

Revisability versus defeasibility

BARTOSZ BROZEK*

Jagiellonian University, Kraków, Poland

IN this paper I would like to address two issues. First, the problem of distinguishing between revisability and defeasibility will be investigated. I will try to argue that one should differentiate between the two. Second, the role of both concepts in the context of formal reconstruction of legal reasoning will be analysed.

Both aims are crucial from the point of view of the contemporary debates in legal theory. It is not difficult to find authors advocating the thesis that there is no difference between both concepts.¹ Furthermore, there are some other notions often associated with defeasibility and revisability, notably open texture and vagueness.² I believe it is important to differentiate carefully between those four concepts. This paper is, therefore, an attempt to fulfill partially this project and clarify the distinction between defeasibility and revisability.

1 Revisability and defeasibility from epistemological perspective

Let us consider the following example. A new illness has been discovered. Let us call it A. Let us assume, further, that the symptoms of A are denoted by SYMPTOMS_A. The following cognitive rule can be formulated on the basis of the above-mentioned facts:

(1) If x has SYMPTOMS_A, then x has illness A.

Whenever we meet someone with SYMPTOMS_A, we may conclude, on the basis of (1), that she has A. Imagine, however, that we meet a person, Arthur, who has SYMPTOMS_A, but the therapy that helps people with A does not help Arthur. We may reasonably suspect that Arthur has some other illness. We look for other symptoms and find a strange behaviour: whenever Arthur reads Hegel's works, he has spasms. Additional research leads to the conclusion that Arthur's illness should be treated in a different way to illness A. Let us call Arthur's illness *hegeliosis spasmatICA* and the symptom that distinguishes *hegeliosis* from

* This paper was written during my research stay at Professor Robert Alexy's department at Christian Albrechts University in Kiel, sponsored by Alexander von Humboldt-Stiftung. My thanks are due to George Pavlakos for reading a draft of the paper and offering valuable comments.

1 C. Alchourrón, "Philosophical foundations of deontic logic and the logic of defeasible conditionals" in J. Meyer and R. Wieringa (eds), *Deontic Logic in Computer Science* (Chichester: Wiley, 1993) pp. 43–84.
2 J. Hage, *Reasoning with Rules. An essay on legal reasoning and its underlying logic* (Dordrecht-Boston-London: Kluwer, 1997); H. Prakken, *Logical Tools for Modelling Legal Argument. Study of defeasible reasoning in law* (Dordrecht-Boston-London: Kluwer, 1997); B. Brozek, *Defeasibility of Legal Reasoning* (Kraków: Zakamycze, 2004).

the illness A *Schopenhauer's Syndrome* (SS). Now we know that rule (1) is incorrect. We have to revise the rule in the following way:

(2) If x has SYMPTOMS_A and it is not the case that x has SS, then x has illness A.

Imagine further that we meet another patient, John, who has SYMPTOMS_A, does not have SS, but the therapy for A does not help him. After a series of complicated blood tests we determine that John has an abnormal level of calcium in his blood. In this way we identify another illness: B. Therefore we revise our rule (2) once more:

(3) If x has SYMPTOMS_A and it is not the case that x has SS and it is not the case that x has SYMPTOM_B, then x has illness A.

Such revisions of cognitive rules are typical. The most detailed description of these types of activities can be found in textbooks on the philosophy of science.³ There is no doubt, however, that in this way we revise many cognitive rules that we use every day.

Logicians working on this problem developed what are called formal theories of belief revision (they concern a somewhat wider range of problems but nevertheless they include the problem of revision). Formal methods cannot determine how the new rule looks. They can tell us, however, what to abandon from our theory in order to keep it consistent after the introduction of the new rule.

For instance, if our previous theory contained, inter alia, rule (1) and we now introduce rule (2), then (1) together with (2) and some other assumptions would yield contradiction. Therefore, a theory of belief revision indicates that the introduction of (2) should result, speaking somewhat boldly, in the abandoning of (1) and of all the sentences belonging to our theory from which (1) follows.

What has been said may serve as a basis for formulating the following definition of revisability:

- (DEF1) A cognitive rule is revisable if and only if it can be substituted with a rule or a set of rules that:
 - (a) grasp the reality better than the revisable rule; and
 - (b) together with the rule yield a contradiction.⁴

Revisability should be carefully distinguished from defeasibility. The concept of defeasibility was introduced in 1948 by H L A Hart in his "The ascription of responsibility and rights". Here is the passage from Hart's paper where the concept is introduced and labelled:

When the student has learnt that in English law there are positive conditions required for the existence of a valid contract . . . his understanding of the legal concept of a contract is still incomplete . . . For these conditions, although necessary, are not always sufficient and he has still to learn what can *defeat* a claim that there is a valid contract, even though all these conditions are satisfied. The

3 K R Popper, *The Logic of Scientific Discovery* (London: Hutchinson & Co/New York: Basic Books Inc., 1959); K R Popper, *Conjectures and Refutations: The growth of scientific knowledge* (London: Routledge & Kegan Paul/New York: Basic Books Inc., 1963).

4 The phrase "grasp the reality better" is somewhat vague. One can think of it in terms of the Popperian verisimilitude. Furthermore, a cognitive rule can be substituted with a single rule or a set of rules. Think of the following example: after identifying *hegeliosis spasmatia*, apart from introducing (2), we have to introduce also the rule: (R) If x has SYMPTOMS_A and x has SS, then x has *hegeliosis spasmatia*. Both (2) and (R) together with (1) yield a contradiction – in the former case an obvious statement "if someone has *hegeliosis* it is not the case that she has the illness A" has to be added for the contradiction to occur. But one can also think of a situation in which we have to take both (or more) newly introduced rules together with the old one to derive a contradiction.

student has still to learn what can follow on the word “unless”, which should accompany the statement of these conditions. This characteristic of legal concepts is one for which no word exists in ordinary English. The words “conditional” and “negative” have the wrong implications, but the law has a word which with some hesitation I borrow and extend: this is the word “*defeasible*”, used of a legal interest in property which is subject to termination or defeat in a number of different contingencies but remains intact if no such contingencies mature. In this sense, then, contract is a defeasible concept.⁵

The concept of defeasibility, introduced within the framework of legal and ethical theory, proved useful in general epistemology.⁶ Moreover, it prompted important logical research resulting in the development of numerous defeasible logics.⁷ The logical developments also caused a certain terminological shift. While Hart in his seminal paper speaks of the defeasibility of concepts, contemporary accounts investigate the defeasibility of rules. The shift is easy to explain and not as substantial as it may seem. We ascribe certain concepts to certain phenomena on the basis of some rule. For instance, we have a rule that says that if a given creature is capable of thinking then it is a human. So, according to that rule, we should ascribe the concept of being a human only to those creatures that are capable of thinking. If, then, the rule we use in the process of ascribing concepts is defeasible, so is the ascribed concept.

Let us come back to our initial problem. Imagine a patient comes to us who has SYMPTOMS_A. Moreover, he is in a critical condition and our withdrawal from applying a therapy will result in his death. Additionally, we cannot test whether SS or SYMPTOM_B are present (say, because there are no works by Hegel at hand and tests for SYMPTOM_B would take too much time). In such circumstances we know that the patient has one of three things: illness A, *hegeliosis* or illness B. However, we have to assume that the patient has A and immediately start therapy for A. Therefore we reason on the basis of (1):

- (1) If x has SYMPTOMS_A, then x has illness A.

Applying (1) usually has some additional justification like, for example, the fact that A occurs far more often than either *hegeliosis* or B. The point is, however, that in particular cases we do not employ probabilistic reasoning (maybe with the exception of the first time we encounter the problem) because the patient has SYMPTOMS_A and we cannot check whether she has SS or SYMPTOM_B and illness A occurs more often than *hegeliosis* or illness B, therefore we apply the therapy for A. We reason simply using (1).⁸

The main logical problem is that we do not reason here with the use of (3):

- (3) If x has SYMPTOMS_A and it is not the case that x has SS and it is not the case that x has SYMPTOM_B, then x has illness A;

for we cannot determine whether our patient has or does not have SS or SYMPTOM_B. A further point is that if we knew that the patient had SS we would not use (1) but a new rule:

- (4) If x has SYMPTOMS_A and x has SS, then x has *hegeliosis spasmatica*.

5 H L A Hart, “The ascription of responsibility and rights” in A Flew (ed.), *Logic and Language* (Oxford: Blackwell, 1951), p. 152. First published in *Proceedings of the Aristotelian Society*, 1948–49.

6 See J Pollock, “Defeasible reasoning” (1987) 11 *Cognitive Science* 481–518.

7 H Prakken and G Vreeswijk “Logical systems for defeasible argumentation” in D Gabbay and F Guentner (eds), *Handbook of Philosophical Logic* 2nd edn, vol. 4 (Dordrecht-Boston-London: Kluwer, 2002) pp. 219–318.

8 An even better example is the following: “It is sunny today, therefore I will go swimming.” We reason very often using such a rule even though we know that there are various circumstances – apart from the weather – that can stop us going swimming, e.g. an uncle can ask for help, or we can break a leg, etc.

From this it is visible that a logic needed to model this kind of reasoning has to be nonmonotonic.⁹ In the above example, from a set of premises containing (1), (4) and the statement of the fact that the patient has SYMPTOMS_A, the conclusion that the patient has A follows. From a superset of the previous set of premises, containing in addition the information that the patient has SS, the conclusion that the patient has A no longer follows.

The question arises why we have to use defeasible cognitive rules in our daily life. There are several possible answers. Some have to do with the lack of information. Sometimes we use simplified cognitive rules because it is impossible to collect all the information required to apply the full rule (for example, because of time limitations). Another reason for using simplified rules is effectiveness. Sometimes it would be ineffective to try to collect all the required information because the failure in applying a defeasible rule would cost much less than the process of collecting the missing information. A third possibility is that a full rule would be too complicated and therefore useless, let's say, because of our limited brain capabilities.

On the basis of what has been said a definition of defeasibility may be formulated as follows:

- (DEF2) A cognitive rule is defeasible if and only if there are situations in which its antecedent¹⁰ is fulfilled but its consequent does not follow.

One can argue that the use of nonmonotonic logic for modelling reasoning with defeasible rules is unnecessary and, what follows, that the (DEF2) is wrong. The “classical logic” solution would look like this. In the case of our example we know that the patient has SYMPTOMS_A but we cannot know whether he has SS or SYMPTOM_B. Nevertheless, the argument runs, we use (3):

- (3) If x has SYMPTOMS_A and it is not the case that x has SS and it is not the case that x has SYMPTOM_B, then x has illness A.

Two conjuncts of the antecedent – that the patient does not have SS and that he does not have SYMPTOM_B – cannot be proved but they are assumed. In such a case the definition of defeasibility looks like this:

- (DEF3) A cognitive rule is defeasible if and only if there are situations in which it is applied with only assuming and not proving some of the conjuncts constituting its antecedent.¹¹

Postponing the discussion about which of the two – (DEF2) or (DEF3) – is better, we can conclude (on the basis of (DEF1), (DEF2) and (DEF3)) that revisability is different to defeasibility.

2 Defeasibility and revisability in legal discourse

Having established the definitions of revisability and defeasibility, I would like to turn now to legal discourse and consider three problems that are usually discussed in connection with defeasibility. The problems in question are: (a) structural resemblance between legal texts and their formalisations; (b) the problem of the shifting burden of proof; and (c) theory of rules and principles. The aim of this part of the paper is to determine whether legal rules are defeasible and/or revisable in the senses explicated above.

9 Classical logic is monotonic. It means that if something follows logically from a given set of premises, it follows also from a superset of it. A logic which is not monotonic is called nonmonotonic. See J Horty, “Nonmonotonic logic” in L Goble (ed.), *The Blackwell Guide to Philosophical Logic* (Malden-Oxford: Blackwell, 2001) pp. 336-61.

10 “Antecedent” can be substituted here with “conditions of rule’s application”.

11 Or equivalently: “only assuming some of its application conditions”.

2.1 STRUCTURAL RESEMBLANCE

Let us assume we have the following legal provisions:

- (§X) Whoever causes damage to someone is obliged to redress it.
- (§Y) Whoever acts in self-defence is not obliged to redress the damage caused by the action.

A first order predicate logic formalisation of the provisions would look as follows:

$$(4) \forall x (Cx \rightarrow Rx)$$

$$(5) \forall x (Dx \rightarrow \neg Rx)$$

where (4) is a formalisation of (§X), (5) is a formalisation of (§Y), C stands for “causes damage to someone”, R for “has to redress the damage” and D for “acts in self-defence”. There is an obvious difficulty with the proposed formalisation. Imagine John (denoted by *j*) has caused damage to someone

$$(6) Cj$$

and that he acted in self-defence

$$(7) Dj.$$

From (4) and (6) we obtain

$$(8) Rj$$

and from (5) and (7) we have

$$(9) \neg Rj$$

and therefore we have a contradiction.

A natural way out of this trouble would be to revise (4) in the following way:

$$(10) \forall x ((Cx \wedge \neg Dx) \rightarrow Rx).$$

Now, the derivation leading to contradiction is blocked, for (10) together with the sole (6) does not yield (8). There is one problem with this solution, however. (10), as a formalisation of (§X), does not resemble structurally its natural-language counterpart. There is some information in (10) – namely $\neg Dx$ – which cannot be found in (§X). In order to save the structural resemblance between (§X) and its formalisation we have to stick to (4):

$$(4) \forall x (Cx \rightarrow Rx)^{12}$$

But this leads to declaring (4) defeasible, for there are situations (like the one described above in which John has caused damage to someone but acted in self-defence) in which the antecedent of (4) is fulfilled (Cx obtains for a specific x), but nevertheless Rx (for a specific x) does not follow.

Observe that if we want to keep structural resemblance between legal texts and their formalisations we have to declare legal rules defeasible in the sense of (DEF2). (DEF3) would lead us to formalising (§X) in the form of (10):

$$(10) \forall x ((Cx \wedge \neg Dx) \rightarrow Rx)$$

and only assuming $\neg Dj$ in case we have no information concerning John’s acting in self-defence.

It seems therefore that if one wants to keep structural resemblance between legal provisions and their formalisations one has to declare legal rules defeasible in the sense of (DEF2).

12 A better way of formulating (4) would be to write: (4) $\forall x (Cx \Rightarrow Rx)$, where “ \Rightarrow ” stands for the so-called defeasible implication. If the rule is defeasible in the sense of (DEF2), one cannot use the material implication “ \rightarrow ” of the classical logic to model it.

2.2. BURDEN OF PROOF

The next problem I would like to investigate is how to give a formal account of the shifts of the burden of proof that occur in the process of legal argumentation. In a legal dispute there are two parties that try to win the case. When a party puts forward an argument to the effect that p , the burden of proof shifts onto the other party, which may attack either the conclusion or one of the premises of the first party's argument.

Let us come back to (§X) and (§Y):

- (§X) Whoever causes damage to someone is obliged to redress it.
- (§Y) Whoever acts in self-defence is not obliged to redress the damage caused by the action.

Imagine further that John has caused damage to Charles and that he acted in self-defence. In order to reconstruct formally the dispute between Charles and John it is natural to assume that Charles starts with an argument based on (§X). Which of the formalisations of (§X) should we choose? Observe that if we decided to use (10)

$$(10) \forall x ((Cx \wedge \neg Dx) \rightarrow Rx)$$

Charles – in order to argue to the effect that John has to redress the damage – would have to prove not only that John has caused him damage but also that John has not acted in self-defence. This result is incompatible with the idea of the shifts of the burden of proof. From this perspective one would expect Charles to prove only the damage caused by John, and John to show that he acted in self-defence.

A way out of this trouble is to declare (10) defeasible in the sense of (DEF3). In this case Charles would have to prove that John has caused him damage but he would only assume that John did not act in self-defence. Charles's argument becomes then:

$$(10) \forall x ((Cx \wedge \neg Dx) \rightarrow Rx)$$

$$(11) C_j$$

$$(12) \text{ assumed: } \neg D_j \text{ therefore}$$

$$(13) R_j^{13}$$

John's response would be to attack the premise assumed by Charles and prove that he acted in self-defence. The problem with this solution is that the formalisation of Charles's argument would have to take into account and assume all the possible exceptions to the general rule (§X). Therefore the structure of Charles's argument is (where EXC_i ranges over the respective exceptions to (§X)):

$$(14) \forall x ((Cx \wedge \neg Dx \wedge \neg EXC_1 \wedge \neg EXC_2 \dots \wedge \dots \wedge \neg EXC_n) \rightarrow Rx)$$

$$(15) C_j$$

$$(16) \text{ assumed: } \neg D_j$$

$$(17) \text{ assumed: } \neg EXC_1$$

$$(18) \text{ assumed: } \neg EXC_2$$

.....

$$\text{assumed: } \neg EXC_n \text{ therefore}$$

$$(19) R_j$$

Similarly, John's counterargument would have to take into account all the possible exceptions to exceptions, becoming – in some cases – extremely complicated.

13 The obvious step of universal instantiation is omitted throughout the paper for the sake of readability.

Another way of dealing with the problem of the shifts of the burden of proof is to declare (§X) defeasible in the sense of (DEF2). Then Charles's argument becomes:

(20) $\forall x (Cx \Rightarrow Rx)$ ¹⁴

(21) Cj therefore

(22) Rj

And John's counterargument is:

(23) $\forall x (Dx \Rightarrow \neg Rx)$

(24) Dj therefore

(25) $\neg Rj$

Of course with such a formal reconstruction we need a logical system – different from classical logic – capable of comparing both arguments and deciding which of the competing conclusions – Rj or $\neg Rj$ – prevails. If John's argument wins then it is clear that the rule used in Charles's argument has to be defeasible in the sense of (DEF2) for there exists a situation in which Cx (for a specific x) obtains but Rx (for a specific x) does not follow.

One more observation that has to be added here is that revisability cannot help us in any conceivable way to deal formally with the problem of the burden of proof.

To summarise: in order to give a formal account of the shifts of the burden of proof, one has to declare legal rules defeasible either in the sense of (DEF2) or of (DEF3).

2.3 RULES AND PRINCIPLES THEORY

The last problem I would like to address in the context of applying the notions of revisability and defeasibility to legal argumentation is the distinction between legal rules and principles. I will not go into the details of the distinction.¹⁵ Instead I would like to concentrate on one particular issue, namely on the fact that legal principles can “produce” exceptions to legal rules.

Let us consider the following example.¹⁶ Assume we have a legal rule stating that:

- (§Z) Vehicles are not allowed into the public park.

Let V stand for “is a vehicle” and P for “may enter the public park”. The formalisation of (§Z) looks like follows:

(26) $\forall x (Vx \rightarrow \neg Px)$

Imagine now that we have an ambulance carrying a badly injured person. The shortest way to the hospital is through the park. If we applied (26), however – with an obvious assumption that an ambulance is a vehicle – the result would be that the ambulance cannot enter the park:

(26) $\forall x (Vx \rightarrow \neg Px)$

(27) Va (a stands here for our ambulance) therefore

(28) $\neg Pa$

14 For the reasons of using “ \Rightarrow ” see n. 12.

15 For details see: R Dworkin, *Taking Rights Seriously* (Cambridge, Mass: Harvard UP, 1977); R Alexy, *A Theory of Constitutional Rights*, J Rivers (trans.) (Oxford: OUP, 2002); B Brozek, “The logic of rules and principles” (forthcoming).

16 It is a modified version of H L A Hart's example from *The Concept of Law* (Oxford: OUP, 1994, 1st edn 1961).

This answer to the question whether the ambulance can drive through the park is troublesome. It seems that the “correct” answer would be to let the ambulance through. The theory of rules and principles offers a very persuasive diagnosis of such cases. According to the theory we have here a legal rule, that is (26), which in this particular case is in conflict with a legal principle stating that human life and health should be protected by the law. Leaving aside the description of the exact mechanism of the interaction between rules and principles, one may say that in our case the principle produces an exception to the rule (26): vehicles are not allowed into the public park but ambulances carrying injured persons are.

The question we have to answer now is what the proper formal account of the described situation is. One way is to say that (26) is revisable. The revision occurs when there is a collision between the rule and a principle. Let A denote “is an ambulance carrying an injured person”. Now, the collision between (26) and a principle in the above presented case causes the following revision of (26):

$$(29) \forall x ((\forall x \wedge \neg Ax) \rightarrow \neg Px)$$

Such revisions have to occur in every single case in which the rule collides with a principle and the principle prevails. As the list of cases of collision is potentially endless, the potential number of revisions is likewise endless.

Another way of giving a formal account of the interaction between rules and principles is to say that legal rules are defeasible in the sense of (DEF2). In our example the rule

$$(26) \forall x (\forall x \rightarrow \neg Px)$$

is defeasible for there are situations – like the one with the ambulance – when $\forall x$ (for a specific x) is fulfilled but $\neg Px$ (for a specific x) does not follow. Observe that now the formulation of (26) does not change. Having declared (26) defeasible in the sense of (DEF2), we have the rule “prepared” for any possible conflict with a principle. In case of such a conflict in which the principle prevails, the formulation of (26) stays intact but the logical system blocks the derivation of $\neg Px$ (for a specific x).

One more observation has to be added here. There seems to be no conceivable way of dealing with the problems of the theory of rules and principles with the use of the notion of defeasibility in the sense of (DEF3).

To summarise: collisions between rules and principles can be handled formally either with the use of the notion of revisability or the notion of defeasibility in the sense of (DEF2).

3 Conclusions

In this paper I have tried to distinguish between revisability and defeasibility and to show what role both notions play in legal discourse. I have defined revisability and distinguished between two senses of defeasibility. Out of the three legal-theoretic problems investigated, all can be accounted for formally with the use of the notion of defeasibility in the sense of (DEF2). Defeasibility in the sense of (DEF3) helps us to deal with the problems of the burden of proof. Finally, revisability offers a way of reconstructing formally collisions between rules and principles.

It is tempting to say, therefore, that the (DEF2)-defeasibility, which helps to deal with all three problems, is the concept to be used in addressing them. On the other hand, one may argue that as we have to do with three different phenomena, we should explain them with three different formal structures. The choice between the two strategies is not an easy one. A full discussion of the issue requires significantly more than this short essay. Nevertheless, the following observation can be made. Both the idea of structural resemblance between legal texts and their formalisation and the problems surrounding

shifts of the burden of proof require developing a formal theory of legal reasoning that adapts a different perspective than the usual theorising on legal argumentation.

The classical way of approaching legal reasoning is from the point of view of a judge justifying his or her decisions.¹⁷ From this perspective a maximally rational procedure is needed and therefore fully reconstructed legal norms and indefeasible conclusions are desired. Theories of legal reasoning that adapt this attitude can be seen as purely normative and, moreover, counterfactual. In actual situations it is difficult for a judge to take into account all the possible exceptions to a norm that serves as the basis for a decision (think of a legal norm prohibiting killing, which usually has many exceptions of which only the relevant few are directly addressed in legal disputes).

Such unachievable, normative models of legal reasoning are well supported by the concept of revisability. On the other hand, nonmonotonic logics are designed for “subideal” situations, for example those in which there is limited access to information. Both the problem of structural resemblance and that of the shifts of the burden of proof are issues connected with such “subideal” and “dynamic” situations. Therefore it is reasonable to say that defeasibility in the sense of (DEF2) is well suited for dealing with the first two problems described above. (DEF3)-defeasibility is not a good answer to the troubles caused by shifts of the burden of proof because with the complicated argumentation structures it embraces it cannot be seen as a proper account of a “subideal” situation.

As to the problem of rules and principles it can be said that the choice between revisability and (DEF2)-defeasibility depends on the kind of theory of legal reasoning one wishes to construct. If an “ideal” theory is pursued, then revisability seems to be the answer. If, on the other hand, one develops a more descriptive theory then one should make use of defeasibility in the sense of (DEF2).

17 A Grabowski, *A Judicial Argumentation and Pragmatics* (Kraków: Księgarnia Akademicka, 1999).

