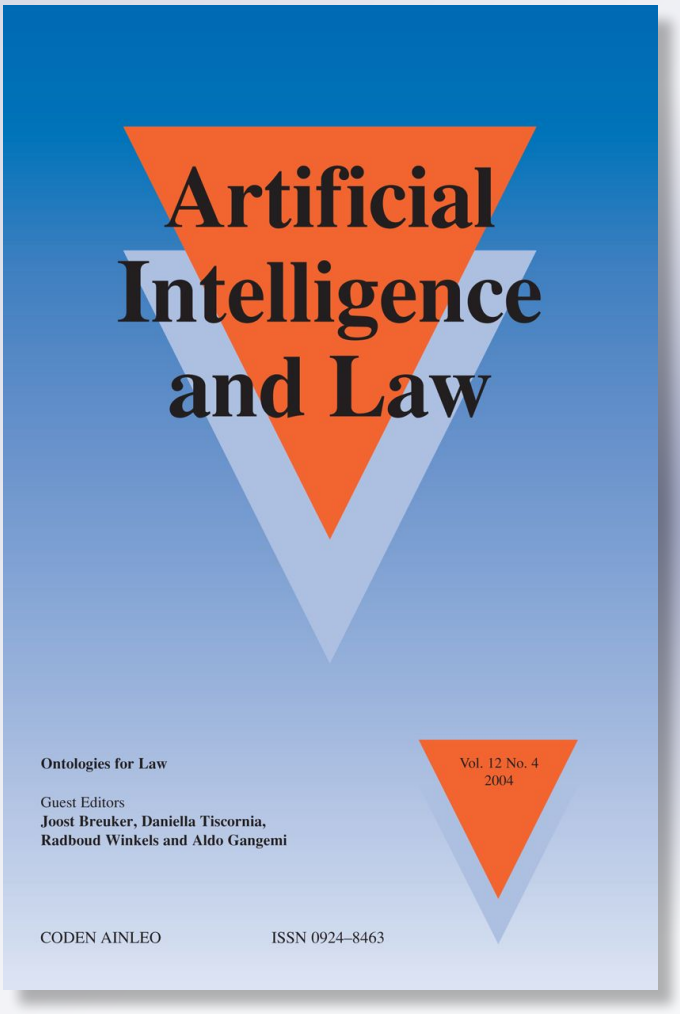
Moral particularism in the light of deontic logic

Xavier Parent

Artificial Intelligence and Law

ISSN 0924-8463

Artif Intell Law
DOI 10.1007/
s10506-011-9108-1



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Author's personal copy

Artif Intell Law
DOI 10.1007/s10506-011-9108-1

Moral particularism in the light of deontic logic

Xavier Parent

© Springer Science+Business Media B.V. 2011

Abstract The aim of this paper is to strengthen the point made by Horty about the relationship between reason holism and moral particularism. In the literature prima facie obligations have been considered as the only source of reason holism. I strengthen Horty's point in two ways. First, I show that contrary-to-duties provide another independent support for reason holism. Next I outline a formal theory that is able to capture these two sources of holism. While in simple settings the proposed account coincides with Horty's one, this is not true in more complicated or "realistic" settings in which more than two norms collide. My chosen formalism is so-called input/output logic. A bottom-line example is introduced. It raises the issue of whether the conventional wisdom is right in assuming that normative reasons run parallel to epistemic ones.

Keywords Moral particularism · Prima facie obligation · Conflict resolution · Priorities · Contrary-to-duties · Default reason · Input/output logic

1 Introduction

In this paper I present a variant to Horty's theory for reasoning with prioritized (deontic or epistemic) defaults presented in Horty (2007a, b). There the framework is used to shed light on the on-going debate on so-called moral particularism, launched in moral philosophy by the British philosopher Dancy (1993, 2006). Neither Dancy nor Horty see the connection with contrary-to-duties and norm violations. I will argue that these are highly relevant to the latter debate, and

X. Parent (✉)
Individual and Collective Reasoning (ICR) Group, Department of Computer Science and Communications (CSC), Faculty of Sciences, Technology and Communication (FSTC), University of Luxembourg, 6, rue Richard Coudenhove—Kalergi, 1359 Luxembourg City, Luxembourg
e-mail: xavier.parent@uni.lu

Published online: 04 August 2011

Author's personal copy

investigate how Horty's framework copes with them. Next I will present an alternative way to handle prioritized defaults in the setting of input-output logic (IOL), and show that while in simple settings both accounts coincide, this is not true in more complicated or "realistic" settings in which more than two norms collide. The bottom-line is whether normative reasons (reasons for action) run in parallel to epistemic ones (reasons for belief).

2 Moral particularism

Not long ago there was a debate running on Brian Leiter's Legal Philosophy blog[1] about why there is so little interest in deontic logic (and in the formal semantics of normative concepts) on the part of other academic disciplines, like moral philosophy. To get them more interested in deontic logic, we need to show them how the latter one can contribute to the issues they are discussing. The paper (Horty 2007b) by Horty can be seen as an attempt to do it. There the focus is on so-called moral particularism, a justly popular "cutting-edge" topic in contemporary ethics that has been most notably defended by the British moral philosopher Dancy (1993, 2006).[2] Moral particularism—or situationism, as it has sometimes been called—seems to present an especially radical objection to the enterprise of moral theory, and hence of deontic logic. To put it simply, particularism is the view that there are no moral principles, only moral intuitions. In his paper Horty examines and criticizes the main argument Dancy gives in support of his view.

First, Dancy's argument. Particularism is a negative thesis on how moral reasoning works. Dancy calls "generalism" the view that "the very possibility of moral thought and judgement depends on the provision of a suitable supply of moral principles" (Dancy 2006, p. 73). This is the traditional approach in moral theory. Particularism can be viewed as the negation of the latter view. It is the claim that "the possibility of moral thought and judgement does not depend on the provision of a suitable supply of moral principles" (Dancy 2006, p. 73). There is, then, the obvious issue of explaining what moral reasoning is. Particularists do not say much about it, and here no attempt will be made to discuss this issue.[3]

In this paper I shall focus on another point of controversy. It is that Dancy grounds his claim on a form of reason holism, holding that what is a reason in one case need not be any reason in another: "Reasons are like rats, at least to the extent that two rats that are supposedly on the same side may in fact turn and fight among themselves" (Dancy 2006, p. 15). The main support he provides in support of holism in the theory of reasons comes in the form of examples, in which the notion of *prima facie* obligation takes centre stage. The latter notion was introduced in moral philosophy by Ross (1930). A prima facie obligation is one that binds unless

[1] http://leiterlegalphilosophy.typepad.com/leiter/2007/10/green-v-shapiro.html.

[2] Dancy's views were historically developed in response to Hare.

[3] For an attempt to shed light on this matter using developments in statistical learning theory, see Harman (2005).

Author's personal copy

overridden by another stronger obligation, and so it is defeasible: it leaves room for exceptions.

It is natural to ask if moral particularism is any different from moral relativism, which is sometimes said to be "self-defeating".[4] Before explaining why I think the worry is misconceived, I would like to make a clarification. In fact, Dancy emphasizes that holism about reasons is consistent with the existence of what he calls "invariant" reasons—considerations that function as reasons with the same valence in all circumstances. An example is: causing gratuitous pain to an unwilling victim. Holism about reasons only maintains that there is nothing about being a reason that precludes the context-sensitivity of reasons. It is compatible with this that certain kinds of reasons are not context-sensitive. Such reasons are invariant not because they are reasons but because of their specific content.[5]

Here, I will put the above subtlety to one side, and I will assume that the particularist thesis is best formulated as the denial of moral principles. It is under this formulation that it is more likely to be vulnerable to the charge of incoherence. Relativism says that truth or knowledge is always relative to some parameter (historical circumstances, culture, conceptual framework, etc.). The usual objection to relativism, as first developed by Plato against Protagoras in the *Theaetetus* (see Cooper and John 1997), involves applying relativism to itself, in order to undermine it. In Plato's days such an objection used to be described as a *peritrope* (a turning of the tables).[6] In the case of particularism, the principle that is turned against itself is "moral principles do not exist". The position becomes self-refuting in the moment it is used as a reason to motivate moral maxims such as "don't act in accordance to any principle", which is itself a principle. Husserl (1988, A) acknowledges the conceivability of using moral relativism this way.[7] I think such a move is understandable, but just wrong. It amounts to conflating two kinds of principles that are, in nature, very different. Following Bennet (1995), Mackie (1977, p. 16, 106) distinguishes between first-order and second-order moral views. A first-order moral view is a claim about what is good or bad. A second-order moral view is a claim about the nature of first-order moral views (e.g., what it is for something to be good or bad). First-order and second-order views address separate and independent matters. The particularist views his original charge only as a second-order claim, without committing himself to any first-order moral view. Turned into a first-order moral view, particularism would be closed to moral nihilism, that is the view that anything is morally permissible. According to Mackie, the idea that there is a logical connection between (second-order) relativism and (first-order) moral nihilism rests on a mistake. I think the same can be said of particularism. Indeed, the relativist is

[4] This objection was raised by a reviewer for DEON'10.

[5] Cf. Dancy (2006, p. 77). The assumption that reasons are general as part of their meaning plays an important role in Hare's analysis of the universalizability of ought-statements. Asking why the word "ought" has the property of universalizability, Hare gives the following reason among others: an "ought"-judgment must be supported by a reason; and this implies that the same judgment must be made whenever the same reason holds (cf. Hare 1952, p. 176, 1963, p. 21).

[6] A good exposition and judicious appraisal of the arguments by which the Ancients sought to refute the assumptions they took to underpin Protagorean relativism can be found in Lee (2005).

[7] This was pointed out to me by one referee.

Author's personal copy

pluralist. He believes that on any moral question there can be more than one correct answer. The particularist still maintains that each particular action is either objectively obligatory or not, depending on the precise details of the situation in question.

Here I find it convenient to use the by now familiar type-token distinction. A type is a concept, and its tokens are the particular objects that instantiate it. For example, the particular bicycle in your garage is a token of the type of thing known as "The bicycle". The bicycle in your garage is in a particular place at a particular time. This is not true of "the bicycle" as used in the sentence "The bicycle has become more popular recently". A particularist denies that there is one "correct" moral answer in any situation viewed as a type. But he admits that there can be one correct answer only in a given token situation (e.g. this lying, *hic et nunc*). What the particularist denies is that an answer, which appears to be right in a given token situation, can be codified in the form of a principle, which will be applied to the same situation again and again.

With these preliminaries out of the way, I can get back to Horty's paper. What he does in there can be described as an attack on attack. He builds a formal theory of reasons, which challenges the connection made between reason holism and particularism. The proposed framework supports reason holism: a reason need not retain its supporting value across contexts. Yet, it is based on a system of principles. These are thought of as defeasible generalizations. So the framework provides a counter-example to the claim that reason holism implies particularism.

I agree with the main point made by Horty. However what is missing from his discussion is the notion of contrary-to-duty (CTD) obligation—a highly problematic topic in deontic logic. Neither Dancy nor Horty see the relevance of CTDs to the topic under discussion. But let us have a look at the kind of examples used by the former to justify reason holism:[8]

The book

Case 1 My borrowing a book from you is a reason to return it to you
Case 2 My borrowing a book from you is not a reason to return it to you, since you've stolen it from the library

This example has the form:

Case 1 *A* is a reason for *B*
Case 2 *A* is not a reason for *B*, since *C*

The example originally used by Chisholm to make his point about contrary-to-duties (see Chisholm 1963) can also be given this form. The example consists of the following sentences: (i) it ought to be that I go to the assistance of my neighbours (ii) it ought to be that if I go I tell them I am coming (iii) it ought to be that if I do not go I do not tell them I am coming (iv) I do not go. From (i)–(iv), we can extract:

[8] This example can be found in Dancy (1993, p. 60).

Author's personal copy

The neighbours

Case 1 My being obliged to go is a reason for telling
Case 2 My being obliged to go is not a reason for telling, since it turns out that I do not go.

In a violation context, the primary obligation to go to the assistance remains in force. Is it to say that it remains a reason for telling? Obviously, not. For (iii) provides (iv) as a (stronger) reason not to tell.

The above considerations show that some CTD scenarios give another independent support for reason holism. Horty's framework was not devised to deal with the latter scenarios. In the next section, I investigate how the theory copes with them. The hope is to show that a stronger statement about the relationship between reason holism and particularism can be made by showing that the framework can deal with both sources of holism.

The idea is to analyze contrary-to-duty reasoning as one that is defeasible. A reasoning is defeasible if is based on reasons that may be defeated in the light of further information. Originality is not my claim. This kind of approach to CTDs has a long tradition which begun with Nute and Donald (1997). Still, I am not aware of a theory (from that tradition) that takes reasons as the basic normative concept.

3 Horty's framework

First, I give an outline of Horty's framework, focusing for the purposes of the present analysis on the special case where all priority relations among defaults are fixed in advance. I will incorporate the amendments to the framework proposed in Horty (2007a).

Reiter's default logic (Reiter 1980) is one of the best known and most widely used formalisms in the Artificial Intelligence (AI) research community. To make his point, Horty could have tried to use one of the prioritized variants of it that have been described in the literature (see, e.g., Brewka and Eiter 2000). However, none of them seem to be very suitable for the specific purpose Horty has in mind, primarily because particularists speak in terms of reasons. Following a suggestion of Raz (1999), they take reasons as the basic normative concept. Horty's account can be viewed as an attempt to recast Reiter's default logic into particularists' language, so they can relate to it better.

An ordered default theory is notated $(W, D, <)$. W is a set of propositional formulae. They represent the knowledge the agent has about the current situation. D a set of defaults subject to a strict partial ordering $<$. Defaults are of the form $A \rightarrow B$, where $\rightarrow$ stands for material implication. They can be given both an epistemic and a deontic reading. In the first case, the conclusion or head B denotes a fact, and $A \rightarrow B$ expresses a reason to believe. In the second case, B denotes the content of an obligation, and $A \rightarrow B$ expresses a reason for acting. $\delta < \delta'$ means that δ' has a higher priority than δ.

The representation of norms as "A is a reason for B" may make some readers feel uncomfortable. At first sight this kind of representation does not allow us to capture a number of distinctions that have proved to be useful, like the distinction between

Author's personal copy

so-called regulative and constitutive norms. The first describe obligations, prohibitions and permissions. The second make possible basic "institutional" actions such as the making of contracts, the issuing of fines, the decreeing of divorces. Basically they tell us what counts as what for a given institution.[9] The framework described here was not devised to deal with constitutive norms. Still, room can be made for them if we accept to interpret constitutive norms as a type of epistemic reasons. On this account, a proposal followed by an acceptance is an epistemic reason to consider the contract as binding. In this respect, a proposal followed by an acceptance counts as a contract. By contrast, the parties promising to make a contract is a normative reason for them to make the contract. Both norms have the same conclusion "contract". The constitutive norm describes it as a fact, which happens to be institutional. The regulative norm describes it as a cue to action.[10]

Before discussing Horty's framework, I need to introduce some further notation. If δ is the default $A \rightarrow B$, for example, then $Premise(\delta)$ is the statement A and $Conclusion(\delta)$ is the statement B. Where D is a set of defaults, $Conclusion(D)$ is a shorthand for $\{Conclusion(\delta) : \delta \in D\}$. $\vdash$ denotes classical consequence. Where X is a set of formulae, $Cn(X)$ denotes the closure of X under $\vdash$.

Like in Reiter's default logic, the goal is to determine the extension(s) associated with a default theory $(W, D, <)$. Intuitively, an extension gathers all the agent's obligations that follow from what it knows about the world. However, the notion of extension is here not as central as it is in Reiter's theory. The key concept is that of (as Horty calls it)"proper scenario(s)" (Horty 2007b, p. 8) based on a default theory. A proper scenario is notated S, and is always a subset of the set of all the defaults D. Intuitively, the defaults in a proper scenario tell us what counts as a binding (good, satisfactory, etc.) reason for what. Thus, if $A \rightarrow B$ is in the proper scenario S based on a given default theory, then S is said to provide A as a reason for B. The idea is to assume that the agent derives its obligations (part of the extension) from justifications or reasons for those obligations: in particular, that the agent is bounded by an obligation if it possesses a satisfactory reason for that obligation.

Formally, the notion of proper scenario is defined using three other notions. Each corresponds to a condition that a default must meet in order to be binding.

The first notion is that of a default being triggered in scenario S. Let us use the notation $Triggered_{(W,D,<)}(S)$ to refer to the set of all such defaults. The definition runs as follows:

$$Triggered_{(W,D,<)}(S) = \{\delta \in D : W \cup Conclusion(S) \vdash Premise(\delta)\}$$

The second notion is that of a default being conflicted in S. Let $Conflicted_{(W,D,<)}(S)$ denote the set of all such defaults. The definition reads:

[9] The idea of constitutive rule is a theme that goes back to Searle (1964, 1969), and before that to Rawls (1955). For a comprehensive overview of the state of the art about logics for constitutive norms, see Grossi and Jones (2011).

[10] The promise example is for illustrative purposes only. Here I take the principle of promise-keeping to be a regulative norm, as does Ross in his theory of prima facie duties. He sees the principle of promise-keeping as falling under (as he calls it) the "duty of fidelity". It is common in the philosophy of language to think of such a rule as constitutive of the act of promising. This view, which was championed by Searle among others, is not germane to my present purpose.

Author's personal copy

$$Conflicted_{(W,D,<)}(S) = \{\delta \in D : W \cup Conclusion(S) \vdash \neg Conclusion(\delta)\}$$

The third notion is that of a default being defeated in S. Let $Defeated_{(W,D,<)}(S)$ denote the set of all such defaults. For $D, D' \subseteq S$, put $D < D'$ if $\delta < \delta'$ for all δ in D and δ' in D'. Let $S^{S'/D'}$ denote the result of replacing S' by D' in S, viz. $S^{S'/D'} = (S - S') \cup D'$. The definition of defeat reads:

$$\begin{aligned} Defeated_{(W,D,<)}(S) = \big\{\delta \in D : \exists D' \subseteq Triggered_{(W,D,<)}(S)\ s.t.\\ a)\ \{\delta\} < D'\ and\\ b)\ \exists S' \subseteq S\ with\ S' < D'\ such\ that\\ W \cup Conclusion(S^{S'/D'})\ is\ consistent\\ W \cup Conclusion(S^{S'/D'}) \vdash \neg Conclusion(\delta)\big\} \end{aligned}$$

These three concepts are used to define the notion of a default being binding in S. Let $Binding_{(W,D,<)}(S)$ denote the set of all such defaults. The definition reads:

$$\begin{aligned} Binding_{(W,D,<)}(S) = \{\delta \in D : \delta \in Triggered_{(W,D,<)}(S)\\ \delta \notin Conflicted_{(W,D,<)}(S)\\ \delta \notin Defeated_{(W,D,<)}(S)\} \end{aligned}$$

The notion of proper scenario is defined using what is sometimes called a "quasi-induction" or, in the terminology of Makinson (1994), an "end-regulated induction". This looks like an ordinary induction except that the induction step makes reference to the object to be defined. While end-regulated inductions look like genuine inductions, logically they behave like fixpoint definitions.

Definition 1 (*Proper scenario*) Let S be a scenario based on the ordered default theory $(W, D, <)$. Then S is a proper scenario based on $(W, D, <)$ just in case $S = \bigcup_{i \geq 0} S_i$ where

$$\begin{aligned} S_0 &= \emptyset\\ S_i &= \{\delta \in D : \delta \in Triggered_{(W,D,<)}(S_{i-1}), \delta \notin Conflicted_{(W,D,<)}(S),\\ &\qquad \delta \notin Defeated_{(W,D,<)}(S)\} \end{aligned}$$

Definition 1 exemplifies an approach to handling inconsistency that is familiar from the literature on so-called non-monotonic reasoning. I shall refer to it as the "increment idea". The key to the proposal is to restrict the step-by-step application of defaults in order to guard against possible contradictions. The agent begins its reasoning process, at the initial stage S_0, without believing in any defaults. Then, at each successive stage S_i, it supplements its stock of defaults with those that have been triggered at the previous stage S_{i-1} as long as they are neither conflicted nor defeated. Note that, at each stage S_i, the constraining scenario against which the agent checks defaults for conflict or defeat is S itself.

Author's personal copy

With these concepts in place, it is a straightforward matter to define the notion of extension. Intuitively, the extension of a default theory can be viewed as the set of obligations generated by a proper scenario. The following formal characterization suggests itself—it is taken from Horty (2007a).

Definition 2 (*Extension*) Let $(W, D, <)$ be an ordered default theory and $\mathcal{E}$ a set of formulae. Then $\mathcal{E}$ is an extension of $(W, D, <)$ just in case $\mathcal{E} = Cn(W \cup S)$, where S is a proper scenario based on this default theory.

It seems that the intuitions behind Horty's framework are fundamentally different from those behind the usual non-monotonic systems for legal argumentation, like the system of Prakken and Sartor (1997). Such systems separate the process of argument construction from that of argument evaluation. By contrast, Horty's framework interleaves the construction and evaluation of arguments. Furthermore, in such systems acceptability in the face of iterated undermining and support is decided using the so-called principle of reinstatement. It is the idea that an argument should count as acceptable even if it is defeated, as long as all the arguments defeating it are themselves defeated. In Horty (2001), Horty gives a counter-example (known as the "Microsoft scenario") to such a principle. We will see that his own framework gets the example right.

Below I apply the construction to the two examples from Sect. 2 used to justify reason holism, and show that the construction yields the correct outcomes.

Example 1 (The book) Let b, s, y and l represent the respective propositions that I borrowed the book from you, that you stole it from the library, that I return the book to you, and that I return it to the library. Put $D = \{b \rightarrow y, s \rightarrow l\}$ with $s \rightarrow l > b \rightarrow y$. Assume $W = \{y \rightarrow \neg l\}$—I cannot simultaneously return the book to you and the library.

Case 1 Suppose that W contains, in addition, the sole formula b. In this case, the ordered default theory $(W, D, <)$ yields $S_1 = \{b \rightarrow y\}$ as its unique proper scenario, providing b as a reason for y.

Case 2 Suppose that W contains, in addition, the two formulae b and s. In this case, the ordered default theory $(W, D, <)$ yields $S_2 = \{s \rightarrow l\}$ as its unique proper scenario, and thereby provides s as a reason for l, rather than providing b as a reason for y.

Example 2 (The neighbours) Let a and t represent the respective propositions that I go to the assistance of my neighbours, and that I tell them I am coming. Put $D = \{\top \rightarrow a, a \rightarrow t, \neg a \rightarrow \neg t\}$ with $<$ the empty relation. Note that in the case of an empty priority relation the notion of defeat plays no role, because $Defeated_{(W,D,\emptyset)}(S) = \emptyset$

Case 1 Suppose that W contains no formula, $W = \emptyset$. In this case, the ordered default theory $(W, D, <)$ yields $S_3 = \{\top \rightarrow a, a \rightarrow t\}$ as its unique proper scenario, providing a as a reason for t.

Case 2 Suppose that W contains the formula $\neg a$. In this case, the ordered default theory $(W, D, <)$ yields $S_4 = \{\neg a \rightarrow \neg t\}$ as its unique proper scenario, and thereby provides $\neg a$ as a reason for $\neg t$, rather than a as a reason for t.

Author's personal copy

Multiple level of violation may be analyzed in the same way.

Example 3 (Multiple level of violation) Put $D = \{\top \rightarrow \neg a, a \rightarrow x, a \wedge \neg x \rightarrow y\}$, where a is for breaking one's promise, x is for apologizing, and y is for being ashamed. Let $<$ be the empty relation.

Case 1 Assume $W = \emptyset$. In this case, the ordered default theory $(W, D, <)$ yields $S_5 = \{\top \rightarrow \neg a\}$ as its unique proper scenario.

Case 2 Assume $W = \{a\}$. In this case, the ordered default theory $(W, D, <)$ yields $S_6 = \{a \rightarrow x\}$ as its unique proper scenario.

Case 3 Assume $W = \{a, \neg x\}$. The solution is $S_7 = \{a \wedge \neg x \rightarrow y\}$.

These examples are simple cases of reason holism. Horty's theory appears to handle them relatively well. In the next section I will argue that the theory is not adequate for a more complex case of reason holism. Before showing it, I need to introduce an alternative framework based on so-called constrained input/output logic (IOL), which will allow me to handle this more complex case of reason holism in—I believe—a more satisfactory way.

4 Constrained IOL with priorities

For clarity's sake, I proceed in three steps. I first give an outline of unconstrained IOL, on top of which constrained IOL is built. Next, I introduce constraints. Finally I add priorities to the framework.[11]

To keep things simple, I only consider obligations—permissions are put to one side. In input/output logic, a normative code is a set G of conditional norms, which is a set of ordered pairs (a, x). Here a and x are two formulae of propositional logic. Each such pair will be referred to as a generator. The body a is thought of as an input, representing some condition or situation, and the head x is thought of as an output, representing what the norm tells us to be obligatory in that situation.

Some notation. L is the set of all formulae of propositional logic. Given an input $A \subset L$, and a set of generators G, $G(A)$ denotes the image of G under A, i.e., $G(A) = \{x : (a, x) \in G \text{ for some } a \in A\}$.

Definition 3 (*Output operations*) Let A be an input set, and let G be a set of generators. The following input/output operations can be defined, where a complete set is one that is either maximal consistent[12] or equal to L:

$$out_1(G, A) = Cn(G(Cn(A)))$$
$$out_2(G, A) = \cap\{Cn(G(V)) : A \subseteq V, V \text{ complete}\}$$
$$out_3(G, A) = \cap\{Cn(G(B)) : A \subseteq B \supseteq Cn(B) \supseteq G(B)\}$$
$$out_4(G, A) = \cap\{Cn(G(V)) : A \subseteq V \supseteq G(V), V \text{ complete}\}$$

[11] My exposition of (unprioritized) IOL is confined to what is essential in the present context. For a more comprehensive account, the reader is referred to Makinson and van der Torre (2000, 2001).

[12] The set is consistent, and none of its proper extensions is consistent.

Author's personal copy

$out_1(G, A)$, $out_2(G, A)$, $out_3(G, A)$ and $out_4(G, A)$ are called *simple-minded* output, *basic* output, *reusable simple-minded* output and *reusable basic* output, respectively.

When analyzing examples, I will occasionally choose to instantiate *out* into either out_3 or out_4. This is because these two operations satisfy plain transitivity ("from (a, x) and (x, y) infer (a, y)"). With such a property it is much easier to get an intuitive feeling of what is going on.

It will help to recast Horty's theory into the Input/Output (I/O) idiom in order to facilitate comparison between the two accounts. In fact, the definition of the key concepts in Horty's approach can easily be adapted to the notational conventions used in I/O logic. His W corresponds to our A. The set of all defaults D corresponds to our G. We write (a, x) where Horty uses its materialization $a \rightarrow x$. The definition of a default being triggered in scenario $S \subseteq G$ becomes the following, where $h(S)$ is the set of heads y of elements of S:

$$Triggered_{(A,G,<)}(S) = \{(a,x) \in G : A \cup h(S) \vdash a\}$$

Some might wonder if $\vdash$ can be replaced with some appropriately chosen output operation. The answer is yes, but I do not think that in the present context this would yield any useful insight. For instance, Stolpe (2010) points out a limiting case where out_1 collapses into classical logic. It is the case where the set of generators is the diagonal relation over L.[13] We get a reformulation along the lines:

$$Triggered_{(A,G,<)}(S) = \{(a,x) \in G : a \in out_1(L \times L, A \cup h(S))\}$$

For the other notions, the reformulation is similar. The notion of extension requires one more small change. We need to put: $\mathcal{E}$ is an extension if $\mathcal{E} = Cn(A \cup m(S))$, where S is a proper scenario, and $m(S)$ is the materialization of S, i.e. the set of all formulae $a \rightarrow x$ where $(a, x) \in S$. Alternatively, $\mathcal{E}$ is an extension if $\mathcal{E} = out_4^+(S, A)$.[14]

Now we can turn to IOL with constraints. The motivation is best explained by considering the case of a conflict between two obligations. Put $G = \{(a,x), (a,\neg x)\}$, and consider input a. The reader may easily verify that, for all output operations, $out(G, a) = L$. This shows that none of the operations considered so far are conflict-tolerant. The idea is to cut back the set of norms to just below the threshold of yielding excess, and consider the resulting output. To do that, we look at the maximal non-excessive subsets, i.e. the maximal $G' \subseteq G$ such that $out(G', A)$ is consistent. The family of all such G' may be called the maxfamily of A, and the family of outputs $out(G', A)$ for G' in the maxfamily, may be called the outfamily of A. Below this maximal subsets strategy is referred to as the "threshold idea". It gives IOL its distinctive flavour compared to other non-monotonic frameworks. The main idea is to perform on G the minimal changes required in order to avoid contradictions (or any other consequences specified as

[13] This is (Stolpe 2010, theorem 2).

[14] Cf. Makinson and van der Torre (2000, observation 16). $out_i^+ (i = 1, \ldots, 4)$ is the throughput version of out_i that allows inputs to reappear as outputs. In terms of the semantical definitions, we have $out_i^+(G, A) = out_i(G \cup I, A)$, where $I = \{(y, y)$: *y is a formula*$\}$.

Author's personal copy

undesirable). This is slightly different from restraining the step-by-step application of rules. Formulated in such general terms, this may seem a rather nebulous and inconsequential distinction to make. In fact, this makes a real difference, as we will see in due course.[15]

The example I have just used to motivate the approach involves what might be called a "strict" or logical conflict: the head of one rule contradicts the head of the other. The point that has been made about such conflicts also applies to "natural" (or non-strict) conflicts. These are of the form: (a, x) and (a, y), with x incompatible with y in the sense of natural or physical necessity, broadly conceived. The treatment of this type of conflict necessitates the use of a set C of "integrity constraints" (as they are sometimes called) that the output is never allowed to contradict. In this case, $C = \{x \rightarrow \neg y\}$. This is meant to indicate that x and y cannot be simultaneously true, given the agent's present physical and psychical capabilities, etc. Compared to the treatment of a strict conflict, the main difference is that we take the maximal $G' \subseteq G$ such that $out(G', A)$ is consistent with C. Example 4 below provides an illustration.

For contrary-to-duties, the idea is similar. Typically we take the maximal $G' \subseteq G$ such that $out(G', A)$ is consistent with A. It is easy to check that not doing so would create the same problems as those encountered in Standard Deontic Logic (SDL). This is illustrated clearly by example 6.

The formal definition below is general, covering as special case both inconsistency of output and its inconsistency with input.

Definition 4 (*Threshold*) Let G be a set of generators and *out* be an input/output operation. Let C be an arbitrary set of formulae, which we may call "consistency constraints". We define:

- $maxfamily(G, A, C)$ is the set of $\subseteq$-maximal subsets G' of G such that $out(G', A)$ is consistent with C.
- $outfamily(G, A, C) = \{out(G', A) \mid G' \in maxfamily(G, A, C)\}$.

The cases $C = \emptyset$ and $C = A$ express consistency of output, and its consistency with input, respectively. In practice, there is no general rule for choosing the specific instantiation. It all depends on the nature of the particular example we are analyzing. The instantiation $C = A$ is expedient for those cases where the input describes a state of affairs that is "settled" as true, and cannot be changed afterwards.[16] In such cases, it makes sense to require that the output be consistent with the input: "ought" implies "can".

Notably, a set of generators and an input do not have a set of propositions as output, but a set of sets of propositions. So, like in the logics of belief change and nonmonotonic inference, we can infer a set of propositions by taking either a credulous or a skeptical approach. In Makinson and van der Torre (2001), the two

[15] The contrast between the two approaches is emphasized by Makinson in (2005, pp. 107–108).

[16] This reading is related to Hansson's interpretation of circumstances in so-called dyadic deontic logic (see Hansson 1969).

Author's personal copy

resulting operations are called full meet and full join constrained output, and they are noted $\cap$ *outfamily*(G, A, C) and $\cup$ *outfamily*(G, A, C), respectively.

Example 4 below illustrates the notions of maxfamily and outfamily.

Example 4 (Natural conflict) Put $G = \{(a, x), (a, y)\}$ with $C = \{x \rightarrow \neg y\}$ and $A = \{a\}$. For all the output operations, $out(G, A) = Cn(x, y)$, which is inconsistent with C. The maxfamily has two elements $\{(a, x)\}$ and $\{(a, y)\}$, and thus the outfamily has two elements $Cn(x)$ and $Cn(y)$. Let the final output be calculated using the full meet operation. We have

$$\cap outfamily(G, A, C) = Cn(x) \cap Cn(y) = Cn(x \vee y)$$

Since no rule has priority over the other, the most that comes is that the disjunction of x and y is obligatory.

Now I turn to the task of adding priorities. Here I follow the most natural approach. I assume that conflicts are resolved based on a priority ordering on the power set of generators, such that only a suitably chosen subset of the maxfamily is used to generate the output. This is implemented using a relation on sets of rules. However, in practical applications, one uses a relation on rules, not a relation on sets of rules. So the question is: given a relation on rules, how can it be lifted to a relation on sets of rules? The definition of lifting given below is taken from Brass (1991).[17] The relation $\geq$ is read "at least as strong as", and the superscript s (mnemonic for "set") is used to distinguish between the two relations.

I assume that $\geq$ is a pre-order, i.e., the relation is reflexive and transitive. $>$ denotes the strengthened complement of $\geq$, defined by putting $a > b$ whenever $a \geq b$ and $b \ngeq a$. Each of the latter two notions has a counterpart in terms of ordering on sets. And $a \in S$ is called a $\geq^s$-maximal element of S if, for all $b \in S, b \geq^s a$ implies $a \geq^s b$.

Definition 5 (*Lifting*) Let G be a set of generators equipped with a pre-order $\geq$. For any $G_1, G_2 \subseteq G$, we define $G_1 \geq^s G_2$ to hold if for every $\delta_2 \in G_2 - G_1$ there is $\delta_1 \in G_1 - G_2$ with $\delta_1 \geq \delta_2$.

Different proposals on how to lift an ordering of propositions to sets of propositions can be found in the literature. Although it is not the purpose of this paper to assess their relative merits, I will motivate the move from the Horty $\forall\forall$-definition to the Brass one by means of an example in a moment.

Now comes the main construction.

Definition 6 (*Outfamily with priorities*) Let A and G be an input set and a set of generators equipped with a pre-order $\geq$, respectively. Let $\geq^s$ be defined as in definition 5. We put

– *maxfamily*(G, A, C) is the set of $\subseteq$-maximal subsets G' of G such that *out*(G', A) is consistent with C

[17] Paternity questions are not always easy to answer. I follow Halpern (1997, p. 4), who seems to give Brass credit for the idea of using such a definition in the area of non-monotonic reasoning. The definition popped up in several places at around the same time, and had been used before in other areas.

Author's personal copy

- *filterfamily*(G, A, C) is the set of $G' \in$ *maxfamily*(G, A, C) that "maximize" the output, i.e., that are such that $out(G', A) \subset out(G'', A)$ for no $G'' \in$ *maxfamily* (G, A, C)
- *preffamily*(G, A, C) is the set of $\geq^s$-maximal elements of *filterfamily*(G, A, C)
- $preffamily_d(G, A, C)$ is the set of elements G' of *preffamily*(G, A, C) stripped of all the pairs (a, x) that are "inactive" in G', in the sense that $out(G', A) = out(G' - \{(a, x)\}, A)$.

The subscript "d" is short for "distilled".

Some further comments on definition 6 are in order. I suggest viewing each member of the distilled preffamily as the analogue of a proper scenario in Horty's theory. There are four steps involved in their construction. We start by determining the maxfamily. It gathers all the maximal subsets of G whose output remains consistent with the constraints C. This step is mandatory to guard against possible contradictions when applying rules to the input set. The filterfamily, then, selects the elements of the maxfamily with the most informative output, i.e., the elements of the maxfamily that "maximize" the number of conclusions that can be drawn. This is needed to avoid an unwanted loss of information created by the threshold idea as applied to rules. Example 9 below provides a good illustration of this. The preffamily, then, determines the most preferred elements in the filterfamily by use of the given priorities between the norms. This step is needed to resolve conflicts between norms. As will become clear in the treatment of the examples, each element in the preffamily may contain norms that are in force, but are not triggered. The last step consists in removing them from the preffamily, because they have no effect on the output. In Boella and van der Torre (2008), such norms are called "redundant". In the present context I prefer to call them "inactive". Like in Horty's framework, the rules in the distilled preffamily tell us what counts as a reason for what.

Although it is not essential for present purposes, I introduce below the associated output operation. It can be viewed as the analogue of the notion of extension in Horty's account. The subscript p is short for "preferred".

Definition 7 (*Preferred output*) Let G, A and C be a pre-ordered set of generators, an input set, and a set of integrity constraints, respectively. We define

$$x \in out_p(G, A) \quad \textit{iff } x \in \cap\{out(G', A) \mid G' \in preffamily_d(G, A, C)\}$$

Below I apply the account to the examples from Sect. 2, and I show that it yields the same results as Horty's.

Example 5 (The book) Let b, s, y and l be instantiated as in example 1. Put $C = \{y \rightarrow \neg l\}$. Let $G = \{(b, y), (s, l)\}$ with $(s, l) > (b, y)$.

Case 1 Assume $A = \{b\}$. For all output operations, $out(G, A)$ equates $Cn(y)$, which is consistent with C. In this case, the outfamily/filterfamily/preffamily has one element $\{(b, y), (s, l)\}$. And the distilled preffamily has one element $\{(b, y)\}$, providing b as a reason for y.

Case 2 Assume $A = \{b, s\}$. For all output operations, $out(G, A)$ equates $Cn(y, l)$, and thus it is inconsistent with C. The maxfamily has two

Author's personal copy

elements, $\{(b, y)\}$ and $\{(s, l)\}$, and so has the filterfamily. Furthermore, $\{(s, l)\} >^s \{(b, y)\}$ since $(s, l) > (b, y)$. So the preffamily/distilled preffamily has only one element $\{(s, l)\}$, providing s as a reason for l.

Example 6 (The neighbours) Put $G = \{(\top, a), (a, t), (\neg a, \neg t)\}$, where a and t are instantiated as in example 2. Let $\geq$ be the empty relation, and $out \in \{out_3, out_4\}$. Put $C = A$.

Case 1 $A = \emptyset$. $out(G, A)$ is $Cn(a, t)$ which is consistent. So the outfamily, filterfamily and preffamily are the same. They have G as element. And the distilled preffamily has one element $\{(\top, a), (a, t)\}$, providing a as a reason for t.

Case 2 $A = \{\neg a\}$. $out(G, A)$ equates $Cn(a, t, \neg t)$, which is inconsistent. The outfamily/filterfamily/preffamily has one element $\{(a, t), (\neg a, \neg t)\}$. The distilled preffamily has one element $\{(\neg a, \neg t)\}$, providing $\neg a$ as a reason for $\neg t$.

Example 7 (Multiple level of violation) Put $G = \{(\top, \neg a), (a, x), (a \wedge \neg x, y)\}$, where a, x and y are instantiated as in example 3. Let $\geq$ be the empty relation. Put $C = A$.

Case 1 $A = \emptyset$. $out(G, A)$ is $Cn(\neg a)$ which is consistent. The outfamily, filterfamily and preffamily are the same. They have G as element. And the distilled preffamily has one element $\{(\top, \neg a)\}$.

Case 2 $A = \{a\}$. The outfamily/filterfamily/preffamily has one element $\{(a, x), (a \wedge \neg x, y)\}$. The distilled preffamily has one element $\{(a, x)\}$, providing a as a reason for x.

Case 3 $A = \{a, \neg x\}$. The outfamily/filterfamily/preffamily/distilled preffamily has one element $\{(a \wedge \neg x, y)\}$, providing $a \wedge \neg x$ as a reason for y.

Example 8 below explains why the "$\forall\forall$" definition of lifting has been replaced with the Brass one. This is a variation on an example by Goble (2011, section 4.1).

Example 8 (Mission) Put $G = \{(\top, a), (\top, b), (\top, c)\}, A = \emptyset, C = \{a \rightarrow (b \rightarrow \neg c), a \rightarrow \neg b\}$ and $(\top, a) > (\top, b) > (\top, c)$. Intuitively, the three options are inconsistent together, and so are the best and the second-best ones. As an illustration, one might imagine the case of a company who has identified new business partners in three different cities. An employee of the company (call him Brown) is asked to go and meet them. a is the obligation to go to London, b is the obligation to go to Paris, and c to Amsterdam. The ordering tells us how important to the company the missions are. c is a low priority for them, b is a higher priority than c, and a is an even higher one. Assume that, for one reason or another, Brown cannot go to all three cities.[18] Given the priorities the company has set, Brown

[18] One could imagine that the meetings must be held within a relatively short time frame, and that Brown is already very busy. He can take at most two consecutive days off within the time frame in question—the other days he has to attend to his other duties, and these are stronger than the missions he has just been assigned. He will not be able to have the three meetings in two days. Still, he should be able to arrange two meetings.

Author's personal copy

should go to London and Paris, and postpone his trip to Amsterdam. Now assume that, for one reason or another, he cannot combine a trip to London with one to Paris.[19] What should he do?

The outfamily and the filterfamily are the same. They have two elements. One is $\{(\top, a), (\top, c)\}$, and the other is $\{(\top, b), (\top, c)\}$.

If $\geq^s$ is defined as in definition 5, then $\{(\top, a), (\top, c)\} >^s \{(\top, b), (\top, c)\}$. So the preffamily and the distilled preffamily have one element $\{(\top, a), (\top, c)\}$, and thus $out_p(G, A) = Cn(a, c)$. This is as it should be. The best outweighs the second-best, when each is combined with the worst. Intuitively, Brown should make the necessary arrangements to go to both London and Amsterdam.

Now, suppose $G_1 \geq^s G_2$ means that $\delta_1 \geq \delta_2$ for each δ_1 in G_1 and δ_2 in G_2. This alternative definition makes $\{(\top, a), (\top, c)\}$ and $\{(\top, b), (\top, c)\}$ incomparable under $\geq^s$. So in that case, the preffamily and the distilled preffamily both contain $\{(\top, a), (\top, c)\}$ and $\{(\top, b), (\top, c)\}$, and thus $out_p(G, A) = Cn((a \vee b) \wedge c)$. This is counter-intuitive. It says that c is obligatory along with the disjunction of a and b. So Brown might comply with his obligation by just making b and c true. Does c not come last in the ordering? Compared to a, is b not sub-optimal?[20]

Example 9 below shows why the filterfamily step is needed in definition 6.

Example 9 (The birthday) Assume $G = \{(\top, b), (b, t)\}$, where b denotes the proposition that I go to a birthday party, and t the proposition that I tell their organizers that I am coming. Assume the organizers need to know the exact number of persons who will come. In this respect, $(b, t) > (\top, b)$. Let $out \in \{out_3, out_4\}$. Put $A = \{\neg t\}$ and $C = A$. This means that $\neg t$ is settled as true. The outfamily has two elements, $\{(\top, b)\}$ and $\{(b, t)\}$. The filterfamily has one element $\{(\top, b)\}$, because $out(\{(b, t)\}, A) = Cn(\emptyset) \subset out(\{(\top, b)\}, A) = Cn(b)$. So the preffamily and the distilled preffamily have one element $\{(\top, b)\}$, and thus $b \in out_p(G, A)$.

Now take definition 6. Assume the filterfamily step is removed from it, and the notion of preffamily is amended accordingly: *preffamily*(G, A, C) is the set of $\geq^s$-maximal elements of *maxfamily*(G, A, C) (the notion of distilled preffamily remains the same). Since $(b, t) > (\top, b)$, the preffamily and the distilled preffamily have $\{(b, t)\}$ as unique element, and thus $out_p(G, A) = Cn(\emptyset)$, which might be considered to be counter-intuitive.

The reader is invited to verify that Horty's account yields the same outcome in example 9. That is, the subset $S_8 = \{(\top, b)\}$ is delivered as unique proper scenario, and thus $b \in \mathcal{E}$.

Another interesting feature of the filterfamily step is that it clarifies how the present account connects with Reiter's one. We can read each element of G as a Reiter normal rule $a{:}b/b$. So let (G, A) be a Reiter normal default system. Let *extfamily*(G, A) denote the family of all the extensions of (G, A) in the sense of

[19] One could imagine that a strike in France disrupts flights and rail services between Paris and London.

[20] For another illustration, suppose I am in charge of hiring two people. Candidate a is my first choice, candidate b is my second choice, and candidate c is my third choice. Now, a and b cannot stand each other, and refuse to work at the same place. They come to me and say "If you hire him, I will turn your offer down". If I stick to my initial ranking, I will pick up a and c.

Author's personal copy

Reiter. Using an observation from Makinson and van der Torre (2000) it is straightforward to show that the account proposed here is a conservative generalization of the Reiter account in the following sense.

Remark Let *out* be out_3^+.[21] If $\geq$ is the empty relation, then

$$extfamily(G,A) = \{out(G',A) \mid G' \in preffamily(G,A,C)\}$$

where C is either A or $\emptyset$.

Proof The proof is straightforward, combining the observation from Makinson and van der Torre (2000, section 4) that (for $out = out_3^+$, and $C = A$ or $\emptyset$) *extfamily*(G, A) consists of exactly the maximal (w.r.t. set-inclusion) elements of *outfamily* (G, A, C)

$$extfamily(G,A) = max(outfamily(G,A,C))$$

with the observation that

$$max(outfamily(G,A,C)) = \{out(G',A) \mid G' \in preffamily(G,A,C)\}$$

when $\geq$ is the empty relation.

Now I present a bottom line example showing why the IOL account is different from Horty's one. The specificity of the threshold idea (as opposed to the increment idea) is brought to the surface in this example. It is the only deontic example I can think of for which the two accounts do not give the same solution.

Example 10 (Bottom line example, with Horty's solution) Assume we have the set $G = \{(a,\neg b),(a,b),(b,c)\}$ with $(a,\neg b) < (a,b) < (b,c)$.

Case 1 Assume $A = \{a\}$. In this case, the ordered default theory (A, G, $<$) yields $S_9 = \{(a, b), (b, c)\}$ as its unique proper scenario, providing a as a reason for b, and b as a reason for c.

Case 2 Assume $A = \{a,\neg c\}$. The ordered default theory (A, G, $<$) yields $S_{10} = \{(a,b)\}$ as its unique proper scenario, hence still providing a as a reason for b.

Example 11 (Bottom line example, with the IOL solution) The setting is as in example 10. Let $out \in \{out_3, out_4\}$. Put $C = A$.

Case 1 Assume $A = \{a\}$. The outfamily and the filterfamily are the same; they have two elements $\{(a, b), (b, c)\}$ and $\{(a,\neg b),(b,c)\}$. We have $\{(a,b),(b,c)\} >^s \{(a,\neg b),(b,c)\}$. So the preffamily/distilled preffamily has one element $\{(a, b), (b, c)\}$, providing a as a reason for b and b as a reason for c.

Case 2 Assume $A = \{a,\neg c\}$. The outfamily and the filterfamily are the same; they have two elements $\{(a, b)\}$ and $\{(a,\neg b),(b,c)\}$. We have $\{(a,\neg b), (b,c)\} >^s \{(a,b)\}$. So the preffamily has one element $\{(a,\neg b),(b,c)\}$,

[21] Cf. footnote 14.

Author's personal copy

> and the distilled preffamily has one element $\{(a, \neg b)\}$, hence providing a as a reason for $\neg b$.

In case 2, Horty's construction keeps the default (a, b), but discards the default (b, c) because it is conflicted, given that $\neg c$ is in A. In the IOL framework the pair (b, c) is kept in the (only) set in preffamily, excluding further addition of (a, b), given that $\neg c$ is in A, but allowing further addition of $(a, \neg b)$.

There is a way to look at the example that makes the IOL solution more intuitive than Horty's solution. The question is: shall I do b or not? The ordering $(a, \neg b) < (a, b)$ says that b has priority over $\neg b$. So it would seem to follow that I should do b. But, in reply, it can be said that the ordering $(a, b) < (b, c)$ says that compliance with the stronger of the two conflicting norms triggers an obligation of even higher rank, namely the obligation to do c. Furthermore, c is already (settled as) false. Hence if I go for b I will put myself in a violation state with respect to a norm with an even higher rank. I might wish to avoid the violation of the most important norm. And so I shall not do b. This is the outcome that definition 6 predicts.

Here is a natural language version of the above example that might be of help to check intuitions.[22] Chemotherapy (also known as chemo) is the general term for any treatment involving the use of chemical agents to stop cancer cells from growing. For each type of cancer, there is a protocol, which describes in detail the modalities, the potential complications and the particular precautions that should be taken during the treatment. Chemo is typically given in cycles, with rest periods between the cycles. Many doctors agree that sticking to the schedule is the first step to success. However, it is often necessary for physicians to postpone chemo. This usually happens when the patient has a low White Blood Cell (WBC) count. According to the textbook (Mehta and Bansal 2004),

> "Chemotherapy should be discontinued or postponed if the total WBC counts drop below 3,500/mm^3." (Mehta and Bansal 2004, p. 144)

White blood cells are part of the body's defense against disease. Chemotherapy may reduce patients' WBC count, making them vulnerable to infection, including life-threatening infection, especially among the elderly.[23] There are drugs (called WBCs growth factors) physicians can prescribe to help preserve WBCs during chemo. However, they cannot always be administrated. For instance, the drug called "Neulasta"[24] is contraindicated in patients with a history of serious allergic reactions to pegfilgrastim or filgrastim. If it is true that sticking to the prescribed chemo schedule is the first step to success, then (in the case of abnormalities) the best is to administer a WCBs booster. And, if (for some reason or another) this

[22] Thanks to Marek Sergot for suggesting me this example, though perhaps not in these exacts terms.

[23] Chemotherapy works by killing fast-growing cancer cells. Unfortunately, chemotherapy drugs cannot always tell the difference between cancer cells and fast-growing healthy cells, including red and white blood cells.

[24] http://www.neulasta.com/.

Author's personal copy

cannot be done, the second best is to postpone chemo until the passage of time brings WBCs count back to a safe level.

In this medical example, the following instantiations suggest themselves: *a* is for the set of data used to set up the treatment. *b* is for receiving chemo as per the protocol. This one contains very specific information on e.g. the dosage, the day, the hour, the order and the duration of drug administration. And *c* is for keeping WBCs count to a safe level using a drug. In the extreme case where having one more chemo means certain death for the patient, surely enough the oncologist will decide to interrupt the treatment.

This is also an example of polarity reversal. In case 1, the consideration *a* is a reason for *b*, and *b* is a reason for *c*. In case 2, *a* is no longer a reason for *b*, and *b* is no longer a reason for *c*. Hence Horty's theory gets the polarity reversal right in simple cases of reason holism, but appears not to be well-suited for slightly more complex cases like this one, where more than two norms collide—viz. where two obligations conflict, and disregard for the more important one is the only way to avoid the violation of an even more important obligation.

Clearly there is a disagreement between the two accounts. How should it be diagnosed? Here is a possible explanation—there might be others. Dancy believes that reasons behave the same way whether epistemic or deontic. It is even this very belief that led him to formulate the moral particularism thesis.[25] Horty seems to follow this line. First, he suggests using the same notation for both kinds of reason, and warns the reader that throughout the discussion he will switch back and forth, rather casually, between the two readings (see Horty 2007b, p. 4). Next, in order to account for the valence-switching capacity of deontic reasons, Horty uses a mechanism initially devised for and tailored to the epistemic domain. A default is disregarded when its head is inconsistent with the facts, either directly or indirectly (using reasons that have previously been accepted). This feature is the main reason why in case 2 of example 10 the theory supports an arguably unintuitive scenario. In the epistemic domain, it makes sense to disregard a default for the sole reason that its head appears to be inconsistent with the rest of the data. Indeed the latter provide an exception to the normality claim the default reason in question rests on. What of the deontic domain? The conventional wisdom is to assume that deontic reasons run parallel to epistemic reasons. It makes perfect sense to assume such a parallelism, at least in the absence of clear evidence to the contrary. Prima facie such an assumption is useful too, for it allows to transfer well-defined and well-documented techniques from one area to the other. However, if correct, the bottom line example casts doubt on the latter assumption. Here lies, I believe, the explanation. To say that holism is a general phenomenon that applies to both normative and epistemic reasons is one thing. To say that reasons behave the same way whether epistemic or normative is another thing.

One would make a stronger statement, by showing that there is a sensible interpretation suggesting that for epistemic reasons the solution supplied by the Horty account is the correct one. The reliability interpretation outlined in Horty

[25] In Dancy (2006, chapter 4.2), he points out that holism is uncontested for epistemic reasons, and therefore it would be surprising if practical reasons behaved differently.

Author's personal copy

(2007a, p. 391) might be the one. Under the latter interpretation, each default indicates something like a high conditional probability that its conclusion is satisfied, and the priority ordering measures relative strength of these conditional probabilities. Turning back to the example, the default (a, b), call it δ_2, tells us that b follows with a high probability, given that a holds. The potential competing argument $(a, \neg b)$, call it δ_1, has no force. For δ_1 supports not-b given a less strongly than δ_2 supports b given a. Not-c has no force either. It is true that the default reason (b, c), call it δ_3, supports c given b more strongly than δ_2 supports b given a. However, the contrapositive of δ_3, "if not-c then not-b" is logically unrelated to it, and cannot be used to defeat δ_2.[26] The solution that is naturally suggested by the reliability interpretation is, thus, the outcome predicted by the Horty account.

There is a possible escape to the bottom line example.[27] Horty (2007, p. 391, 394) distinguishes between a conditional command and the (unconditional, categorical) command of a conditional. A conditional command has the form (a, x), while the command of a conditional has the form $(\top, a \rightarrow x)$, where $\rightarrow$ is material implication. I am not able to explain (in terms of reasons) the difference in meaning between the two. But let us see what happens if the bottom line example is represented as imperatives of conditionals rather than as conditional imperatives. Suppose the input is $A = \{a\}$. In that case, the unique proper scenario associated with the default theory is $S_{11} = \{(\top, a \rightarrow b), (\top, b \rightarrow c)\}$. At the first stage all the defaults are triggered. At the second stage the default $(\top, a \rightarrow \neg b)$ is conflicted since $A \cup h(S_{11}) \vdash a \wedge b$. Now suppose the input is $A = \{a, \neg c\}$. In that case, the unique proper scenario associated with the default theory is $S_{12} = \{(\top, a \rightarrow \neg b), (\top, b \rightarrow c)\}$. At the first stage all the defaults are triggered. At the second stage $(\top, a \rightarrow b)$ is conflicted in S_{12}. On the other hand, the counter-intuitive $S_{13} = \{(\top, a \rightarrow b)\}$ is no longer supported. This is because $(\top, a \rightarrow b)$ is defeated by $(\top, b \rightarrow c)$ in S_{13}. Therefore, if the bottom line example is represented as imperatives of conditionals rather than as conditional imperatives, then Horty's account gives the right result.

I note an apparent oddity. The rule $(\top, b \rightarrow c)$ cannot be removed from the proper scenario S_{12}. This is because $(\top, a \rightarrow \neg b)$ is defeated by $(\top, a \rightarrow b)$ in $S_{12}^{\#} = \{(\top, a \rightarrow \neg b)\}$, although it is not in S_{12}.[28] There was the question of explaining (in terms of reasons) the difference in meaning between a conditional command of the form (a, x) and its counterpart in the 'command of a conditional' language $(\top, a \rightarrow x)$. There is now the related question of explaining what the function of the proper scenario is under the 'command of a conditional' interpretation. I would say that it must be enlarged to include counterfactual thinking. The proper scenario gathers the reasons that are binding in a given context, but also (sometimes) tells us what *would*

[26] For illustration purposes, assume that these conditional probabilities encode statistical assertions about some population groups, and instantiate a, b and c into (this is the classical example of non-transitivity of default rules) *being a student, being an adult*, and *being employed*. The fact that I am unemployed does not imply that I am not an adult.

[27] I am indebted to Jeff Horty for drawing this possible reply to my attention.

[28] Why should the presence of $(\top, b \rightarrow c)$ in the current scenario make a difference? This is essentially due to the consistency requirement in condition b) of the definition for defeat, viz the restriction according to which $A \cup h(S^{S'/D'})$ should be consistent. Take $D' = \{(\top, a \rightarrow b)\}$. The only S' that qualifies is $\{(\top, a \rightarrow \neg b)\}$. $A \cup h((S_{12}^{\#} - S') \cup D'$ is consistent, but $A \cup h((S_{12} - S') \cup D')$ is not.

Author's personal copy

qualify as a reason for what *if* the context was different. Here b would have counted as a good reason for c in other circumstances.

The move to the 'command of a conditional' idiom is perfectly possible. Still, the question comes up to what extent independent motivation can be given for it. I am not able to provide one. Noticeably, the translation into the 'command of a conditional' idiom gives the contraposition of every rule with exceptions as long as it is part of a proper scenario. This is illustrated with example 12 below.

Example 12 (Contraposition) $A = \{\neg x\}, G = \{(\top, a \rightarrow x)\}$. The proper scenario is $S_{14} = \{(\top, a \rightarrow x)\}$, and thus $\neg a \in \mathcal{E}$.

Generally, authors agree in thinking that an appropriate translation of rules should not give the automatic contraposition of the rules.[29] The following example, due to Hansen (2008, section 3.2), shows that contraposition is undesirable for obligations.

Example 13 (Best suit) $A = \{\neg c\}$, and $G = \{(\top, a \rightarrow b), (\top, \neg a \rightarrow c)\}$, where a, b, and c represent the respective propositions that it rains, that I wear a rain coat, and that I wear my best suit. The proper scenario is $S_{15} = \{(\top, a \rightarrow b), (\top, \neg a \rightarrow c)\}$, and thus $b \in \mathcal{E}$. In the event that I cannot wear my best suit (e.g., it is in the laundry), it seems strange to say that I should wear a rain coat.

Therefore, as a general policy, the translation into the 'conditional command' idiom should be preferred unless something goes wrong with it—in which case the translation into the other idiom becomes appropriate. Still, some might find such a move somewhat *ad hoc*, if no independent motivation can be given to support it.

One last remark. I have taken the standard approach to be proceeding in increment steps, and I have suggested that the latter method is not well-suited to deal with obligations. As a matter of facts, it is possible to imagine variations on the increment method. One such variation in the literature is worth a mention, because it gives the same solution for the bottom line example as the I/O account. It is the increment method of Hansen (2008, section 5.2). When building the preferred subset of norms, the idea is to check for the collective "obeyability" of the defaults, rather than their collective consistency as done in most non-monotonic frameworks. This is implemented by testing for consistency not the heads of the rules, but their materializations. In the particular case where the ordering is total, the procedure boils down to the following. Suppose the rules in G are ordered in the matter written, $g_1, g_2, g_3, \ldots$, the convention being that concatenation is value decreasing: each element in the sequence takes priority over its immediate successor. Define the preferred maximally obeyable subset $\mathcal{G} \subseteq G$ of norms from a sequence of subsets, G_n, thus:

- $G_0 = \emptyset$
- $G_{n+1} = \begin{cases} G_n \cup \{g_{n+1}\} & \text{if } A \cup m(G_n \cup \{g_{n+1}\}) \text{ is consistent} \\ G_n & \text{otherwise} \end{cases}$
- $\mathcal{G} = \cup\{G_n : n < \omega\}$

[29] Cf. e.g. Lewis (1973, section 1.8), Adams (1975, chapter 1), Bennet (2003, chapter 59), Makinson (1999, section 3.2) and Hansen (2008, section 3.2).

Author's personal copy

I give without proof the solution we obtain for the bottom line example. If $A = \{a\}$, then $\mathcal{G} = \{(a, b), (b, c)\}$. If $A = \{a, \neg c\}$, then $\mathcal{G} = \{(a, \neg b), (b, c)\}$. Troublesome for the account is, I think, the birthday scenario (example 9). We have $G = \{(\top, b), (b, t)\}$ with $(b, t) > (\top, b)$ and $A = \{\neg t\}$. Since $(b, t) > (\top, b), (b, t)$ is added first, preventing the further addition of $(\top, b)$. Hence $\mathcal{G} = \{(b, t)\}$.

5 Conclusion

This paper focused on the divide between particularism and generalism, and argued for a middle ground position between these two extremes. This view was already defended by Horty, who made his point by showing (*pace* Dancy) that reason holism does not imply particularism. The aim of this paper was to strengthen Horty's argument. In the literature only the notion of prima facie obligation has been used to motivate reason holism. I strengthened Horty's point, by first showing that contrary-to-duties provide another independent support for reason holism, and then outlining a formal theory that is able to capture these two sources of holism. While in simple settings the proposed account coincides with Horty's one, this is not true in more complicated or "realistic" settings in which more than two norms collide.

The aim of the paper was more general than just comparing two particular logical systems. Its purpose was to contrast two strategies of formalizing defeasible deontic reasoning. The first one, referred to as the "increment" idea, has been around for a while in the non-monotonic literature. It consists in restraining the step-by-step application of defaults. The second one, referred to as the "threshold" idea, is much less known, and certainly deserves further study. It consists in cutting back the set of defaults to just below the threshold of yielding excess. Comparing the two approaches to see if one provides a better account of the valence-switching capacity of reasons, we have seen that in the main example where their outcomes differ, the intuitions seem to favor the outcome of the second strategy. The comparison has given rise to a more general observation. It is that the commonly made assumption that normative reasons run parallel to epistemic ones is not as plausible as might seem at first sight.

Needless to say, much work remains to be done to get a full-blooded formal account. For one thing, I have been concentrated on two specific frameworks. Future studies should include alternative approaches available from the default logic literature (like those documented in Delgrande et al. (2004)). And a wider set of examples might be used to test and compare the approaches. For another, axiomatisation would be a good way of exporting deontic logic to moral particularism. However, just as default logic lacks a proof-theory, so also constrained I/O logic lacks a proof-theory. This issue must be left for future research.

Of course, I do not claim that the threshold approach is completely free of technical problems, especially when it comes to epistemic reasons. For instance, I can think of an epistemic scenario that is structurally different from the one I have

Author's personal copy

mostly been concerned with in this paper, and for which the I/O account does not provide the correct solution. It is the Microsoft example from Horty (2001). Earlier I mentioned this example in connection with argumentation theory. There is something peculiar about the scenario. It is that the consequent of the least preferred default is logically stronger than that of the most preferred default. Since the example is somewhat involved I use mnemonic abbreviations, focusing on a particular individual, Beth, and letting $NMEb$, MEb, $1Mb$, and $\frac{1}{2}Mb$ represent the respective propositions that Beth is a new Microsoft employee, a Microsoft employee, millionaire, and has half a million dollars.

Example 14 (Microsoft) $A=\{NMEb\}, G=\{(MEb,1Mb),(NMEb,\neg\frac{1}{2}Mb),(\top,\frac{1}{2}Mb)\}$ where $(MEb,1Mb)<(NMEb,\neg\frac{1}{2}Mb)<(\top,\frac{1}{2}Mb)$, and $C=\{NMEb\rightarrow MEb, 1Mb\rightarrow\frac{1}{2}Mb\}$. The input is that Beth is a new Microsoft employee. As a first and very weak default, we have that (in virtue of stock options accrued over the years) Microsoft employees (including Beth) are millionaires. As a slightly stronger default, we have that new Microsoft employees have not yet accumulated so much as half a million dollars. As a third and very strong default, we have that Beth does happen to have half a million dollars.

The maxfamily and the filterfamily are the same. They have the following two elements: $\{(NMEb,\neg\frac{1}{2}Mb)\}$ and $\{(MEb,1Mb),(\top,\frac{1}{2}Mb)\}$. The former is in the preffamily and the distilled preffamily, but not the latter. Therefore $out_p(G,NMEb)=Cn(1Mb,\frac{1}{2}Mb)$.[30] Intuitively, we would expect $out_p(G,NMEb)=Cn(\frac{1}{2}Mb)$.

The reader is invited to verify that, on Horty's account, the subset $S_{16}=\{(\top,\frac{1}{2}Mb)\}$ is delivered as unique proper scenario.

I do not claim that one and the same formalism should be able to deal with both epistemic and deontic reasons. Unless one can instantiate the scenario in the deontic realm, I will not worry too much about the latter finding. At first, it corroborates further that deontic and epistemic reasons do not operate the same way. Now, just supposing someone can come up with a deontic illustration, there might be a way around in I/O logic. One might pre-process the set of generators in order to "freeze" the application of the less preferred/specific default. Call (a, x) a safe (there might be a better name in the present context) element of G if (a, x) is not a minimal element (under $\geq$) of any minimal (under $\supseteq$) subset of G whose output is inconsistent with C. Write G_s for the set of safe elements of G. The idea is to assume that definition 6 is applied to G_s rather than G itself. In the example, we have $G_s=\{(NMEb,\neg\frac{1}{2}Mb),(\top,\frac{1}{2}Mb)\}$, and thus the final solution is $\{(\top,\frac{1}{2}Mb)\}$ as required. It remains to be seen whether the pre-processing idea would always give the right result.

Acknowledgments I wish to thank the three anonymous referees for their criticisms. Special thanks are due to Marek Sergot for his example, and to Jeff Horty for his enthusiasm about the project and his insights. Thanks are also due to Jörg Hansen, David Makinson and Leendert van der Torre for valuable comments. The feedback received at DEON'10 and the workshop on the "Dynamics of Normative

[30] The formula $1Mb$ is also part of the output that is delivered by an argumentation framework based on the notion of reinstatement (cf. p. 8).

Author's personal copy

Reasoning" held on 6 Apr 11 at Ghent was also useful. Support from the National Research Fund (F.N.R.) of Luxembourg (grant PDR-08-013) is gratefully acknowledged.

References

Adams EW (1975) The logic of conditionals. D. Reidel, Dordrecht
Bennett J (1995) The act itself. Oxford University Press, New York
Bennett J (2003) A philosophical guide to conditionals. Oxford University Press, New York
Boella G, van der Torre L (2008) Institutions with a hierarchy of authorities in distributed dynamic environments. Artif Intell Law 16(1):53–71
Brass S (1991) Deduction with supernormal defaults. In: Brewka G, Jantke K, Schmitt P (eds) Nonmonotonic and inductive logics, volume 659 of LNCS. Springer, Berlin, pp 153–174
Brewka G, Eiter T (2000) Prioritizing default logic. In: Intellectics and computational logic (to Wolfgang Bibel on the occasion of his 60th birthday). Kluwer, B.V., Deventer, pp 27–45
Chisholm RM (1963) Contrary-to-duty imperatives and deontic logic. Analysis 24(2):33–36
Cooper, J (eds) (1997) Plato: complete works. Hackett, Indianapolis
Dancy J (1993) Moral reasons. Blackwell, Oxford
Dancy J (2006) Ethics without principles. Oxford University Press, Oxford
Delgrande JP, Schaub T, Tompits H, Wang KA (2004) Classification and survey of preference handling approaches in nonmonotonic reasoning. Comput Intell 20(2):308–334
Gabbay D, Horty J, van der Meyden R, van der Torre L (eds) (forthcoming) Handbook of deontic logic and normative systems. College Publications, London
Goble L (2011) Prima facie norms, normative conflicts and dilemmas. To appear in Gabbay et al. (forthcoming)
Grossi D, Jones A (2011) Constitutive norms and counts-as conditionals. To appear in Gabbay et al. (forthcoming)
Halpern J (1997) Defining relative likelihood in partially-ordered preferential structures. J AI Res 7:1–24
Hansen J (2008) Prioritized conditional imperatives: problems and a new proposal. J Auton Agent Multi Agent Syst 17(1):11–35
Hansson B (1969) An analysis of some deontic logics. Noûs 3:373–398
Hare R (1952) The language of morals. Clarendon Press, Oxford
Hare R (1963) Freedom and reason. Clarendon Press, Oxford
Harman G (2005) Moral particularism and transduction. Philos Issues 15(1):44–55
Horty J (2001) Argument construction and reinstatement in logics for defeasible reasoning. Artif Intell Law 9(1):1–28
Horty J (2007a) Defaults with priorities. J Philos Logic 36:367–413
Horty J (2007b) Reasons as defaults. Philos Imprint 7(3):1–28
Husserl E (1988) Vorlesungen über Ethik und Wertlehre (1908-1914). In: Melle U (ed) Husserliana, vol 28. Kluwer, Dordrecht
Lee M-K (2005) Epistemology after Protagoras. Responses to relativism in Plato, Aristotle, and Democritus. Clarendon Press, Oxford
Lewis DK (1973) Counterfactuals. Clarendon Press, Oxford
Mackie JL (1977) Ethics: inventing right and wrong. Viking Penguin Inc., New York
Makinson D (1994) General patterns in nonmonotonic reasoning. In: Gabbay D, Hogger C, Robinson J (eds) Handbook of logic in artificial intelligence and logic programming, vol 3. Oxford University Press, Inc, New York, pp 35–110
Makinson D (1999) On a fundamental problem of deontic logic. In: McNamara P, Prakken H (eds) Norms, logics, and information systems, vol 49 of frontiers in artificial intelligence and applications. IOS Press, Amsterdam, pp 29–53
Makinson D (2005) Bridges from classical to nonmonotonic logic. College Publications, London
Makinson D, van der Torre L (2000) Input/output logics. J Philos Logic 29(4):383–408
Makinson D, van der Torre L (2001) Constraints for input/output logics. J Philos Logic 30(2):155–185
Mehta A, Bansal SC (2004) Diagnosis and management of cancer. Jaypee Brothers Medical Publishers, New Delhi
Nute, D (eds) (1997) Defeasible deontic logic, vol 263 of synthese library. Kluwer, Dordrecht

Author's personal copy

Prakken H, Sartor G (1997) Argument-based extended logic programming with defeasible priorities. J Appl Non Class Logics 7(1):25–75
Rawls J (1955) Two concepts of rules. Philos Rev 64:3–32
Raz J (1999) Practical reasons and norms. 2nd edn. Oxford University Press, Oxford
Reiter R (1980) A logic for default reasoning. Artif Intell 13(1-2):81–132
Ross WD (1930) The right and the good. Oxford University Press, Oxford
Searle J (1964) How to derive 'ought' from 'is'. Philos Rev 73:43–58
Searle J (1969) Speech acts. Cambridge University Press, Cambridge
Stolpe A (2010) Norm-system revision: theory and application. Artif Intell Law 18(3):247–283