

MODELING LEGAL RULES

RICHARD HOLTON, MIT

Common law rules admit of exceptions. When a court, especially a higher court, finds that the routine application of a rule would result in an injustice, it is likely to distinguish. It will concede that yes, the case does appear to fall under the rule as it is currently understood; but will insist that there are further factors, not mentioned in the rule (though perhaps acknowledged in other rules in other parts of the law) that distinguish this case from the cases that the existing rule was meant to cover. The court will conclude that in this case the verdict that the existing rule suggests would be wrong. Nevertheless, the old rule does not die.¹ When the writers of case books come to accommodate the new ruling it will come in as an amendment: the old rule was correct except under these new circumstances.

So how should we understand the form of legal rules? A simple minded approach is to see them as universally quantified claims: whenever this holds, then this is the right verdict. But exceptions make this hard to maintain. A universally quantified sentence cannot have exceptions, only counterexamples; and counterexamples show that the sentence is false. Of course, one could insist on the approach, maintaining that a simple legal rule is indeed false, and that the need to amend it shows this to be so. But since every common law rule, however amended, is very likely to admit of further amendment, this leaves us in the uncomfortable position of saying no rule is strictly true. We might try to soften the blow by saying that as they are amended the rules get closer to the truth, but proposals to explain such an idea have had a very chequered history.²

So we need an approach that allows rules to have exceptions. Elsewhere I have developed an account that sees legal rules as universals containing implicit *unless*-clauses; the idea is that the exceptions trigger the clauses.³ The challenge comes in doing this in such a way that the rules do not become trivial. In the first part of this paper I reformulate and extend that account. In the second part I ask how it fares against some alternatives. I see two. One treats legal rules not as universal generalizations at all, but as generics. I argue that whilst this has some plausibility for legal *principles*, it doesn't do the job for legal *rules*. The second alternative is more radical, treating legal rules as default rules within a nonmonotonic logic. Here I argue that the move to nonmonotonic logic does not bring the advantages claimed for it, and further that it fails to explain something that is handled very nicely by the approach I

¹ Or at least, hardly ever. When the Australian High Court gave the *Mabo* decision, possibly the whole set of rules based around *terra nullius* were thrown out. But that is unusual.

² This was part of Karl Popper's idea of verisimilitude. David Miller and Pavel Tichy raised the classic problems with it. For review of the debate since see Graham Oddie 'Truthlikeness' *Stanford Encyclopedia of Philosophy* <<http://plato.stanford.edu/entries/truthlikeness/>>

³ See 'The Exception Proves the Rule', *Journal of Political Philosophy* forthcoming.

favor: how it is that a legal decision can be criticizable, even though the court used the legal rules that were in force at the time.

PARTICULARISM AND RULES

In *The Concept of Law* Hart writes:

We promise to visit a friend the next day. When the day comes it turns out that keeping the promise would involve neglecting someone dangerously ill. The fact that this is accepted as an adequate reason for not keeping the promise surely does not mean that there is no rule requiring promises to be kept, only a certain regularity in keeping them. It does not follow from the fact that rules have exceptions incapable of exhaustive statement, that in every situation we are left to our discretion and are never bound to keep a promise. A rule that ends with the word 'unless ...' is still a rule.⁴

So here Hart at least implicitly accepts that some valid rules have exceptions that cannot be exhaustively stated. The passage comes after his famous discussion of the open texture of law that results from the open texture of language—Is a child's bicycle covered by a regulation that prohibits vehicles from the park?—so one might think that the ideas here can be similarly explained. But clearly they cannot. Whilst there may be vagueness in the idea of a promise—Is it a promise if made under duress, or if the promisor does not understand what they are committing themselves to, or if the promisee is not aware of it?—such vagueness is not what is causing the open-endedness here. There may be no doubt that my promise was as clear and central an example as one is ever likely to find, and yet it still be true that in the circumstances I am not bound to keep it. So we need a different explanation.⁵

Hart's example might put us in mind of recent work on ethical particularism. Those who advocate such an approach contend that any putative rule is subject to exceptions; and as a result they tend to reject any role for rules. Take any rule that links the normative to the descriptive, they say, and we can find an exception to it; amend the rule to embrace the exception and we can find an exception to the amended rule, and so on. So the rules must be false.⁶

But it does not follow from the supposition that there are no exceptionless rules that rules are false or have no useful role to play. As we have seen, Hart suggests a model: perhaps rules end with an (often unstated) *unless*-clause; where this is not triggered,

⁴ H.L.A. Hart, *Concept of Law* (Oxford: Clarendon Press, 1961) p. 136

⁵ One complicating factor is that Hart's discussion here concerns *moral* rules rather than the legal rules that are the main subject of the book. This is somewhat surprising, since it is not obvious that what holds for moral rules also holds for legal, especially if, like Hart, one wants to maintain a broadly positivist framework. I have discussed whether they can be treated in the same way elsewhere; here I am going to assume that they can. See 'The Exception Proves the Rule'

⁶ Jonathan Dancy, 'Ethical Particularism and Morally Relevant Properties' *Mind* 92 (1983) 530–47; *Ethics Without Principles* (Oxford: Clarendon Press 2004).

the rule applies. Hart, however, does not tell us how the *unless*-clause is to be completed. And here it might seem that he is faced with a dilemma. On one approach the *unless*-clause contains a full statement of all the factors that would defeat the rule; but that is clearly incompatible with the idea that the exceptions are incapable of exhaustive statement. On the other approach the *unless*-clause would provide no information: 'One should keep one's promises unless one shouldn't'. But that is clearly trivial. If we are to give substance to an open-ended *unless*-clause we need to find a middle way between these two approaches. This is what I aim to provide.

THAT'S IT

The intuitive idea that I shall work with is that a rule can be over-ruled if there is a justification for the exception.⁷ In Hart's example there are moral grounds for tending to the seriously ill person rather than keeping my promise. But if so, then there is plausibly a moral rule that tells us that there are such grounds. So the *unless*-clause can be read as quantifying over other moral rules. It says that the rule will apply to the case unless there are other moral rules that apply to that case that render the verdict of the first rule wrong. That is: a rule like "Killing is wrong" applies to a case of killing if and only if there are no other moral rules—for instance, "Killing in self-defense is not wrong"—that apply to the case and render the verdict of the first rule wrong.⁸ But the *unless*-clause doesn't list all of the possible further rules that would defeat the application of the initial rule. That would be impossible if they are incapable of exhaustive statement as Hart supposes. It simply quantifies over them. Likewise for legal rules: the *unless*-clause quantifies over other legal rules.

The crucial thought here is that what makes a rule apply to a case isn't just what obtains; it is also what doesn't obtain. So as well as adding an *unless*-clause to the rules, the full form of a moral or legal argument will also require the addition of a premise to the effect that the *unless*-clause is not triggered. We can make these ideas more precise by defining a notion of what it is for one set of considerations to be *superseded* by a second; that is, for the second set to allude to some further consideration that would upset the conclusion that one would reach on the basis of the first. Then we can construct the *unless*-clause, which I call '*That's it*', in terms of that notion of supersession.

⁷ I here broadly follow the account that I gave in 'Principles and Particularisms' *Proceedings of the Aristotelian Society Supplementary Volume 67*, (2002) pp. 191-209. Readers wanting discussion of some of the difficulties it faces, should consult that article. Note that there I used the term 'principle' for what I am here calling a rule. Following Dworkin I now reserve 'principle' for something couched at a greater level of generality; see below.

⁸ Phrased like this, the rules might look like generics, not properly analyzed as universally quantified generalizations at all. Below I try to say why we need more than this.

I start with a definition of supersession for sentences.⁹ More precisely, the definition is for ordered pairs, the first of which consists of a set of sentences that will stand as the premises of the argument (minus the rule), the second of which consists of a single sentence that will stand as the conclusion (i.e. the verdict):

Supersession for sentences

A set of sentences $\{F_1a, F_2a, \dots, F_m a\}$ and a verdict $F_v a$ are *superseded* by another set of sentences $\{G_1a, G_2a, \dots, G_n a\}$ and a verdict $G_v a$ if and only if:

- (i) $(G_1a \ \& \ G_2a \ \& \ \dots \ \& \ G_n a)$ entails $(F_1a \ \& \ F_2a \ \& \ \dots \ \& \ F_m a)$, but not *vice versa*;
- (ii) $G_v a$ is incompatible with $F_v a$.

The first clause here requires that the second set of set of sentences says everything that is said by the first *and* something more; the second clause requires that the second verdict is incompatible with the first. For instance, the singleton $\{a \text{ is a killing}\}$ and the verdict $'a \text{ is wrong}'$ are superseded by the two-membered set $\{a \text{ is a killing}, 'a \text{ was done in self-defense}'\}$ and the verdict $'a \text{ is not wrong}'$.

Using this we can now define supersession for legal arguments of simple *modus ponens* form:

Supersession for simple arguments

A legal argument¹⁰ of the form

$$\begin{array}{l} F_1a \\ F_2a, \dots \\ F_m a \\ \forall x ((F_{1x} \ \& \ F_{2x} \ \& \ \dots \ \& \ F_{mx}) \rightarrow F_v x) \\ \hline F_v a \end{array}$$

is superseded by a legal argument of the form

$$\begin{array}{l} G_1a \\ G_2a \dots \\ G_n a \\ \forall x ((G_{1x} \ \& \ G_{2x} \ \& \ \dots \ \& \ G_{nx}) \rightarrow G_v x) \\ \hline G_v a \end{array}$$

iff

- (i) the set of non-quantified premises, and the verdict, of the first argument

⁹ Here I have somewhat revised the definitions that I gave in my earlier papers to make them more perspicuous. Those given here should be seen as replacing the earlier ones, which involved some rather awkward features such as entailment relations between predicates. But I don't think that there are radical differences.

¹⁰ I.e. one in which the universally quantified conditional—here $\forall x ((F_{1x} \ \& \ F_{2x} \ \& \ \dots \ \& \ F_{mx}) \rightarrow F_v x)$ —is a legal rule.

are superseded by

- (ii) the set of non-quantified premises, and the verdict, of the second argument.

We can now define the *unless*-clause making use of this idea of supersession:

That's it: There is no sound legal argument that supersedes this argument.

But things aren't quite right, since to know whether an argument is sound we will need to know whether the rule it contains is true, and if what we said earlier is correct, true moral rules must contain *That's it* clauses; and that leads to trouble since we haven't defined a notion of supersession for arguments containing such clauses. So let's do so:

Supersession for That's it arguments

A legal argument of the form

$$\begin{array}{l} F_1a \\ F_2a, \dots \\ F_ma \\ \forall x ((F_1x \& F_2x \& \dots \& F_mx \& \textit{That's it}) \rightarrow F_vx) \\ \textit{That's it} \\ \hline F_va \end{array}$$

is superseded by a legal argument of the form

$$\begin{array}{l} G_1a \\ G_2a \dots \\ G_na \\ \forall x ((G_1x \& G_2x \& \dots \& G_nx \& \textit{That's it}) \rightarrow G_vx) \\ \textit{That's it} \\ \hline G_va \end{array}$$

iff

- (i) the set of non-quantified premises minus the *That's it* clause, and the verdict, of the first argument

are superseded by

- (ii) the set of non-quantified premises minus the *That's it* clause, and the verdict, of the second argument.

This leads to something a little strange: *That's it* is defined in terms of argument supersession; but the relevant notion of argument supersession itself makes reference

to *That's it*. I don't think that this is pernicious; it just means that we need to understand supersession and *That's it* together.¹¹

So, for example, we get legal arguments like this:

- 1
- P₁ A killed another human being
 - P₂ $\forall x ((x \text{ killed another human being} \ \& \ \textit{That's it}) \rightarrow x \text{ is guilty of murder})$
 - P₃ *That's it*
 - C A is guilty of murder

Recall that Hart claimed that the *unless*-clause cannot be exhaustively stated: no matter how many exceptions are given to a rule, one can always imagine further exceptions that have not been captured. Equivalently, no matter how much is built into the content of the rule itself, one can always imagine further factors that will render the rule invalid. In our current framework, this can now be understood as the claim that any legal argument like 1 is bound to be superseded by other valid arguments: take any legal argument, we can always find another that supersedes it. So for instance, we can say that 1 is superseded by the valid argument:

- 2
- P₁ A killed another human being
 - P₂ A killed in self-defense
 - P₂ $\forall x ((x \text{ killed another human being} \ \& \ x \text{ killed in self-defense} \ \& \ \textit{That's it}) \rightarrow x \text{ is not guilty of murder})$
 - P₃ *That's it*
 - C A is not guilty of murder

Similarly 2 would be superseded by an argument that added the claim that the killing was not necessary for self-defense, and that in turn would be superseded by one that added that the defendant didn't realize this to be so, and that by one that he could have known it had he only paid due care, and so on. If Hart is right, no matter how complicated the rule gets, we will always be able to think of an argument that supersedes it. But the fact that every legal argument is superseded by some *valid* argument does not mean that it is superseded by a *sound* argument, that is, by a valid argument that has *true* premises. If the killing was not done in self-defense, and there is equally no other excusing condition, then the *That's it* premise in the original argument 1 will be true, 1 will not be superseded by any sound argument, and the conclusion, that the defendant is guilty of murder, will be simply true. Put another way: the fact that every legal argument would be superseded by a sound argument were certain facts to obtain does not show that every legal argument is in fact soundly superseded. A good legal argument is one that is not.

So the approach meets one of the desiderata with which we started: we have found a way to interpret the *unless*-clause that does not involve a simple list. The approach also meets the second, for clearly it does not lead to triviality. The *That's it* condition is a substantial one. Many real legal arguments go wrong exactly because it is not met: a

¹¹ I discuss this circularity a little further in 'Principles and Particularisms'.

court will reach a mistaken legal conclusion because there is some further relevant factor that it is overlooking.

DISTINGUISHING, OVERTURNING AND OTHER LEGAL ACTIONS

Let's see then how this machinery works to account for how courts can distinguish a case, and how it might be extended to some other phenomena. Sometimes courts will simply overturn previous rulings. This might be an appellate court overturning a lower court's decision; or, less commonly, it might be a court overturning a whole run of precedent. In such cases the overturning court will hold that the legal rule implicit in the earlier rulings was simply mistaken. The account offered here has nothing to say about them.

However, as we have seen, very often a court's attitude to precedent will be more subtle. The court will distinguish the rule that is applicable to the case at hand from the established rule. It is this that the current account is meant to explain. For in judging that the new case should be handled differently, the court implicitly invokes the *That's it* clause. The court claims that there is a sound legal argument that supersedes the argument that would come from applying the existing rule. Indeed, the rule in the superseding argument—the *superseding rule*, as I shall call it—comes into existence as a result of the very judgement that the court is making. If this is a subordinate court, then the judgment, and the superseding rule itself, will be provisional, pending appeal to a higher court: the higher court could find that there is no superseding rule, a possibility that we shall discuss later. But if the decision stands, then the superseding rule will have been established by the action of the court.¹²

I have written as though there is a partition into the cases that involve over-ruling and the cases that involve distinguishing. In fact though things are often messier. For the law is often unclear, and it is often contradictory, so courts will find themselves in the business of making unclarities precise, or resolving contradictions. Nevertheless, even here we can make a distinction between, on the one hand, judgments that precisify or resolve in a way that is tantamount to overturning; and, on the other, judgments that precisify or resolve in a way that is tantamount to distinguishing. The idea is this: sometimes the precisification of a rule is designed to apply to all its applications; but at other times it is limited to those cases where some further distinguishing feature is in play. Likewise for resolution. Sometimes contradictory rules can be resolved so that one of them holds across the board; at other times they can be resolved only for cases that have some further distinguishing feature.

We can accommodate these further cases by extending our definition of supersession. So far we have required superseding sentences to involve a verdict that is straightforwardly incompatible with the verdict of the sentences they supersede. But

¹² This means, I think, that the account is compatible with positivism: we do not need to assume that the superseding rules exist independently of the actions of the courts. I discuss this further in 'The Exception Proves the Rule'. In effect we have here a process of accommodation. For an overview of the process see David Lewis 'Scorekeeping in a Language Game' in his *Philosophical Papers* Vol 1 (New York: Oxford University Press 1983) pp. 233–49. Note though that since a higher court can find that the lower court made a mistake about what the relevant rules are, the lower court must be implicitly quantifying over the set of the rules that the higher court will recognize. I discuss this below (p. 16 of the current manuscript).

the incompatibility might be less absolute than that. We can think of an unclear or a contradictory legal argument as containing not a single conclusion but a class of conclusions. An argument can then supersede it so long as its conclusion contradicts at least one member of that class. So thinking in terms of a verdict set rather than a single verdict, we can redefine supersession for sentences as follows:

Supersession for sentences

A set of sentences $\{F_1a, F_2a, \dots, F_m a\}$ and a verdict set $\{F_{v_1}a, F_{v_2}a, \dots\}$ are *superseded* by another set of sentences $\{G_1a, G_2a, \dots, G_n a\}$ and a verdict $G_v a$ if and only if:

- (i) $(G_1a \ \& \ G_2a \ \& \ \dots \ \& \ G_n a)$ entails $(F_1a \ \& \ F_2a \ \& \ \dots \ \& \ F_m a)$, but not *vice versa*;
- (ii) $G_v a$ is incompatible with one member of $\{F_{v_1}a, F_{v_2}a, \dots\}$

The original definition is the special case where the verdict class contains a single verdict.

So much for my presentation of the account. I now want to turn to compare it with some alternatives.

ALTERNATIVES I: COULD LEGAL RULES BE GENERIC?

English contains a generic construction; or, rather, a family of related generic constructions. If someone says

Chimpanzees have ten toes; or
The Chimpanzee has ten toes; or
A Chimpanzee has ten toes

the claim (assuming that the same claim is made by all these sentences) is true, and it isn't made untrue by the rare cases of chimpanzees born with more, or fewer, than ten toes, or by those who have lost them in accidents. So generic sentences are sentences that admit of exceptions.

There is still much debate over how the semantics of generics works. But there is now a large body of evidence that they shouldn't be seen as universal quantifications at all, but rather as involving some innate faculty of generalization: children understand them long before they understand universally quantified sentences that apply to whole kinds.¹³ Could legal rules be examples of them?

Let us start, once again, by considering moral sentences. It would be odd to use the bare plural, as in the first of the Chimpanzee sentences, and say

Killings are wrong.

¹³ That is, they understand sentences like 'Chimpanzees are fierce' before they understand sentences like 'All chimpanzees are fierce'. For review see Sarah-Jane Leslie 'Generics: Cognition and Acquisition', *Philosophical Review* 117 (2008) pp. 1-47. The restriction to whole kinds is important since the understanding of generics does not seem to precede understanding of sentences like 'All the chimpanzees in this cage are fierce'. I'm grateful to Leslie for discussion here.

But we might naturally use the singular

Killing is wrong.

'Killing' in that case though looks to be a verb and not a noun, since we can modify it by adding a noun:

Killing people is wrong.

But take a case like:

Murder is wrong

'Murder' here is clearly an abstract noun. We cannot add a further noun as we could with 'killing':

*Murder people is wrong

So we plausibly have a generic: not a count-term generic as illustrated by the Chimpanzee sentences, but a mass-term generic, like

Water is liquid at room temperature

Note that 'murder' is already a moralized term: to count something as a murder is already to count it as at least *prima facie* wrong. It is, in other words, a thick moral (and legal) term, one containing both normative and descriptive information. And once we think of thick moral (and legal) terms we find that that examples are legion, whether critical (theft/coercion/bribery is wrong) or complimentary (loyalty/bravery/ kindness is good).¹⁴

We might well appeal to such sentences in justifying legal rules, but I want to argue that they are not themselves legal rules of the kind in which we are interested. Many of what Dworkin terms 'principles' look as though they might have this form.¹⁵ One of his examples can be formulated as what looks like a generic, here with a bare plural:

Wrongdoers should not benefit from their own wrongdoing.¹⁶

Similarly, the principle that is often said to underpin contract law is put very naturally in bare plural form:

¹⁴ Interestingly it is harder to find purely descriptive non-count terms that work in this way. It looks as though we may be readier to use abstract non-count terms for virtues and vices than for actions in general; but if this is so, I do not know why.

¹⁵ Ronald Dworkin *Taking Rights Seriously* (London: Duckworth, 1977) pp. 22ff.

¹⁶ Though he actually formulates it as a universally quantified claim: No man may profit from his own wrong; Dworkin, *op cit* p. 26.

Agreements are to be kept.¹⁷

As I say, these look rather like generics; I am not sure that is right, but let's assume that they are. Still, these are not yet the kinds of rule that we find in the details of common law texts; they are too general and open-ended. As Dworkin stresses, principles can clash. An argument that contained only principles would not culminate in a conclusion for how to decide a particular case: we would need to know whether the principles applied in this case, and that requires something more specific than a principle.

Let us look further at the case of contract law. When we move beyond the general principles to something more specific we find things like this:

A contract is only valid if it is accepted;
but its acceptance does not count if it is not heard by the offerer (for
instance, if it is drowned out by an aeroplane);
but it does count if the offerer did not hear it, but it was clearly made;
but it does not count if it was clearly made but the offerer made it
known to the acceptor that he did not hear it.¹⁸

And this is just one of the necessary conditions for a contract to be valid; stating the necessary and sufficient conditions requires a hefty text book, which even then will leave a great deal out. These are the materials from which legal arguments are actually made. Now they do culminate in a binding conclusion, and they do so because they plausibly have the form, not of generics, but of universally quantified sentences, along the lines of:

$\forall x$ (If x is a valid contract \rightarrow (x is accepted & (the acceptance is heard
by the offerer \vee (the acceptance is clearly made & the offerer does not
make it clear that he does not hear the acceptance))))

Nevertheless, despite the complexity of the condition we can easily imagine circumstances in which the very considerations that prompt the complexity require further complexity: what if the offerer's statement that he did not hear the acceptance was itself drowned out by an airplane, and so on. So even here we need *That's it* clauses once we try to define necessary and sufficient conditions for being a valid contract.

We can think of the development of the common law as the process of arriving at rules from a set of principles—that, I think, is a plausible account of at least one aspect of its history. If we understand principles as generics, we can therefore think of it as a move from generics to universally quantified sentences. If, in a parallel way, we think of moral principles as generics, there is a question of whether we have need of the move to moral rules at all: and therefore of whether a *That's it* account for moral rules (something that I have advocated elsewhere) has any role to play. On the one hand, we

¹⁷ *Pacta sunt servanda*; here again, at least in its use in international relations, the principle is usually qualified with an open-ended *unless*-clause: *clausula rebus sic stantibus* (things remaining as they are).

¹⁸ Denning LJ in [1955] 2 QB 327.

don't have need of rules to play the roles of public justification and of guidance that they play in the common law. On the other, if we stop just with principles then we do not get moral arguments that entail their conclusions. I leave the question open.

ALTERNATIVES II: ARE LEGAL RULES GOVERNED BY NONMONOTONIC LOGIC?

The proposal developed so far uses classical logic; indeed, it does not even involve any new connectives. An alternative, proposed in a set of important articles by John Horty, seeks to achieve a similar end by very different means.¹⁹ A similar end, in that Horty too wants to develop a framework in which legal rules play an important role but can be superseded. Different means, in that he proposes a logical framework that moves a long way from classical logic.

Classical logic is monotonic: adding new premises to a valid argument cannot result in the argument becoming invalid. In contrast a nonmonotonic logic does allow for this possibility. So let me start by spelling out the logic that Horty wants to use, before asking what benefit is gained.

Horty suggests using a logic modeled on Reiter's default logic. Typically this is presented as an interpretation of generic sentences, but here I will consider the idea that it can serve to provide a logical interpretation for legal rules. Reiter's logic was developed as a logic for drawing consequences from a set of premises, given that further premises might serve to undermine those consequences. In classical first order logic understood proof-theoretically, the consequences of a set of sentences are simply the sentences that follow from that set given the rules of inference, rules that operate on the original sentences in virtue of their logical form. Thus if the premise set contains P and $(P \rightarrow Q)$, and *modus ponens* is one of the rules of inference, then Q will follow from the original set. We can think of the *conclusion set* as being the set of sentences that is obtained by collecting all the logical consequences so understood.²⁰

But classical logic is monotonic: adding extra sentences to the premise set cannot result in sentences being deleted from the conclusion set. In moving to a nonmonotonic logic, Reiter takes the premise set to consist not just of the set of sentences as classically conceived, but in addition a set of *default rules*. Although these are often referred to as rules of inference, they are not akin to classical rules of inference like *modus ponens*. Rather they are more like additional premises. Reiter introduces them as something like a three-place conditional, roughly of the form: If A , then B , provided that C is compatible with the conclusion set. He symbolizes this as

¹⁹Reasons as Default's *Philosopher's Imprint* 7 (2007); 'Reasons and Precedent' ms. For discussion of the nonmonotonic logic that lies behind his discussion, see his 'Nonmonotonic Logic', *The Blackwell Guide to Philosophical Logic*, L. Goble (ed.), (Oxford: Blackwell, 2001), pp. 336–361.

²⁰ More standardly in classical logic we would simply call a set closed under logical entailment a *theory*. But I will try to keep the discussion parallel to Reiter's.

$(A: C / B)$.²¹

There is a special case of the default rule—the so called ‘normal rule’—where the role of C is taken by B itself

$(A: B / B)$

We can read this as: if A, then B provided that B is compatible with the conclusion set. Let us write it in abbreviated form, using a special arrow to distinguish it from the material conditional, as

$A \mapsto B$

Following most writers on the topic, from now on we shall just work with normal rules of this form.

We can see now why the approach is nonmonotonic. If our initial set contained just A and $A \mapsto B$, then we would be entitled to add B to the conclusion set. But if we were to add $\sim B$ to the initial set, this making it inconsistent with B, we should no longer be entitled to add B. How do we construct the conclusion set given this approach? The idea, of course is that, given both A and $A \mapsto B$ in our initial set, B should end up in the conclusion set just in case $\sim B$ is not in the conclusion set. But it turns out that constructing that set is no simple matter, since, even if $\sim B$ is absent from the initial set, it might arrive in the conclusion set as a result of other inferences. In consequence there is no simple iterative procedure that will take us from the initial set to the conclusion set. We need instead to determine the conclusion set as a kind of equilibrium set that accommodates all of the information in the initial set. Reiter suggests doing this using a fixed point approach: posit a conclusion set, and then show that it stands as a limit of operations performed on the initial set.²² However, even on this approach, a given initial set will sometimes enable us to arrive at different putative conclusion sets—or *extensions* as he calls them. So how should we determine *the* conclusion set given different consistent extensions? One approach is to simply pick an arbitrary extension as the conclusion set. Another is to say that a sentence will go in the conclusion set just in case it is in *any* extension (though such a set may not be consistent). A third is to say that the conclusion set contains only those sentences that appear in *every* extension (though such a set may be empty).

²¹ In Horty's discussion the default rules are presented in ways that make them seem more like classical inference rules—the idea being that given A, one is committed to inferring B provided that C is compatible with the conclusion set. But they are not logical rules; they are not supposed to apply to a sentence just in virtue of that sentence's logical form. Their application crucially depends upon the interpretation of the non-logical vocabulary. For some discussion of the issues here, and of whether some presentations of nonmonotonic logic muddle the idea of the frame and the interpretation function see Robert Stalnaker, 'What is a nonmonotonic consequence relation?', *Fundamenta Informaticae*, vol. 21, (1994) pp. 7-21. In general Stalnaker argues that it might be helpful to recast nonmonotonic logic, understood in terms of the consequence relation, as a theory of nonmonotonic operators—i.e. operators that do not obey the rule of inference (If $p \models q$ then $Op \models Oq$)—within a monotonic logic.

²² For summary of the fixed point approach see Horty, 'Nonmonotonic Logic'.

Deciding which of these approaches is right is beyond the scope of our enquiry here; I will assume that the issues can be resolved satisfactorily.²³ Let us instead focus on how the default rule approach, with its new nonmonotonic logic, gives us any advantages over the *unless*-clause approach that I have been advocating. Horty argues that the default logic has two things in its favor.²⁴ However, what he has in mind as the alternative is not the *That's it* proposal, but a rather more simple-minded theory that uses a material conditional and then simply lists the possible *unless*-conditions as conditions on the antecedent. So instead of the simple conditional

$$A \rightarrow B$$

it will have the qualified conditional

$$(A \ \& \ \sim C_1 \ \& \ \sim C_2 \ \& \ \sim C_3 \ \dots) \rightarrow B$$

where C_n is the n th *unless*-condition. Call this the *list account*. The differences between the list account that the quantified approach involved in the *That's it* account are important, and we shall return to them shortly. But it will help fix the issues if we first consider this simpler list account, since I am not convinced that the default rule approach brings a real advantage even over it.

The first problem that Horty identifies with the list account is that 'the list of circumstances that might interfere ... is open-ended. No conceivable list of possible interfering circumstances could be complete'²⁵ That is a legitimate worry, and it is the reason that I tried to capture the defeating conditions using quantifiers rather than a list. But does the default rule account escape the worry? Defeating conditions only get into the default approach once they are incorporated into the default theory. One obvious way of incorporating them is to add to the simple default rule $A \mapsto B$ a set of further default rules that capture the unless clauses:

$$\begin{aligned} C_1 &\mapsto \sim B \\ C_2 &\mapsto \sim B \\ C_3 &\mapsto \sim B \dots \end{aligned}$$

If the worry is that no finite list account could list all of the possible defeating conditions, then it is equally true that no finite set of default rules could capture them either. In both cases the theory that we have will only accommodate some of the defeating conditions. We can always extend it by adding more—more clauses in the antecedent for the list account, more default rules for the default account—but the two are on a par.

Let us turn then to the second advantage that Horty claims for the default account. He writes:

²³ In recent work Horty has greatly reduced the indeterminacy here by introducing an ordering on default rules: see 'Defaults with Priorities' *Journal of Philosophical Logic* 36 (2007) 367–413. But since he only requires it to be partial, indeterminacy will remain.

²⁴ Horty 'Nonmonotonic Logic' p. 7

²⁵ *Ibid.*

The second problem is more subtle, and would arise even if we did have a relatively exhaustive list of qualifications. The point of placing preconditions in the antecedent of a [principle] is that we must verify that the preconditions are satisfied before concluding that the [principle applies]... But it seems less reasonable to suppose that we must actually have to verify that all of the various weird circumstances that might interfere with this [principle] do not occur ... It would be better to be able simply to assume that weird circumstances like these do not occur unless there is information to the contrary.²⁶

This was an important consideration in the development of default logic—it is the initial motivation in Reiter’s original article for instance—so we should spend some time on it. Part of the reason for providing default logics was to provide a logic that enabled one to move forwards—to draw conclusions, or to perform actions—in the absence of information. Default logics don’t give us a decision procedure for determining membership of the conclusion set.²⁷ But the idea was that, to take the standard example, on learning that Tweety is a bird, we should be able to defeasibly conclude that Tweety can fly. We do not want to have to hold off from drawing that conclusion while we make sure that Tweety is not a penguin, or an emu, or an ostrich, or has had his wings clipped, or was born deformed, or whatever. But clearly if we had built these conditions into the antecedent of the conditional

If Tweety is a bird, & Tweety is not a penguin & Tweety is not an emu... then Tweety can fly

we would need to establish each of the conjuncts of the antecedent before we could detach the consequent.

In contrast, in the absence of defeating conditions, the default approach will enable us to move forward. We do not need to establish that Tweety is not a penguin, not an emu and so on. Provided that we do not have evidence that he is, the instantiation of the default rule

(x is a bird \mapsto x can fly)

will enable us to (defeasibly) conclude, from the information that Tweety is a bird, that Tweety can fly.

Call this the *no stalling* feature of the default account. Sometimes no stalling will indeed be a feature we want. Typically whether we do so will depend upon the number of exceptions that there are to the rule, the benefits of moving on and the costs of the

²⁶ *Ibid.*

²⁷ Indeed, Reiter’s default logic is even worse off than first order predicate calculus, in that it is not even semi-decidable, i.e. not only is there no procedure for demonstrating that a sentence isn’t in the set, there is also no procedure for showing that a sentence is. See Reiter pp. 104ff.

occasional mistake. In making rough and ready predictions about a creature's flying capacity a no stalling feature might be very useful. Is it something we want in moral and legal reasoning though?

This is a difficult question. In moral reasoning I think that it generally is not. We put high store on getting the conclusion right. By and large, if there are known defeating conditions, we want people to ensure that they are not met before they move to a moral judgment. In the terms of default approach, we ask that they check that none of the known defeating conditions obtain; and that involves adding to the initial evidence set information on whether they do or do not. So it seems to me that the no stalling feature is not much of an advantage in moral reasoning. In fact, it can look to be something of a liability.

In legal reasoning the situation is rather different. Courts and jurisdictions differ, but in the adversarial Anglo-American system, it isn't normally the role of the judge and jury to seek out new information; decision are made on the basis of the (admissible) evidence presented. To that extent then the default account might look to give a good basis for legal reasoning. For the court needs to not stall: it needs to make a decision on the basis of the evidence it has.

Even here though the advantages of the default account over the list account should not be overplayed. Courts have a complex set of default assumptions stemming from the idea of the burden of proof. They are allowed to *assume* many things unless proved otherwise, and perhaps this would allow them to complete the list: items in the list could be assumed to hold unless they were shown not to. But I shall not pursue the question of whether this will work. For it strikes me that both the list approach and the default approach share a fundamental flaw, a flaw that can only be avoided by moving to a quantified approach like the one embodied in the *That's it* account.

So far we have been concerned with *recognized* defeating conditions: with clauses that are contained in the list or embedded in the default rules. When it comes to moral and legal reasoning though, we ask for more than that. We ask that agents be sensitive to considerations that are relevant to the judgments they make but that have not been previously recognized. Admittedly when it comes to moral reasoning, this is to place the bar very high: it is only a somewhat idealized moral judge who will recognize the significance of factors that have not been recognized before. But in the case of legal reasoning the requirement is far more realistic. It is quite common for a case to come before a court that involves factors that have not previously been dealt with by the law. And in such a case, as we said before, the court will have either to reaffirm the existing rule, thereby in effect judging that the new factors are not sufficient to warrant a different verdict; or it will have to distinguish.

Let us take the case where the court reaffirms the existing rule. It is important to note that here, just as much as in the cases where they distinguish, courts are open to criticism. Most obviously, if the case goes to appeal, the decision may be rightly overturned.²⁸ What are we to make of such a happening? On either the list approach or the default approach, it looks as if the court did nothing wrong: it applied the existing rules, and arrived at the conclusion. The obvious response is that applying such

²⁸ I leave open the question of whether the lower court can be wrong in their decision even if it is not overturned. Obviously we might think that they are morally wrong, but can they be legally wrong? This opens difficult questions about the strength of positivism on which the present approach can remain neutral.

rules is not good enough. But if that is so, in virtue of what is it not good enough? It seems that the flaw must stem from some further feature. Neither the list account or the default account has the resources to account for it.

In contrast, on the account I am offering, it is clear where the mistake lies. In endorsing a legal argument, the court is implicitly endorsing the *That's it* condition. It is committed to the claim that the argument is not superseded, and so that there are no further relevant facts and rules. And in the circumstances imagined, where the verdict is subsequently overturned, that was not true.

I say that this is clear, but if it is right it does bring out a further feature of the account. For if the lower court is making a mistake in not distinguishing, then the class of rules over which it quantifies cannot just consist in the previously acknowledged rules, and whatever further rule it recognizes. By hypothesis, there is no superseding rule there. We have to say that the domain of quantification includes the previously acknowledged rules, and the rules that the court should have recognized. I don't think that this requires us to give up positivism, but it does mean that the court sees itself as bound by requirements that go beyond those of the extant rules.

A parallel issue arises when the court does distinguish, and where on appeal the move is rejected. Here again, we want to know what it was that the court did wrong. And again neither the list account nor the default approach have anything to say on the matter, whereas the approach that I have offered does: the court held that the *That's it* clause in the argument using the original rule was false, whereas it was not.

How might the proponent of the default account respond? One way is just to reject the idea that the lower court has made a mistake; one might say instead that it has simply lost out to a more powerful player. But short of embracing legal realism, it is hard to maintain that that is true in all such cases. Another response is to say that the mistake is a mistake of prediction: the lower court predicted that the appellate court would support its judgment. But that doesn't seem right either: the lower court is making a ruling of law, not trying to second guess what the appellate court will say. More plausibly the proponent of the default account might insist that all this can be explained by a further requirement to which they are perfectly well entitled: a requirement on courts to come up with the right default rules. Courts make mistakes when they fail to conform to that requirement. I shan't try to guess exactly what such a proposal would look like, but a natural idea would be that the court will be committed to the claim that there is no further default rule that, added to the existing rules, would overturn the verdict that it reached. But what we have now is looking very much like the *That's it* clause. And if we add it to the default account, we have to ask whether it cannot do the work that the default rules are doing. The default account is in danger of collapsing into the *That's it* account, in which case the use of nonmonotonic logic looks redundant.

The proponent of the default account might take another tack, denying that we can ever know whether the *That's it* condition is met. The issue would gain importance if we agreed with Horty's implicit requirement that we must *verify* the premises in an argument. The *That's it* condition is equivalent to a universally quantified sentence, and Karl Popper built a hugely influential theory on the idea that such sentences can *never* be verified. We need not follow him in that to think that we are in no position to verify that the *That's it* condition is met. But what I think we should conclude is that

verification is far too high a standard for the premises in legal argument. As with much else in law, we might want them to be beyond reasonable doubt. However, we want to concede that judges need to act in the awareness that they might be making mistakes; and this strikes me as placing the fallibility in just the right place.

In responding to Horty's arguments I have made no use of the familiar idea that classical logic is well behaved and well understood, whereas nonmonotonic logic is not. I do think that that is true; and so if two theories did the job equally well, I think that we would still have grounds to prefer a theory that used classical logic over one that did not. But if I am right that the default theory does not do the job equally well, I have had no need of the argument.

CONCLUSION

In relation to its rivals, the *That's it* approach stands up surprisingly well. It does require us to posit implicit claims, but they are far from outlandish. Perhaps it has little role to play in moral argument: that will depend how well a generic account can be developed. But in law, where we do have need of tight, precedence-justifying arguments that nonetheless admit of exceptions, I know of no better approach.²⁹

²⁹ Thanks to Irene Heim, Jeff Horty, Sarah-Jane Leslie for discussion of some of these issues; and to the participants at the USC law and language seminar where the paper was presented, especially Andrei Marmor and Scott Soames.