises of argument C, then one must also accept its conclusion, since argument C instantiates the deductively valid scheme of modus ponens. And if one accepts argument C's conclusion, one must, of course, reject the third premise of argument B. In the latter case we say that argument C not only attacks but also defeats argument B. Let us assume that the latter is indeed the case.

In sum then, what we have so far is that all three arguments are deductively valid but that argument A is defeated by argument B on its third premise while argument B is in turn defeated by argument C on its third premise. This implies that it is rational to accept the conclusions of arguments A and C: even though A is defeated by B, it is *defended* by C, which defeats A's only defeater.

This leads to a very important insight. In order to determine what to believe or accept in the face of a body of conflicting arguments it does not suffice to make a choice between two arguments that directly conflict with each other. We must also look at how arguments can be defended by other arguments. In our example this is quite simple: it is intuitively obvious that C defends A so, since C is not attacked by any argument, both argument A and argument C (and their conclusions) are acceptable. However, we can easily imagine more complex examples where our intuitions fall short. For instance, another argument D could be constructed such that C and D defeat each other, then an argument E could be constructed that defeats D but is defeated by A, and so on: which arguments can now be accepted and which should be rejected? Here we cannot rely on intuitions but need a *calculus,* or an *argumentation logic*. Its input will be a collection of arguments plus an assessment of which arguments defeat each other, while its output will be an assessment of the *dialectical status* of these arguments in terms of three classes (three and not two since some conflicts cannot be resolved). Intuitively, the *justified* arguments are those that survive all conflicts with their attackers and so can be accepted, the *overruled* arguments are those that are attacked by a justified argument and so must be rejected; and the *defensible* arguments are those that are involved in conflicts that cannot be resolved. Furthermore, a statement is justified if it has a justified argument, it is overruled if all arguments for it are overruled, and it is defensible if it has a defensible argument but no justified arguments. In terms more familiar to lawyers, if a claim is justified, then a rational adjudicator is convinced that the claim is true, if it is overruled, such an adjudicator is convinced that the claim is false, while if it is defensible, s/he is neither convinced that it is true nor that it is false.

Before an argumentation logic can be presented, a subtlety concerning the defeat relation between arguments must be explained. Above we made a distinction between attack and defeat. An argument A *defeats* an argument B if A *attacks* B and is *not inferior* to B (according to the appropriate criteria for comparing arguments). This definition allows that two arguments defeat each other, namely, if neither argument is inferior or superior to the other. In such cases we say that the two arguments *weakly defeat* each other; otherwise (if one argument is superior to the other) we say that one argument *strictly defeats the other*. Suppose in our example that further evidence was given by both John and his father on whether John's mother consented. John presents a testimony by his brother that his mother consented, while his father presents a testimony by John's mother that she never consented:

Argument C (continued)
John's brother says that John's mother consented
What a witness says is usually true
Therefore, John's mother consented

Argument D
John's mother says that she never consented
What a witness says is usually true
Therefore, $\neg$ John's mother consented

(In Section 4 we return to the question whether these arguments can be reconstructed as deductively valid and how it can be that they attack each other on their conclusions instead of on their premises.) Suppose that the court cannot find a reason why one witness testimony is stronger than the other: then it must conclude that both conflicting arguments (weakly) defeat each other.

3 Logic of Argumentation

Let us now discuss how an argumentation logic looks like. Just as in deductive logic, there is no single universally accepted one and there is an ongoing debate in AI on what is a good argumentation logic. However, we need not go into the details of this debate, since it turns out that there is a simple and intuitive definition that suffices for most applications. The idea is to regard an attempt to prove an argument justified as a *debate* between a proponent and opponent of the argument.[2] The proponent starts with the argument that he wants to prove justified and then the turn shifts to the opponent, who must provide all its defeating counterarguments. It does not matter whether they weakly or strictly defeat their target, since the opponent's task is to interfere with the

[2] The proponent and opponent should not be seen as real human beings; they are a metaphor for the dialectical nature of the reasoning process, where both pros and cons are considered. Such a dialectical reasoning process can just as well take place in the mind of a single reasoner.

proponent's attempt to prove his argument justified. For each of these defeating arguments the proponent must then construct one strict defeater (it has to be a strict defeater since the proponent must prove his argument justified). This process is repeated as long as it takes: at each of her turns, the opponent constructs all mutual and all strict defeaters of the proponent's previous arguments, while at each of his turns, the proponent constructs a strict defeater for each of the opponent's previous arguments, and so on. The idea is that our initial argument is justified if proponent can eventually make opponent run out of moves in every of opponent's lines of attack.

This process can be visualised as follows (note that this figure does *not* model the above example but is a new, abstract example).

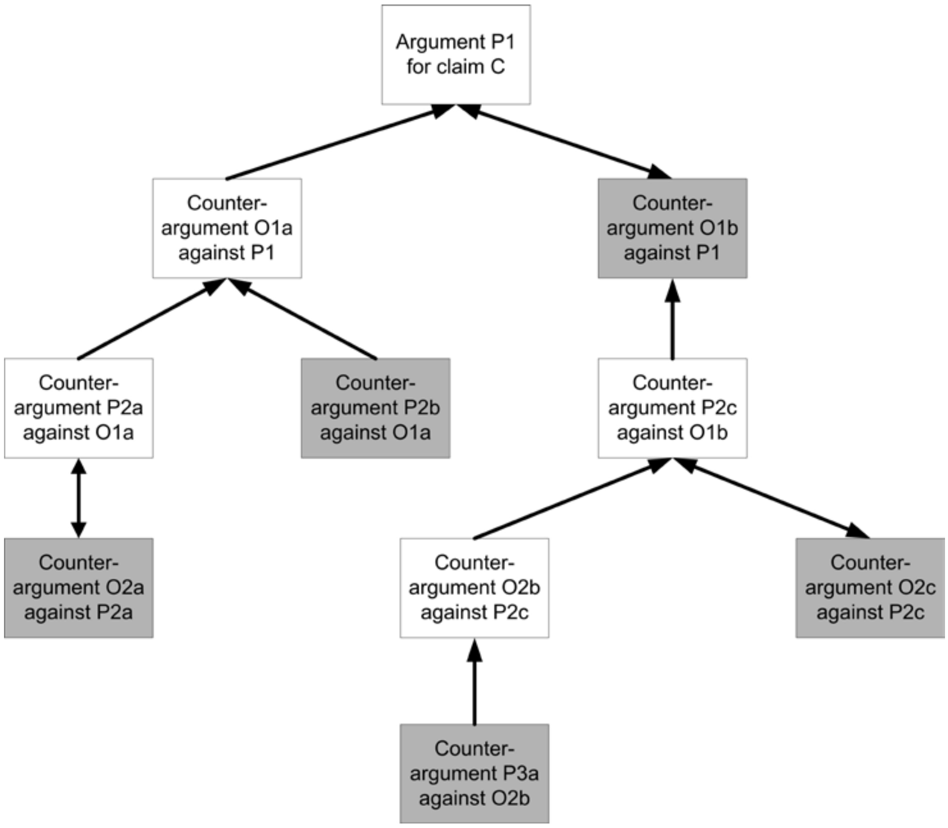

Figure 1: a dialectical tree

Note that if an argument is justified this does not mean that the proponent will in fact win the game: he could make the wrong choice at some point. All that it means is that the proponent will win if he plays optimally. In terms of game theory, an argument is justified if the proponent has a so-called *winning strategy* in a game that starts with the argument. In fact, there is a simple way to verify whether the proponent has a winning strategy. The idea is to label all arguments in the tree as *in* or *out* according to the following definition:

1. An argument is *in* if and only if all its defeating counterarguments are *out*
2. An argument is *out* if and only if it has a defeating counterargument that is *in*

In the figures *in* is coloured as grey and *out* as white. It is easy to see that because of (1) all leaves of the tree are trivially *in*, since they have no counterarguments. Then we can work our way upwards to determine the colour of all other arguments, ultimately arriving at a colour of the initial argument. If it is grey, i.e., *in*, then we know that the proponent has a winning strategy for it, namely by choosing a grey argument at each point he has to choose. If, on the other hand, the initial argument is white, i.e., *out*, then it is the opponent who has a winning strategy, which can be found in the same way. So in the above figure the opponent has a winning strategy, which she can follow by choosing argument O1b at her first turn.

Suppose now that new information becomes available that gives rise to a strictly defeating counterargument P3b against O2c. Then the situation is as follows:

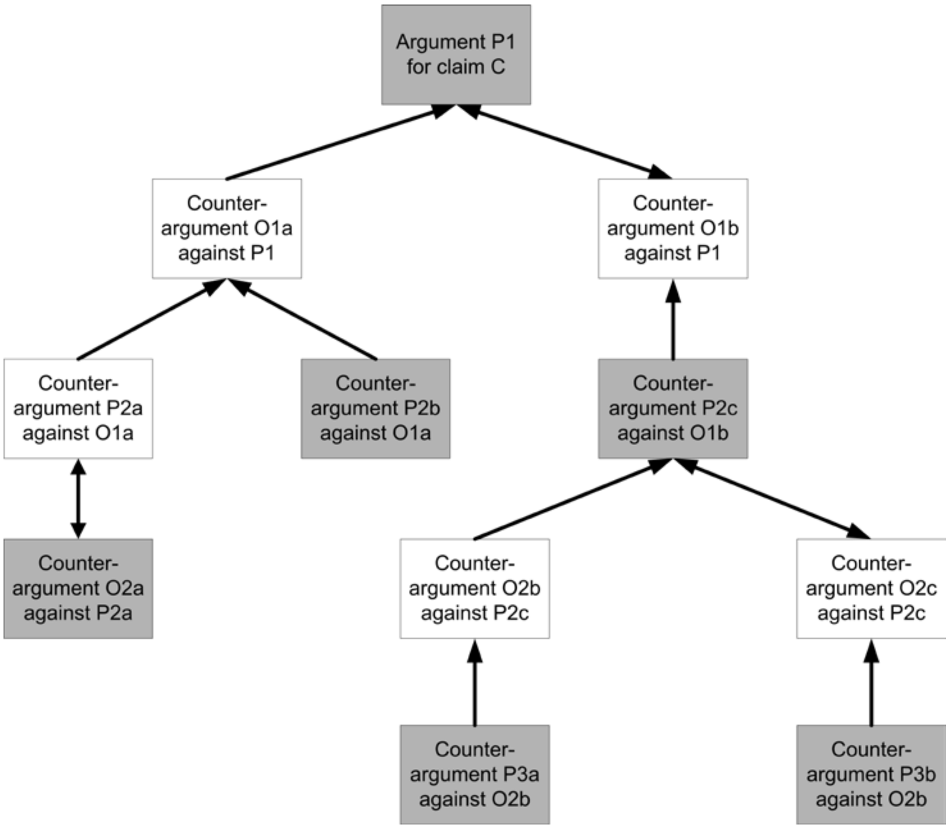

Figure 2: an extended dialectical tree.

Now argument P1 is *in* so now it is the proponent who has a winning strategy, viz. choosing P2b instead of P2a when confronted by O1a. This illustrates that when new information becomes available from which new arguments can be constructed, the dialectical status of arguments may change.

It should be noted that each argument appearing as a box in these trees has an internal structure. In the simplest case it just has a set of premises and a conclusion, but when the argument combines several inferences, it has the structure of an inference tree as is familiar from standard logic. This is illustrated by Figure 3, which displays all four arguments of our example, plus their defeat relations (solid lines represent inferences while dashed lines stand for defeat relations):

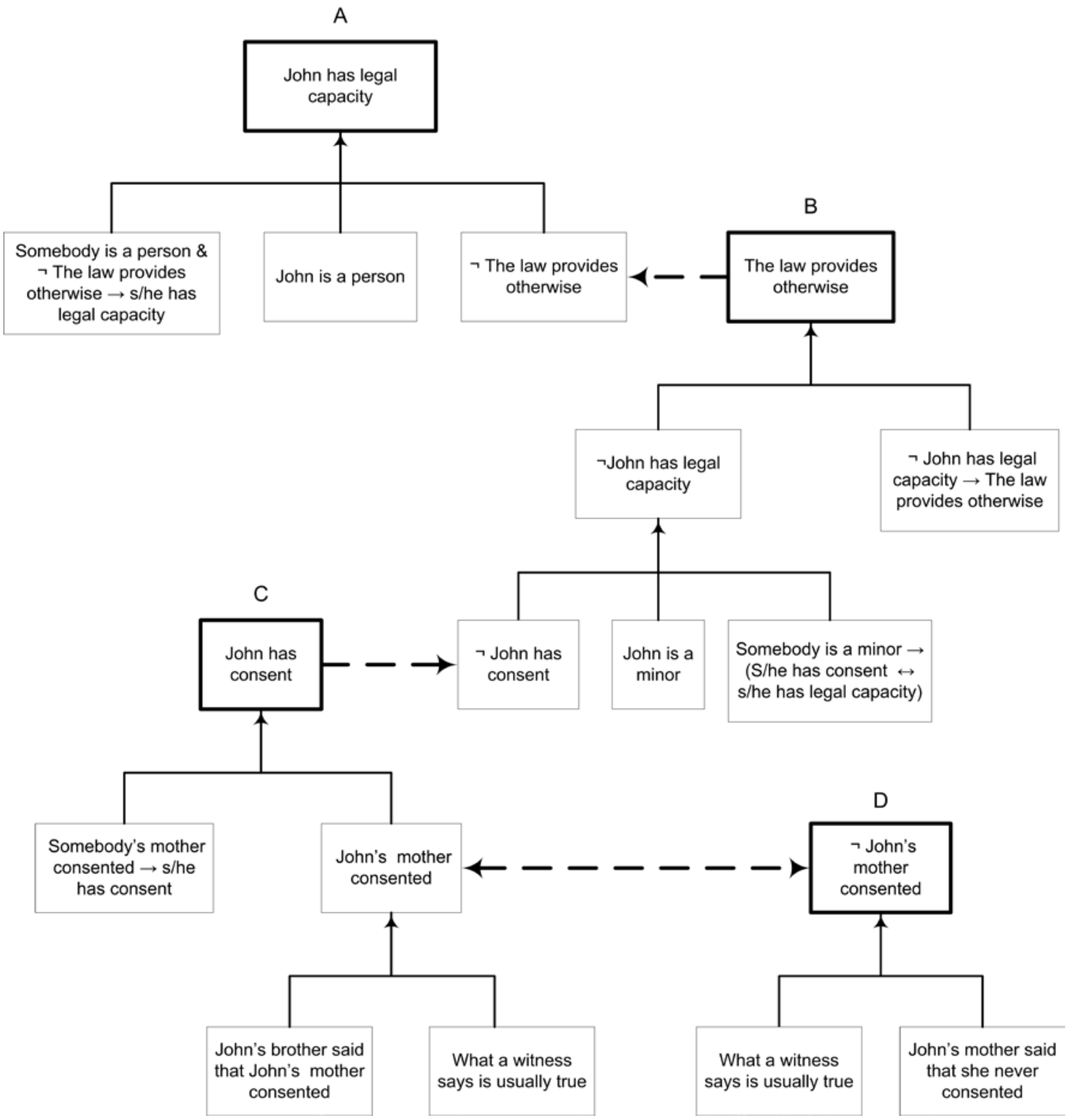

Figure 3: four arguments and their defeat relations

Abstracted to a dialectical tree as in Figure 1 we obtain Figure 4, which shows the two game trees for arguments A and B (both trees are very simple, having just one branch, since neither the proponent nor the opponent has any choice):

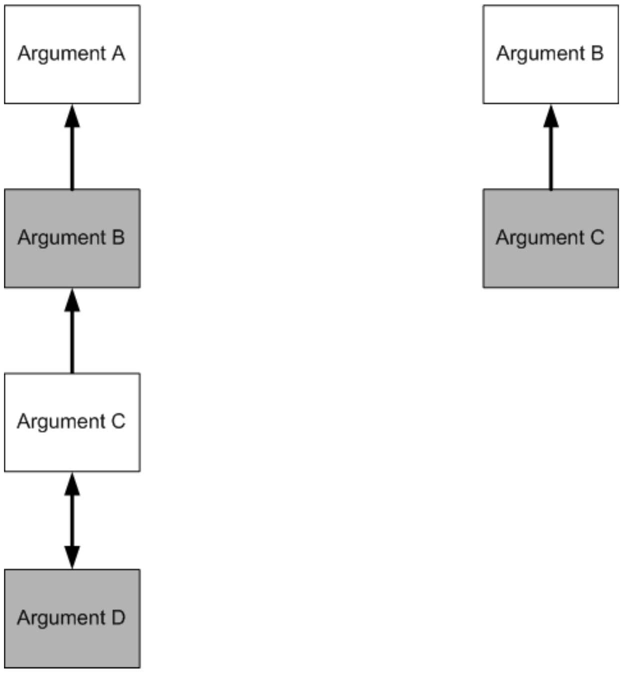

Figure 4: two abstract dialectical trees corresponding to Figure 3

The tree on the left of Figure 4 displays the game tree for argument A: it shows that argument A is not justified, since argument A is labelled *out*. The tree on the right displays the game tree for argument B: it shows that argument B is also not justified since it is also labelled *out*. Hence both argument A and argument B are not justified. Let us next examine whether they are defensible or overruled. Argument A has just one defeater, namely, argument B, but we have just seen that B is not justified. So A is not overruled but defensible. Argument B also has just one defeater, namely, argument C. Is C justified? No it is not: if the proponent starts a game with C, then the opponent can reply with argument D, which weakly defeats C. Then the proponent cannot find an argument that strictly defeats D (since C only weakly defeats C), so the game ends with a win for the opponent. So B's only defeater is not justified (in fact it is defensible), so B is also defensible.

4 Defeasible Rules and Generalisations

So far we have only considered deductive argument schemes and we have modelled the fallibility of arguments as the possibility of premise attack (with one exception, namely, arguments C, continued, and D, to be further discussed below). At first sight, it might be thought that this is all we need: if we adopt a suitable logic for adjudicating conclusion-premise conflicts between arguments, then the only argument schemes we need are those of deductive logic. However, if we have a closer look at arguments as they are constructed and attacked in practice, we see that they can often be attacked even if all their premises are accepted. In fact, we have already encountered one such case. Let us again look at the above arguments for and against the claim that John's mother consented. Both arguments used a premise 'What witness say is usually true' but from this premise and a given witness statement it does not *necessarily* follow that what the witness says is true. The circumstances may be unusual: for example, the witness may have bad memory or may have reason to lie (in our example mother might be afraid for father, if he is known to be violent). In such cases the conclusion only *presumptively* follows, namely, under the presumption that everything is as usual. This explains why a rational person can accept both premises of arguments C and D but still deny their conclusion, namely, if there is evidence that things are not as usual.

Many factual generalisations that are used in daily life and also many that are used in in legal proof are of this presumptive, or *defeasible* nature. At the beginning of this chapter we gave the example *usually married people live where their spouses live*, which can have exceptions, for instance, when a spouse temporarily works abroad. An evidential example is the generalisation *fleeing from the crime scene indicates consciousness of guilt*, which has exceptions in case the person fleeing from the crime scene has another reason for avoiding the police, such as being an illegal immigrant. Or consider *This type of radar is a reliable source of information about speed of traffic*; but what if it has a hidden defect? Or *Confessing suspects are guilty*; but what if the suspect later retracted his confession? Not only factual generalisations are defeasible but also, for example, interpretation rules or reasons for action. An example of a defeasible interpretation rule in American contract law is *A statement 'I accept ...'* is an acceptance, but an exception was *A statement 'I accept'* followed by terms that do not match the terms of the offer is not an acceptance (cf. Gardner 1987). Reasons for action are also often defeasible, for example: *When I have an exam, I should study hard*; but not necessarily, for example, when a close friend or family member is seriously ill. It has even been argued that legal rules are also defeasible, since there can always be unforeseen cases in which a rule should be set aside because of higher principles or unwanted consequences. A famous example is the *Riggs v. Palmer* case in American inheritance law (discussed by Dworkin 1977), in which a grandson had killed his grandfather and then claimed his share in the inheritance. The court made an exception to inheritance law based on the principle that no person shall profit from their own wrongdoing.

Now an important point is that the application of defeasible generalisation cannot be regarded as an instance of the *modus ponens* argument scheme, since $P \rightarrow Q$ means that *always* when P is true then Q is true, and this is not the same as saying that *usually* when P is true then Q is true. So we need a new argument scheme, called *defeasible modus ponens*:

Defeasible Modus Ponens:
If P then usually Q
P
Therefore (presumably) Q

The qualifier 'presumably' of the inference indicates that this is not a deductively valid scheme. Even if both premises are accepted, it may be rational not to accept the conclusion, namely, if there is an exception to the first premise. However, whether this is rational depends on whether an acceptable counterargument can be constructed: if this is impossible, that is, if there is no evidence that there is an exception, or this evidence is unconvincing, then an argument that instantiates the defeasible modus ponens scheme must be accepted. Now in the previous section we saw how an argumentation logic systematises this process of testing an argument in light of all possible counterarguments. The only thing left to do is to allow that an argument can be attacked not only on its premises but also on its conclusion.

In fact, a conclusion of an argument can be attacked in a stronger and a weaker way. The strong way is to build an argument with the opposite conclusion, as we did above with argument D. Such a conclusion-to-conclusion attack is called a *rebutting* attack. A rebutting counterargument may attack the final conclusion of its target but it may also attack an intermediate conclusion. For example, as Figure 3 shows, argument D above attacks argument C by rebutting its intermediate conclusion that John's mother consented. However, sometimes an argument can be attacked in a weaker way, namely, by saying that the premises, even if true, do not support their conclusion in the case at hand, because the case at hand is an exceptional case. Consider again the factual generalisation *What witnesses say is usually true*: if, for example, it turns out that the witness has a poor memory, this is a reason not to infer that what he says is true, but of course, this does not imply that the opposite of what he says is true. This weaker form of attack is often called *undercutting* attack. Undercutting counterarguments do not attack a premise or the conclusion of their target but instead deny that the scheme on which it is based can be applied to the case at hand. Obviously, such a denial does not make sense for deductive argument schemes. In sum, while deductive arguments can only be attacked on their premises, presumptive arguments can also be attacked on their conclusion and on their inference steps.

5 Presumptive Argument Schemes

Our analysis can be further refined. When looking at defeasible rules or generalisations, we see that often they are not just specific statements about the world but conform to certain reasoning patterns. For instance, evidential arguments are often based on stereotypical evidential sources, such as expert or witness testimony, observation or memory. Other evidential arguments apply the scheme of causal abduction: if we know that A causes B and we observe B, then in the absence of evidence of other possible causes we may presumptively conclude that it is A that caused B. It is important to note that arguments based on such patterns speak about states of affairs in general: unlike specific generalisations like 'summer in Holland is usually cool' or 'fleeing from a crime scene typically indicates consciousness of guilt' they express general ways of obtaining knowledge from certain information. Nobody would dispute these ways in general; at most their application to specific cases is disputed on the grounds that there is an exception. For these reasons it is natural to regard such patterns not as conditional premises of a *defeasible modus ponens* argument but as independent presumptive (or defeasible) argument schemes. The idea of presumptive argument schemes has been developed in the fields of informal logic and argumentation theory. The standard reference is Walton (1996). In this section I discuss some schemes that are common in legal reasoning, to start with the argument scheme from witness testimony:

Argument Scheme from Witness Testimony:
Person *W* says that *P*
Person *W* was in a position to observe *P*
Therefore (presumably), *P*

As explained in the previous section, the use of presumptive argument schemes in an argument gives rise to two new ways of attack, namely, rebutting and undercutting attack. For example, an application of the scheme from witness testimony can be rebutted by an application of the same scheme to a contradicting witness. Undercutting attacks are in fact based on the idea that a presumptive argument scheme has typical exceptional circumstances in which it does not apply. For example, a witness testimony is typically criticised on the witnesses' truthfulness or the functioning of his memory or senses. In general, then, each argument scheme comes with a set of *critical questions* which, when answered negatively, give rise to undercutting counterarguments (or sometimes to rebutting counterarguments). For example, the Witness Testimony Scheme is often given the following critical questions:

Critical Questions to the Argument Scheme from Witness Testimony
W1: Is the witness truthful?
W2: Did the senses of the witness function properly?
W3: Does the memory of the witness function properly?

A related scheme is the scheme from expert testimony:

Argument Scheme from Expert Testimony:
Person *E* is an expert in domain *D*
Person *E* says that *P*
P is within domain *D*
Therefore (presumably), *P*

Obvious critical questions are:

Critical Questions:
E1: Is the expert truthful?
E2: 'is P consistent with what other experts say?'
E3: 'is P consistent with known evidence?'

Note that the second and third question do not point to undercutting attacks: E2 points to rebutting applications of the same scheme, while E3 points to rebutting applications of any scheme with the opposite conclusion.

Another important scheme for legal evidential reasoning is causal explanation, which is a form of abduction:

Argument Scheme from Causal Explanation:
P causes *Q*
Q is observed
Therefore (presumably), *P* is the case

(Weaker variants of this scheme can be obtained by replacing the first premise with '*P* usually causes *Q*' or '*P* can cause *Q*'.)

Critical questions:
C1: Can Q be caused by something else?

C2: Does P also cause something else which is known not to be the case?

Here is an example of this scheme and its critical questions:

An abductive argument:
This type of gun fires this type of bullet
A bullet of this type was found at the crime scene
Therefore, (presumably) the victim was shot at with this type of gun.

Critical question C1 can be used, for instance, when other types of guns also fire this type of bullet, while question C2 can be used as follows:

A rebuttal of the abductive argument:
This type of gun usually causes this type of wound
The victim had a different type of wound
Therefore (presumably) the victim was not shot at with this type of gun

This argument instantiates the following scheme:

Argument Scheme from Causal Refutation:
P usually causes *Q*
Q is not observed
Therefore (presumably), $\neg P$

The use of presumptive argument schemes is not confined to reasoning about the facts. For instance, arguments from practical reasoning (reasoning about what to do) often conform to the 'argument scheme from consequences':

Argument Scheme from Good Consequences:
If *A* is brought about, then good consequences may plausibly occur.
Therefore (presumably), *A* should be brought about.

An application of the argument scheme from consequences may be criticised by pointing at other ways than *A* to realise the same consequences or at negative consequences brought about by realising *A*:

Critical questions:
GC1: Are there other ways to bring about the good consequences?
GC2: Does *A* also bring about bad consequences?

The scheme also has a negative version:

Argument Scheme from Bad Consequences:
If *A* is brought about, then bad consequences may plausibly occur.
Therefore (presumably), *A* should not be brought about.

Critical question:
BC1: Does *A* also bring about good consequences?

In legal reasoning a typical way in which the schemes from consequences are used is in so-called teleological interpretation arguments. For example:

> If the term 'personal data' of the Dutch Data Protection Act is interpreted to include email addresses, then new legal measures against spam become possible, which is good.
> Therefore, the term 'personal data' of the Dutch Data Protection Act should be interpreted to include email addresses.

An argument using the positive version of the argument scheme from consequences may be rebutted by an argument using the negative version, such as:

> If the term 'personal data' of the Dutch Data Protection Act is interpreted to include email addresses, then new legal measures against spam become possible, which may result in a massive increase in litigation, which is bad.
> Therefore, the term 'personal data' of the Dutch Data Protection Act should be interpreted not to include email addresses.

We next list without illustration argument schemes for two further common forms of presumptive reasoning. The first is (enumerative) induction, which defeasibly infers generalisations from a collection of observations. This scheme can be formulated differently depending on the strength of the conditional connection:

> **Argument Scheme from Induction:**
> Most (all) observed *P*'s were *Q*'s
> Therefore (presumably), most (all) *P*'s are *Q*'s
>
> **Critical Questions:**
> I1: Is the number of observations large enough?
> I2: Was the selection of observations biased?

The following scheme is for analogical reasoning, a form of reasoning often used in the law.

> **Argument Scheme from Analogy:**
> Cases that are relevantly similar should be decided in the same way
> Cases *C1* and *C2* are relevantly similar
> Therefore (presumably), cases *C1* and *C2* should be decided in the same way
>
> **Critical Questions:**
> A1: Are there also relevant differences between the cases?
> A2: Are there relevant similarities with other cases that are decided differently?

But more sophisticated formulations of analogy are possible.

To give a final illustration of reasoning with argument schemes and their critical questions, consider an alleged murder case in which an accused called Bob claims he killed in self-defence and produces a witness called Roy, who testifies that the victim threatened the accused with a knife. The Witness Testimony scheme can then be instantiated as follows (using an intermediate reasoning step to infer the second premise of the scheme):

Argument A:
F1: Roy says that the victim threatened the accused Bob with a knife
F2: Roy was there when the killing took place
G1: If *W* is there when an event *E* takes place, then *W* is in the position to observe event *E*
Therefore, Roy was in the position to observe whether the victim threatened the accused Bob with a knife
Therefore (presumably), the victim threatened the accused Bob with a knife

Let us represent critical questions that point to undercutting counterarguments as rules with a consequent ¬Name, where 'Name' is a placeholder for the name of the undercut argument scheme. Then, for instance, critical question W1 of the Witness Testimony Scheme can be represented as follows:

W1: ¬ Witness W is truthful → ¬ Witness-Testimony-Scheme

Suppose next that it becomes known that the witness is a friend of the accused. Then if it is believed that friends of accused persons, when testifying, try to protect their friend and that this means that they are not truthful, an undercutter of this argument can be constructed in three steps.

Argument B:
G2: If Witness *W* is a friend of accused *A*, then usually Witness *W* tries to protect accused *A*
F3: Witness Roy is a friend of accused Bob
Therefore (presumably), Witness Roy tries to protect accused Bob
R1: Witness *W* tries to protect accused *A* → ¬ Witness *W* is truthful
Therefore, ¬ Witness Roy is truthful
W: ¬ Witness *W* is truthful → ¬ Witness-Testimony-Scheme
Therefore, ¬ Witness-Testimony-Scheme

This argument could in turn be attacked, for instance, by attacking its premise F3 or by arguing that there is an exception to the defeasible generalisation G2.

Finally, how many presumptive argument schemes are there? Here the classical logician will be disappointed. One of the main successes of modern formal logic has been that an infinite number of valid deductive inferences can be captured by a finite and even very small number of schemes. However, things are different for defeasible inference: many different collections of presumptive argument schemes have been proposed; the above schemes are just a few examples of schemes. Moreover, while some schemes, such as abduction and the scheme from consequences, can arguably be used in any domain, other schemes may be domain-dependent. For instance, it can be argued that in legal contexts the witness and expert schemes have a different form and different critical questions than in ordinary commonsense reasoning.

6 Logic of Argumentation Formalised

Readers of this chapter who are interested in formal logic may want to know more about how an argumentation logic can be formally defined. For these readers I now discuss in more detail how the account of an argumentation logic developed throughout this paper can be formalised. I will do so by summarising a formal framework for argumentation logics as developed in artificial intelligence; some aspects will not be presented in their full formal detail but will be semi-formally sketched. I start with the formalisation of the structure of arguments and the nature of attack and defeat, and then I discuss a fully abstract approach to the dialectical evaluation of arguments.

6.1 The structure of arguments and the nature of attack and defeat[3]

In the previous sections arguments were depicted as trees of inferences, where the inferences apply either deductive (or 'strict') or presumptive (or 'defeasible') argument schemes (or 'inference rules'). Informally, that an inference rule is strict means that if one accepts all its antecedents, then one has to accept its consequent *no matter what*, while that an inference rule is defeasible means that if one accepts all its antecedents, then one has to accept its consequent *if one has no good reason not to accept it*.

This notion of an argument can be formalised assuming the following elements:

- A *logical language L* containing a negation symbol ¬.
- Two sets of *strict* and *defeasible inference rules Rs* and *Rd*. Each inference rule is of the form $a_1,\ldots,a_n$ *strictly/defeasibly implies c,* where $a_1,\ldots,a_n$ and *c* are well-formed formulas from *L*. The formulas $a_1,\ldots,a_n$ are called the *antecedents* and *c* is called the *consequent* of the rule.

Arguments are then inference trees, where the nodes are well-formed formulas from the logical language *L* and the links between the nodes are applications of strict or defeasible inference rules. More formally, for any node *n,* if $n_1,\ldots,n_k$ are all children of *n,* then there exists an inference rule in *Rs* or *Rd* of the form $n_1,\ldots,n_k$ *strictly/defeasibly implies c.* The starting points of an argument (the nodes not derived by applying an inference rule) are called the argument's *premises* and all other nodes are called its *conclusions*. The premises of an argument are assumed to be taken from a, possibly inconsistent, *knowledge base* (the available information, from wherever it originates). A conclusion that is not used as an antecedent of another inference in an argument is called the argument's *final conclusion*; the other conclusions are its *intermediate conclusions*.

For example, argument C in Figure 3 has three premises, *John's brother said that John's mother consented, What a witness said is usually true* and *Somebody's mother consented → S/he has consent*; it has one intermediate conclusion *John's mother consented* and a final conclusion *John has consent*.

Note that this notion of an argument does not commit to a particular logical language. For *L* any logical language can be chosen, such as the language of propositional logic, first-order predicate logic or deontic logic. The strict rules over *L*

[3] This subsection is based on the *ASPIC+* framework presented in Prakken (2010). Some definitions of ASPIC+ are here slightly changed for ease of explanation. The changes do not make a difference for the outcome of argument evaluation.

can be based on the semantic interpretation of *L* by saying that *Rs* contains all inference rules that are semantically valid over *L* (according to the chosen semantics). So, for example, if *L* is chosen to be the language of standard propositional logic, then *Rs* can be chosen to consist of all semantically valid inferences in standard propositional logic (whether such an inference is valid can be tested with, for example, the truth-table method).

The defeasible rules *Rd* cannot be based on the semantic interpretation of *L*, since they go beyond the meaning of the logical constants in *L*. Consider, for example, defeasible modus ponens, discussed in Section 4: 'if *P* then usually *Q*' and '*P*' do not together deductively imply *Q*, since we could have an unusual case of *P*. In other words, defeasible inference rules are deductively invalid. They can instead be based on insights from epistemology or argumentation theory. For example, *Rd* could be filled with presumptive argument schemes as discussed above in Section 5.

Arguments can be *attacked* in three ways: on their premises, on conclusions of their defeasible inferences and on their defeasible inferences themselves:

- An argument *A undermines* an argument *B* if a conclusion of *A* contradicts a premise of *B*.
- An argument *A rebuts* an argument *B* if a conclusion of *A* contradicts a conclusion of *B* that is derived with a defeasible inference rule.
- An argument *A undercuts* an argument *B* if a conclusion of *A* says of some defeasible rule used in *B* that it is not applicable.

Undermining and rebutting attacks can be resolved by looking at the relative strength of arguments:

- An argument *A successfully undermines* an argument *B* if *A* undermines *B* and *A* is not weaker than *B* with respect to this conflict.[4]
- An argument *A successfully rebuts* an argument *B* if *A* rebuts *B* and *A* is not weaker than *B* with respect to this conflict.

These definitions presuppose that the relative strength of arguments can be determined. This may not be trivial and may even be the subject of arguments and counterarguments. However, in this chapter we do not go into the details of this issue and simply assume that some definition of the relative strength of arguments is given.

Undercutting attacks succeed irrespective of the strength of arguments, since undercutters state exceptions to inference rules. Then:

- An argument *A defeats* an argument *B* if *A* successfully undermines *B*, successfully rebuts *B* or undercuts *B*.

Let us illustrate these notions with the arguments in Figure 3.

- The logical language *L* in the example is that of standard propositional logic
- The strict rules *Rs* are all propositionally valid inferences.

[4] Here we say "with respect to this conflict" since in general two arguments may simultaneously attack each other in several ways.

- The set of defeasible rules *Rd* contains the defeasible-modus-ponens rule discussed in Section 4.
- Arguments *A* and *B* only use strict inference rules, so they can only be attacked on their premises. Argument *C* first uses a defeasible rule, namely, defeasible modus ponens, and then uses a strict rule, namely, standard modus ponens. So *C* can be attacked on its premises, on its conclusion that mother consented and on the application of defeasible modus ponens to derive that conclusion. Finally, argument *D* uses defeasible modus ponens and can therefore be attacked on its premises, its conclusion and its inference.
- Argument *B* successfully undermines argument *A*, since *B*'s final conclusion contradicts a premise of *A* and since (as assumed above in Section 2) *B* is stronger than *A* with respect to this conflict since *B* provides a statutory exception to the legal rule that is *A*'s first premise (*A*'s third premise in fact assumes that there is no such exception). So *B* strictly defeats *A*.
- Argument *C* undermines argument *B* since *B*'s final conclusion contradicts a premise of *A*. Moreover, in Section 2 we assumed for unspecified reasons that *B* is stronger than *A* with respect to this conflict, so *C* successfully undermines *B* and so strictly defeats *B*.
- Arguments *C* and *D* rebut each other since they have contradictory conclusions *John has consent* and ¬ *John has consent*, both derived with a defeasible inference rule. Furthermore, in Section 2 it was assumed that no reason can be found why one witness is more reliable than the other, so both arguments are equally strong with respect to this conflict. So they successfully rebut each other and hence (weakly) defeat each other.

6.2 Dialectical evaluation of arguments[5]

As explained in Section 2, in order to determine whether an argument can be accepted, it does not suffice to make a choice between two arguments that directly conflict with each other but we must also look at how arguments can be indirectly defended by other arguments. One way to formalise this idea is to fully abstract from the internal structure of the arguments and from the reasons why they defeat each other. Thus in the example of Figure 3 we just say that there are four arguments *A*, *B*, *C* and *D* and that *B* defeats *A*, *C* defeats *B* while *C* and *D* defeat each other. This can be conveniently depicted as a directed graph, which from now on will be called a *defeat graph*:

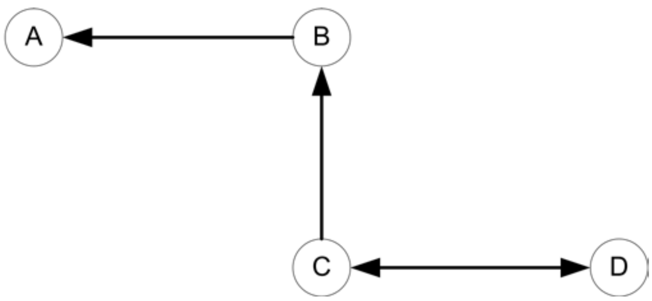

Figure 5: a defeat graph (and a 'grounded labelling').

[5] This subsection is based on Dung (1995) and later work extending it. See Prakken (2011) for a brief overview.

The task now is to determine which arguments in this graph can be accepted and which must be rejected given their various defeat relations. From now on we will speak of the task of labelling nodes in the graph with either the label *in* or the label *out* (but not both). Clearly, not any labelling will do but they must satisfy some conditions. It turns out that the same conditions can be used as above in Section 2:

1. An argument is *in* if and only if all its defeating counterarguments are *out*
2. An argument is *out* if and only if it has a defeating counterarguments that is *in*

However, this time we do not apply these conditions to labellings of a game tree, to see which party has a winning strategy in a game, but to labellings of the defeat graph.

Let us see if in Figure 5 argument *A* can be labelled *in*. To make *A in*, by condition (1) all defeaters of *A* must be *out*, so *B* must be *out*. To make *B out*, condition (2) requires that it has a defeater that is *in*, so *C* must be *in*. In turn, to make *C in*, condition (1) requires that *D* must be *out*. And to make *D out*, condition (2) requires that it has a defeater that is *in*, so *C* must be *in*. At first sight, the definition would here seem to turn out to be circular, but this is not true: all we need is to find a labelling that satisfies conditions (1) and (2) and if we choose the labelling as depicted in Figure 6, this is the case (here grey nodes stand for the label *in* and white nodes for the label *out*).

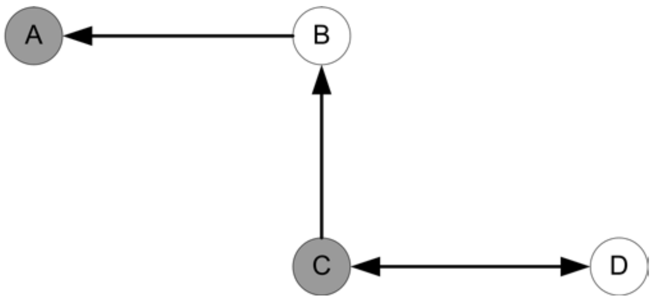

Figure 6: a 'preferred' labelling of the defeat graph.

However, this is not the only possible labelling that satisfies our two conditions. Since the defeat relation between *C* and *D* is symmetric and since neither *C* nor *D* is defeated by any other argument, we can just as well label *C out* and *D in*. But if we do so, then *B* must be labelled *in* (since all its defeaters are now *out*) while *A* must be labelled *out* (since it now has a defeater that is *in*). This alternative labelling is depicted in Figure 7:

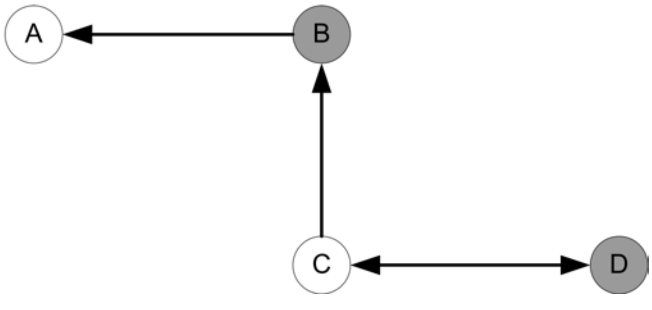

Figure 7: an alternative 'preferred' labelling of the defeat graph.

Thus the labelling approach captures that often in argumentation alternative reasonable positions can be defended. However, this is not yet all to the labelling approach, since we are not forced to assign a label to arguments *C* and *D*: if we do not assign any label to these two arguments, then this also satisfies conditions (1) and (2): that *C* is not labelled *in* is justified by condition (1), since not all defeaters of C are labelled *out*. And that *C* is not labelled *out* is justified by condition (2), since there exists no defeater of *C* that is labelled *in*. However, the choice not to label *C* and *D* implies that *A* and *B* can also not be labelled: *B* cannot be labelled *in* since it has a defeater which is not labelled *out* and *B* cannot be labelled *out* since it has no defeater that is labelled *in*. For the same reasons *A* cannot be labelled. This third labelling that satisfies conditions (1) and (2) is in fact the labelling depicted in Figure 5.

The three alternative labellings are induced by two alternative policies for labelling a defeat graph. The labellings of Figures 6 and 7 are induced by the policy to label as many arguments *in* as possible, while the labelling of Figure 5 is induced by the policy to abstain from labelling a node *in* whenever possible. In formal argumentation theory such policies for labelling defeat graphs are called 'argumentation semantics' and the policies to, respectively, maximise and minimise the sets of arguments that are labelled *in* are called, respectively, *preferred* and *grounded semantics*.[6] In fact, these semantics correspond to, respectively, a 'brave' and 'cautious' attitude in taking positions in argumentation (sometimes also called a 'credulous' and 'sceptical' attitude).

Since defeat graphs can have more than one labelling, a single labelling does in general not suffice to determine whether an argument is dialectically acceptable (with respect to a given semantics). Only if an argument is labelled *in* in all possible labellings according to the given semantics can it be said to be dialectically acceptable according to that semantics. In fact, a distinction can be made into three dialectical statuses of arguments. Let *S* be any given argumentation semantics. Then an argument *A* in a defeat graph *D* is:

1. *S-justified* if and only if *A* is labelled *in* in all *S*-labellings of *D*
2. *S-overruled* if and only if *A* is labelled *out* in all *S*-labellings of *D*
3. *S-defensible* otherwise.

These statuses can be carried over to conclusions of arguments as follows. A well-formed formula *P* for which an argument exists in *D* is:

4. *S-justified* if and only if *P* is the conclusion of some *S*-justified argument
5. *S-overruled* if and only if all arguments that have *P* as conclusion are *S*-overruled
6. *S-defensible* otherwise.

In our example, if *S* is either grounded or preferred semantics, then all arguments are defensible: in grounded semantics since they all have no label and in preferred semantics since all four arguments are labelled *in* in one preferred labelling and *out* in another preferred labelling.

[6] No further meaning should be attached to the terms 'preferred' and 'grounded'. Also, the term 'semantics' is here used in a different way than in deductive logic, where it is used for the meaning of a formal language.

Finally, what about the relation with the argument game that was sketched in Section 3 for determining the dialectical status of a given argument? It has been shown that this game corresponds to grounded semantics in that the proponent in the game has a winning strategy for a given argument just in case that argument is labelled in in the grounded labelling.

6.3 Labellings versus positions

More can be said about the positions that can be taken in argumentation. It can be shown that every defeat graph has a unique grounded labelling but (as shown by our example) defeat graphs can have multiple preferred labellings. This makes preferred labellings particularly interesting from a philosophical point of view, since they combine a foundationalist and coherentist view on knowledge. Very briefly, *Foundationalists* argue that something is knowledge if it can be derived (in some sense) from undisputed givens, while coherentists claim that something is knowledge if it is part of an (in some sense) coherent system of cognitions. To see how preferred semantics combines the two views, let us call any set of arguments from a defeat graph a *position*. Then we define a position *P* as *coherent*[7] if and only if

1. *P* is conflict-free (i.e., no argument in *P* defeats an argument in *P*); and
2. *P* defends all its members (i.e., all defeaters of a member of *P* are defeated by a member of *P*)

In our example, the smallest coherent position is the empty set but this position is, of course, not very interesting. Two more interesting positions are the sets *P1* = {*A*,*C*} and P2 = {*B*,*D*}: both are conflict free (no defeat relations within the sets) while they both defend all their members against defeats from outside: in position *P1* argument *C* defends *A* against *B* while *C* defends itself against *D*; and in position *P2* argument *D* defends both *B* and itself against *C*. So both *P1* and *P2* are coherent. From this it follows that positions {*C*} and {*D*} are also coherent (although not maximally coherent). However, positions {*A*} and {*B*} are not coherent: {*A*} does not defend *A* against *B* while {*B*} does not defend *B* against *C*. Furthermore, position {*A*,*D*} is incoherent since it does not defend *A* against *B*. Finally, any other position is not conflict-free.

The following properties of positions can be proven:

- If position *P* is coherent, then there exists some labelling of the defeat graph satisfying labelling conditions (1) and (2) (with respect to any semantics) in which all arguments in *P* are labelled *in*.
- If an argument *A* is *S*-justified or *S*-defensible (according to any given semantics *S*) then *A* is included in at least one coherent position.
- If two coherent positions *P1* and *P2* both only contain arguments that are labelled *in* in the grounded labelling, then their union *P1* U *P2* is also coherent.

The first two properties state relations between labellings and positions: the first one implies that every coherent position is included in the set of all *in*-arguments of at least

[7] Dung (1995) calls sets satisfying these conditions ‘admissible’.

one possible labelling. The third property in fact says that grounded semantics does not allow for alternative coherent positions (in the sense that their union is incoherent). By contrast, as our example shows, preferred semantics does allow for alternative positions that are internally coherent but jointly incoherent. Now preferred semantics combines the foundationalist and coherentist views in the following way: any argument must be based on premises taken from the knowledge base while to be justified or defensible it must be part of a coherent position.

7 Conclusion

To return to our question in the introduction, we have seen that there is indeed logic in defeasible argumentation: the form of arguments must fit a recognised argument scheme (whether deductive or defeasible), and the dialectical status of an argument must be determined in a systematic dialectical testing procedure. On the other hand, what cannot be provided by a logic for argumentation are the standards for comparing conflicting arguments: these are contingent input information, just like the information from which arguments can be constructed.

Let me finally return to Patrick Glenn's suggestion that to make logical sense of the fact that the world contains many different legal (sub)traditions, often inconsistent with each other and interacting in complex ways, new logics may be needed that are multivalent, paraconsistent or non-monotonic and do not adhere to the classic rules of non-contradiction and the excluded middle. In this chapter I have explored the use of one such new kind of logic, namely, logics of argumentation. It turns out that, even if such a logic includes classical twovalued logic as its strict core (the set of strict inference rules for constructing arguments), the resulting logic is (at least) three-valued: each conclusion of an argument has one of the three statuses *justified*, *defensible* or *overruled*. Another way to capture the multiplicity of legal traditions is in the notion of a coherent position discussed in Section 6.3: in general a body of legal information gives rise to multiple *positions*, each internally coherent but mutually conflicting. Argumentation logics are also clearly *nonmonotonic*. For example, an argument that is justified on the basis of given information can become defensible or even overruled if new information gives rise to new counterarguments. Argumentation logics are also *paraconsistent*: if two arguments have contradictory conclusions, this does not lead to trivialisation as in classical logic but gives rise to two rebutting arguments; at most one of these arguments can be in the same position but both arguments can be in alternative positions. Finally, argumentation logic *invalidates the law of non-contradiction* in that two statements P and $\neg P$ can both be defensible, and it *invalidates the law of the excluded middle* in that it is not the case that for any statement P, either P or its negation $\neg P$ is justified. In sum, argumentation logics satisfy all the properties that according to Patrick Glenn may be needed to capture the multiplicity of legal traditions.[8] Moreover, argumentation logics are arguably close to how lawyers think, since notions like argument, counterargument and rebuttal are natural them.

[8] Strictly speaking this only holds for argumentation logics that are well-defined but explaining this goes beyond this chapter.

References

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

Dworkin, R.M. (1977), *Taking Rights Seriously* (Cambridge, MA: Harvard University Press).

Gardner, A. von der Lieth (1987), *An Artificial Intelligence Approach to Legal Reasoning* (Cambridge, MA: MIT Press).

Glenn, P. (2010), *Legal Traditions of the World*, 4th edition. Oxford: Oxford University Press.

Glenn, P. (2013), *The Cosmopolitan State*. Oxford: Oxford University Press.

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

Prakken (2011), An overview of formal models of argumentation and their application in philosophy. *Studies in Logic* 4 (2011): 65-86.

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*, Aldershot: Ashgate Publishing 2009, p. 223-253.

Prakken, H. & Sartor, G. (2015), Law and logic: a review from an argumentation perspective. *Artificial Intelligence* 227 (2015): 214-245.

Walton, D.N., (1996), *Argumentation Schemes for Presumptive Reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.