

## OBLIGATION, FREE CHOICE, AND THE LOGIC OF WEAKEST PERMISSIONS

ALBERT J.J. ANGLBERGER

Munich Center for Mathematical Philosophy, LMU Munich

NOBERT GRATZL

Munich Center for Mathematical Philosophy, LMU Munich

and

OLIVIER ROY

Universität Bayreuth

**Abstract.** We introduce a new understanding of deontic modals that we call *obligations as weakest permissions*. We argue for its philosophical plausibility, study its expressive power in neighborhood models, provide a complete Hilbert-style axiom system for it and show that it can be extended and applied to practical norms in decision and game theory.

**§1. Introduction.** In this paper we study the logic of what we call *obligations as weakest permissions*. The basic idea is this. An action type  $\varphi$  is obligatory only in cases where the following two conditions hold.

- $\varphi$  is permitted;
- If  $\psi$  is also permitted then one cannot do  $\psi$  in the present situation while not doing  $\varphi$ .

In that sense  $\varphi$  is the logically weakest permitted action type that the agent can achieve in the situation she is in. This notion is a very natural one: Suppose that  $\varphi$  is obligatory, but that there is some logically weaker and permitted  $\varphi'$ . This means that you may do  $\varphi'$  without ensuring the obligatory  $\varphi$ . But this seems odd. In Section 2 we study and motivate our understanding of obligations and permissions in more detail. This understanding of obligations and permissions results in an interesting non-normal deontic logic. In Section 3 we show some of its core properties, and provide a complete Hilbert-style axiom system for it. In Section 4 we show that the system can be extended to capture stronger (neighborhood) frame conditions, and that it has natural applications to rational recommendations in decision and game theory. The view of obligations and permissions that we propose here thus has interesting mathematical properties. Furthermore, it is philosophically plausible. It is this for which we argue first.

**§2. Philosophical motivations.** In this section we present and defend the philosophical view of the deontic modalities that underlie the formal system that we subsequently develop. We start with permissions. The basic idea is that permissions provide sufficient conditions for an action type to be licensed by a given system of norms, what we call being “normatively OK.” This view results in a known, specific interpretation of permissions, the

---

Received: March 16, 2015.

“open reading”. We argue for this reading first. Then we turn to obligations. We view them as the strongest action type that is necessary for an action to be normatively OK. We first argue that, given the open reading of permissions, this is a natural interpretation of obligation to take. To strengthen this point we finish by presenting a number of concrete cases where obligations and permissions can be understood in the sense that we develop here.

**2.1. Permissions as sufficient conditions: The open reading.** Take an (unconditional) permission statement, for instance “it is permitted to board the plane.” Let us write this  $P\phi$ . Suppose that  $\phi$  is an action type. By this we mean that there are many, mutually exclusive action tokens of that type. There might be many ways to board a plane. There might be more than one gate to go through, there might be several times within a fixed period when one can proceed, etc. What does it mean, then, to say that “it is permitted to board the plane”? How to read  $P\phi$ ? This is David Lewis’ famous “problem about permission” (Lewis, 1979). There are two obvious candidates:

- (a) every token of type  $\phi$  is OK according to the boarding regulations;
- (b) at least one token of type  $\phi$  (but possibly not all) is OK according to the boarding regulations.

By saying that an action token “is OK according to the boarding regulations” we mean that this token is legal, that it is among those which the set of boarding regulations licenses. Following Broersen (2004) we will call (a) *the open reading of permissions*. We argue now that (a) is the correct understanding in *some* situations. This is sufficient for this paper. Provided that the open reading is the right one in certain cases, we will argue next that in those cases obligations should be understood as weakest permissions. We do not argue (and do not need to) that the open reading is always the right or the only sensible understanding of permissions.

Under the open reading permission statements identify the sufficient conditions for an action type to be licensed by a given normative system. In Andersonian-Kangerian terminology, permissions statements like  $P\phi$  can be interpreted as

If you perform a  $\phi$ -type action then it is OK

In other words, any action token of type  $\phi$  is OK in a given normative system. Performing a  $\phi$  action is sufficient for legality. This has also been called *strong permissions* (e.g. Kamp, 1973; Asher & Bonevac, 2005).<sup>1</sup>

There are many situations in which permissions behave this way. Suppose Ann is allowed to board through the front or the back door of the plane. She is allowed to do both. She may board through the front door *and* she may board through the back door. Either way is legal, as far as the boarding regulations are concerned. This is only one mundane case.

<sup>1</sup> There are in fact two senses of strong permissions in the literature. One as above: permission defined through the Open Reading. In that sense obligations discussed in this paper are weakest strong permissions, for the simple reason that all permissions discussed here are strong. Strong permissions have also been interpreted in von Wright’s sense (Wright, 1963), i.e. as explicit permissions as opposed to the mere absence of a prohibition. We do not discuss this sense of strong permission in this paper. But since we only have explicit permissions (we do not discuss the respective duals of O and P), it is also true that obligations are weakest strong permissions in that sense. So one could call the obligations discussed here weakest strong permissions. But we refrain from doing so, partly because of the potential ambiguity between two senses of “strong”, and partly because almost everywhere we use “weak” and “strong” to refer to logical strength, which is again different from the two senses of “strong” just discussed.

There are many more situations where permissions identify sufficient conditions for “legality” (see Section 2.3). These apply to a large range of normative systems, from law to morality to rationality. This last is particularly fitting in our view. We develop it further in Section 4.2.<sup>2</sup>

2.1.1. *The open reading and free choice permission.* As the last example suggests, the open reading corresponds to the well-known *free choice permission* principle (see Makinson, 1984):

$$P(\varphi \vee \psi) \rightarrow P\varphi \wedge P\psi \quad (\text{FCP})$$

(FCP) is clearly valid under the open reading. If by  $P(\varphi \vee \psi)$  we mean every way to ensure that  $\varphi$  or  $\psi$  is permitted, then every way to ensure that  $\varphi$  and every way to ensure that  $\psi$  has to be permitted too. So we get  $P\varphi$  and  $P\psi$ . Now suppose the open reading is not true for a specific permitted action type  $\varphi$ . Write  $\tau_1, \dots, \tau_n$  for the action tokens of type  $\varphi$ .<sup>3</sup>  $\varphi$  is equivalent to the disjunction  $\tau_1 \vee \dots \vee \tau_n$ . So we have  $P(\tau_1 \vee \dots \vee \tau_n)$ . But since the open reading is false, there is at least one  $\tau_i$  that is not permitted. So we have  $\neg P\tau_i$ , which is enough to conclude that (FCP) is false.

Of course (FCP) is not uncontested (e.g. McNamara, 2014; Hansson, 2013). We now consider some of the classical objections to it. The first is the known fact that (FCP) has disastrous consequences in Standard Deontic Logic. That logic validates weakening of permission:

$$P\varphi \rightarrow P(\varphi \vee \psi) \quad (\text{PRW})$$

Our reply to that potential problem is simple. (PRW) is not valid in the logic we develop below. Furthermore, as many have observed, (FCP) and (PRW) correspond<sup>4</sup> to two understandings of permissions which should not be confused. In our terminology, these two notions of permissions are the readings (a) and (b) above. This was already observed by von Wright:

“In the monadic weak permission calculi we prove the theorem ‘ $Pp \rightarrow P(p\&q)$ ’. It says that if a certain thing is permitted, then this thing in conjunction with anything else is permitted too. This would be disastrous, if the deontic logic which permits the step from ‘ $Pp$ ’ to ‘ $P(p\&q)$ ’, also permitted the transition from ‘ $P(p\&q)$ ’ to ‘ $Pp$ ’. But it doesn’t. The inferences from ‘ $Pp$ ’ to ‘ $P(p\&q)$ ’ and from ‘ $P(p\&q)$ ’ to ‘ $Pp$ ’ are both valid – but valid for different concepts of permission.” (Wright, 1968, pp. 33–34)

<sup>2</sup> As an historical aside, it is worth mentioning that von Wright himself seems to have endorsed that reading:

“It seems to me that the most natural way of understanding the phrases is this: “It is permitted that  $p$ ”, no circumstances being specified, means that it is permitted that  $p$ , no matter what the circumstances are, i.e. in *all* circumstances.” (Wright, 1968, pp. 34–35)

<sup>3</sup> This assumption is only mentioned to match the finitary character of the language that we present below. The argument could be run in the general case too, but this would bring in unimportant technicalities later on.

<sup>4</sup> Indeed a similar argument as above for (a) shows that (b) corresponds to (PRW).

A second potential objection is that the open reading is too strong. Suppose boarding is permitted. We rarely mean by this that *all* possible ways of boarding are permitted. What is in fact permitted is boarding, provided one has a valid boarding pass, shows it to the flight attendant at the gate, does not have over-sized hand-luggage, etc...<sup>5</sup> Strictly speaking, it is boarding *provided that no other regulations are violated* that is permitted. The objection is that most permissions are like this. We say all the time that this or that is permitted, but we rarely ever mean by this that every possible way to carry out these actions is permitted.

This problem can be spelled out more formally. Suppose  $\varphi$  is permitted. This action type is classically equivalent to its “factorization” into  $(\varphi \wedge \psi) \vee (\varphi \wedge \neg\psi)$ . So by (FCP) we get that both disjuncts are permitted, for any  $\psi$  whatsoever. This includes cases where  $\psi$  itself is forbidden. And a particular case of this factorization is when  $\varphi = \psi$ , yielding the conclusion that whenever something is permitted then the impossible action type  $\perp$  is also permitted. These two consequences, the objection goes, is a *reductio* against (FCP).

In our view these consequences are not sufficient to warrant the desired *reductio*. Start with the fact that  $P\varphi$  implies  $P\perp$ . This might be seen as an oddity, but philosophically this is harmless.  $P\perp$  is a vacuous permission, since its content cannot be performed. And in the system we develop below  $\perp$  is never obligatory, by the “ought implies can” principle.

Now consider the implication from  $P\varphi$  to  $P(\varphi \wedge \psi)$  when the latter action types are not mutually exclusive. This, the objection goes, leads to strongly counter-intuitive results, for instance when  $\psi$  itself is forbidden. This is essentially the point of Hansson’s “Vegetarian’s Free Lunch” example in Hansson (2013). For the sake of the argument let us grant that if  $\psi$  is forbidden then  $\varphi \wedge \psi$  should not be permitted. This is plausible under the open reading. But even granting that there are two possible reactions. One can reject the implication and with it (FCP), *or* one can keep (FCP) and conclude either that it is really  $\varphi \wedge \neg\psi$  which was permitted in the first place, or that the normative system at hand should be revised. (FCP) and the open reading of permissions demand a very precise description of what is permitted. Permission requires circumscription. The objection above does nothing in the way of showing that it is never possible to do so, and even less that (FCP) cannot be fruitfully put to use in a dynamic process of revising the law. In the example above, it seems perfectly possible to pinpoint precisely what is permitted and what is not. So the objection is not decisive. In certain situations (FCP) and the open reading remain plausible.

This reply can be strengthened by pointing out that the purportedly overly strong character of the open reading can be well handled by moving to *defeasible* norms. The open reading of permissions is by no means bound to the view that such permissions are not defeasible. If we call some  $\varphi$  permitted, we of course usually exclude *implicitly* exceptional cases where some non-permitted  $\psi$  is done together with  $\varphi$ . If someone claims that it is permitted for me to donate a certain percentage of my salary each month, she implicitly excludes cases where I donate this percentage and steal back the same amount of money from my neighbor. In legal contexts the situation is exactly the same: whenever a legal authority calls  $\varphi$  permitted, cases in which  $\varphi$  is combined with other, non-permitted  $\psi$ s are implicitly excluded from this permission, i.e. are not permitted. Our view is that the open reading, which requires precise and complete descriptions of a permission’s content, can

<sup>5</sup> A similar phenomenon has been observed in other areas of research, also affecting other notions (not just deontic ones). In the research on reasoning about actions, it is often assumed to be impossible to explicitly state sufficient conditions which imply the executability of an action (the qualification problem, e.g. Brown, 1987; McCarty, 1980). We would like to thank an anonymous referee for making us aware of this connection. This relation has been studied further in Anglberger *et al.* (2014).

be seen as the default position. If some exceptional case shows up then the system of permissions can (and should) be refined. It is not  $\varphi$  but  $\varphi \wedge \neg\psi$  that is permitted. But for that stricter permission the open reading remains in force.

*2.1.2. Obligations under the open reading.* Given that permissions are seen as providing sufficient conditions for legality, the natural option for  $O\varphi$  is to view it in terms of necessary conditions.<sup>6</sup> To go back to our boarding regulation example, we get that no token of type  $\neg\varphi$  is OK according to the boarding regulations.

(a') If you fail to do a  $\varphi$ -type of action then it is not OK

The reading of obligations that we defend in this paper is stronger than (a'). It entails it. Under that reading, when obligations imply permissions, the unique obligatory act type is the set of legal tokens. This is what we argue for in the next section.

*2.2. Argument for obligations as weakest permissions.* We now argue that, in situations where the open reading of permissions applies, it is natural to view obligations not only along (a') but also as strongest necessary conditions, i.e. as those that are entailed by all permissions.

(a'') A  $\varphi$ -type of action is obligatory if and only if for all act-types  $\psi$  that are permitted, if  $h$  is a token of type  $\psi$  then it is also a token of type  $\varphi$ .

If one adds to (a'') that obligatory act types must also be permitted, one gets that obligations are logically weakest permissions. We first argue that being the weakest permission is necessary for an action to be obligatory, under the open reading of permissions. Then we give a similar argument to the effect that the implication also goes the other way around. This latter argument uses an additional assumption, namely that there is always some action type that is obligatory. We are not especially worried about this. The core system that we present and develop in Section 3 only takes the "only if" direction of (a''). In Section 4 we consider an extension of the core system that does capture the "if" direction.

There is a direct argument for the "only if" direction of (a''), given (a') and the open reading. Suppose obligations and permissions give, respectively, necessary and sufficient conditions for legality. Then any token of a permitted type  $\psi$  is OK, which in turn implies that that token is of type  $\varphi$ , for any  $\varphi$  that is obligatory. In short: if  $\varphi$  and  $\psi$  are respectively necessary and sufficient for being OK, we have:

$$\psi \rightarrow OK \rightarrow \varphi$$

This is the "only if" in (a'').

Let us look at a less abstract argument. Suppose I ought to save a child who is drowning, but that this is not the logically weakest permission that I have. Maybe saving the child is not permitted in the first place. Then we are under a system of obligations where some obligatory actions are not permitted. Such a system is not implausible in our view. But it is surely not mainstream in deontic logic. We leave that possibility aside in this paper. So suppose that saving the child is permitted. That it is not the among the logically weakest permissions means that there must be another action type, non-equivalent and logically weaker, which is permitted. Suppose for the sake of illustration that this is "saving the child or not ruining my shoes" (possibly by not jumping in the water with them on), and that there is one token of that type where I save my shoes but not the child. Then the open

<sup>6</sup> The open reading of permissions does not commit in itself to that particular view of the interpretation of obligations. It is simply the natural option to take.

reading of permissions results in this token being permitted. But this is highly counter-intuitive. By performing that action token, viz. not saving the child and not ruining my shoes, I clearly violate my obligation. So to avoid this counter-intuitive consequence, under the open reading (and the assumption that obligation implies permission), saving the child must be my logically weakest permission after all.

Now for the converse direction. For this argument we need an additional assumption, i.e. that some type  $\psi$  is obligatory.<sup>7</sup> We can give again an abstract and a concrete argument. Let us start with the abstract one. Let  $\phi$  be the weakest permission, but assume towards contradiction that  $\phi$  is not obligatory. We know that some other action type  $\psi$  is obligatory, i.e. (by the necessary condition interpretation of obligation) every token that is OK is of type  $\psi$ . Obligation implies permission and the open reading also gives us the other direction, i.e. all OK tokens are  $\psi$ -tokens. Since  $\phi$  is the *weakest permission*, the set of  $\psi$  tokens constitutes a subset of the set of  $\phi$  tokens, i.e. every  $\psi$ -token is a  $\phi$ -token. Hence, every token that is OK is also of type  $\phi$ , which by the necessary condition interpretation of obligation gives us that  $\phi$  is obligatory. Contradiction.

Here is the concrete version. Suppose that saving a child who is drowning is the logically weakest permission that I have, but it is not the case that I ought to do this. As in the previous paragraph, to run this direction of the argument we need to assume that there is something else that I ought to do in that situation. Call it  $\psi$ . We are assuming again that obligation implies permission. So we can conclude that  $\psi$  is permitted too. But since saving the child is my logically weakest permission, there are only two cases to consider.  $\psi$  might be equivalent to saving the child. Then the latter is obligatory after all. Otherwise  $\psi$  is strictly stronger than saving the child. But this is just to say that there are tokens in which I do not  $\psi$  but I save the child. Are these tokens OK? By the open reading they are, but by (a”) they are not. Contradiction.

In the two concrete versions of these arguments we assume that obligated action types are also permitted, that “O implies P”. This is independently plausible. But this assumption takes a particularly strong reading in the present case. We view obligations and permissions as providing, respectively, necessary and sufficient conditions for legality. Necessary conditions need not in general be sufficient. No act token might be morally OK unless it is a token of saving the child. But this doesn’t mean that all “saving the child” tokens are morally OK. Saving the child by sacrificing another one is not, for instance. Assuming that obligation implies permission thus means, in the present setting, that the obligatory action types, i.e. necessary conditions for legality, are also sufficient ones. What is obligatory is to perform a “legal” or an “OK” action. Obligations, under this reading, pinpoint precisely the set of legal action tokens. This is a strong view of obligations. But not an uncommon one. We show some examples now.

**2.3. Examples of obligations as weakest permissions.** We conclude the philosophical part of this paper by providing three examples of cases where obligations are weakest permissions. We encountered two of them above, but it is worthwhile making them more explicit. The first one is a quite mundane case of airport regulation. The second one is the shallow pond example from moral philosophy. The last one involves rational recommendations in games. This last example is in our view particularly compelling. We will be generalizing and illustrating it in Section 4.2.

<sup>7</sup> This assumption is not as strong as it looks. This  $\psi$  could be the trivial action type T. This would merely say that I ought to do something, but nothing specifically.



**Regulative norms of behavior.** According to boarding regulations in many airports, once a passenger has checked in her luggage, she is obliged to board the plane. Now suppose Ann waits at the gate, and has a business class ticket. Many airlines allow passengers like Ann to board at any time while the gate is open. Ann *may* board as soon as the gate is open, she *may* wait until the very last moment, or she *may* proceed anytime in between. But none of these is obligatory. The only thing she *ought* to do is to board according to the regulations. Obeying all the boarding regulations is the strongest necessary condition for a boarding action to be OK. And this is also the weakest permission she has. All other permitted action types for her imply fulfilling this obligation.

**Moral obligations and permissions.** Consider the following textbook example in moral philosophy:

“On my way to give a lecture, I pass a shallow ornamental pond and notice that a small child has fallen in and is in danger of drowning. I look around to see where the parents, or babysitter, are, but to my surprise, I see that there is no one else around. It seems that it is up to me to make sure that the child doesn’t drown. Would anyone deny that I ought to wade in and pull the child out? This will mean getting my clothes muddy, ruining my shoes and either cancelling my lecture or delaying it until I can find something dry to change into; but compared with the avoidable death of a child none of these things are significant.” (Singer, 2011, p. 199)

The agent in this example ought to ensure that the child is prevented from drowning. Is he permitted to take off his jacket first, before wading in and pulling the child out? Only if he does so in a way that does not prevent him saving the child. One should think similarly of other action types that might be permitted in this case. They are all specific ways of, and thus compatible with, saving the child. What should the agent do then? Arguably morality, or most probably utilitarianism in Singer’s case, pinpoints a number of necessary conditions for a saving action to be morally OK. What the agent ought to do here is be moral, i.e. save the child in a way that complies with all other requirements of morality. So in this case there is an “ought” which is the weakest permission too.

**Rational recommendations.** Consider the relatively uncontroversial decision-theoretic principle that agents should *not* choose actions that are strictly dominated by others, viz. actions that make the agent strictly worse off than another in all possible states of the world. In general there will be many non-strictly dominated, mutually exclusive alternatives. To say that the agent ought to do *all* of them would be a blatant violation of the ought-implies-can principle. The more natural reading is that each individual, non-strictly dominated action is permitted. What about obligation? On the ground of dominance alone nothing else distinguishes these non-strictly dominated strategies. They are all rationally permissible, and what the agent ought to do is pick among them. So only the disjunctive action type “either this strategy or that one or...” is obligatory. This is the weakest permission the agent has. We expand on this example in Section 4.2.

**§3. Core system.** We now develop the core deontic logic of obligations as weakest permissions. As mentioned, in this system being the weakest permitted action-type available to the agent is a necessary condition for it being obligated. The deontic logic where the converse direction also holds is studied in Section 4.1. For now we introduce first the syntax

and semantics of our deontic language, then show some basic but interesting properties of the resulting system. We then provide a sound and complete axiom system for that logic.

**3.1. Syntax and semantics.** We work in a propositional modal logic with three “boxes”. Two of them take a deontic interpretation, the third one is an alethic modality to talk about what the agent can and cannot achieve in a given situation. Let PROP be a countable set of propositions. Then  $\mathcal{L}_D$  is the language constructed as follows (with  $p \in \text{PROP}$ ):

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \diamond\varphi \mid P\varphi \mid O\varphi$$

The other propositional connectives are introduced as usual, and  $\square$  is defined as  $\neg\diamond\neg$ . In line with our interpretation presented in the previous section,  $O\varphi$  and  $P\varphi$  should be read as “the agent ought to do a  $\varphi$ -action”, and “every  $\varphi$ -action is permitted”. Under our interpretation of obligation as weakest permission, the two notions are not duals. We will see an example of this later. This is why we introduce them as primitives.  $\diamond\varphi$  is an alethic modality, to be read “it is possible (to ensure) that  $\varphi$ ”. This modality has a dual,  $\square\varphi$ , to be read “it is necessary (to ensure) that  $\varphi$ ”. Formulas of  $\mathcal{L}_D$  are interpreted in *deontic models*:

DEFINITION 3.1 (Frames and models). A deontic frame  $\mathcal{F}$  is a tuple  $\langle H, Alt, n_P, n_O \rangle$  where:

- $H$  is a set of action tokens.<sup>8</sup>
- $Alt$  is the universal relation on  $H$ , i.e.  $Alt = H \times H$ .
- $n_P : H \rightarrow 2^{2^H}$  is a neighborhood function assigning sets of subsets of  $H$  to each  $h \in H$  such that:
  - (Open Reading) If  $X \cup Y \in n_P(h)$  then  $X \in n_P(h)$  and  $Y \in n_P(h)$ .
- $n_O : H \rightarrow 2^{2^H}$  is a neighborhood function assigning sets of subsets of  $H$  to each  $h \in H$  such that:
  - (Ought-Perm) If  $X \in n_O(h)$  then  $X \in n_P(h)$ .
  - (Ought-Can) If  $X \in n_O(h)$  then there is also an  $h' \in X$  such that  $Alt(h, h')$ .
  - (Weakest-Perm) If  $X \in n_O(h)$  then for all  $Y \in n_P(h)$ ,  $Y \subseteq X$ .

A deontic model  $\mathcal{M}$  is a deontic frame together with a valuation  $V : \text{PROP} \rightarrow 2^H$ , assigning to each atomic action type  $p \in \text{PROP}$  a set of action tokens.

We interpret the points in  $H$  as *action tokens* and not as states of the world. When the points are states of the world, sets of them are usually interpreted as events or propositions. Here sets of action tokens are *action types* or properties of action tokens, i.e. the property of being of type  $X \subseteq H$ . The reason we take this non-standard interpretation of the points in our models is that the deontic logic we develop is made to talk about what kind of action types are obligatory and permitted, and the logical relations between these.

The domain  $H$  of a deontic frame can be seen as the *situation* or the *decision problem* the agent is in. It lists the possible action tokens that can be realized. We take different

---

<sup>8</sup> One natural interpretation of action tokens is as the outgoing branches at a particular moment in a model of indeterministic time, e.g. in Belnap *et al.* (2001). Thus our notation  $H$  and the corresponding  $h, h', \dots$  here.



action tokens to be mutually exclusive in a given situation. So an alternative reading of  $Alt(h, h')$  is that either action is *executable* in the current situation.

The deontic modalities are interpreted using the corresponding two *neighborhood functions*  $n_O$  and  $n_P$ . As usual in neighborhood semantics, these functions can be seen as returning, for each state, a list, in this case of action types, that are obligatory and permitted, respectively. For  $n_P$ , the only condition we impose is the one we showed above to be conceptually equivalent to endorsing the open reading of permissions: Free Choice Permission. (Ought-Perm) and the (Ought-Can) are interaction principles. They ensure that obligation implies permission and feasibility, respectively. Note that since  $Alt$  is the universal relation (Ought-Can) can be equivalently formulated as  $\emptyset \notin n_O(h)$ . The key constraint is of course (Weakest-Perm). It makes some  $\varphi$  obligatory *only if* no logically weaker action type is both permitted and feasible. We discuss the converse direction in Section 4.1. Taken all together, this boils down to the following: If  $n_O(h)$  is not empty,<sup>9</sup> then it is a singleton, i.e. it contains *the* weakest permitted action type, and  $n_P(h)$  is the downset, i.e. the set of all subsets of that unique element of  $n_O(h)$ .

Both  $P$  and  $O$  will turn out to be non-normal modalities. Otherwise the interpretation of obligations as weakest permission would be trivial.  $P$  is a modal “box” here. If it was a normal box, it would validate necessitation. So we would always have  $P\top$ . Semantically this would mean that the whole set  $H$  would always be the unique logically weakest permission. Interpreting  $P$  as a normal “diamond” does not work either: A normal diamond is monotone, in the sense that if  $\varphi$  implies  $\psi$ , then  $P\varphi$  implies  $P\psi$ . Given the open reading of permission, this implies that if some  $\varphi$  is permitted, then any arbitrary  $\psi$  is permitted: Assume that  $P\varphi$ . Then, since  $\varphi$  implies  $\varphi \vee \psi$ , we get that  $P(\varphi \vee \psi)$ . By free choice permission, this implies  $P\varphi$  and  $P\psi$ , and so  $P\psi$  in particular. In other words, if  $P$  is a normal diamond, then either everything or nothing is permitted. We need to go non-normal to avoid that.

Truth and validity in deontic models/frames are defined as usual in neighborhood semantics, with the exception that formulas are evaluated “at” action-tokens. One should read  $\mathcal{M}, h \models \varphi$  as “in the situation where  $h$  is a live alternative,  $\varphi$  holds”.<sup>10</sup>

**DEFINITION 3.2 (Truth conditions).** *Let  $\mathcal{M}$  be a deontic model, and  $\|\varphi\|_{\mathcal{M}} = \{h : \mathcal{M}, h \models \varphi\}$ . We omit the subscript to  $\|\varphi\|$  when the context makes it clear. The truth conditions for the Boolean connectives are standard.*

- $\mathcal{M}, h \models \Diamond\varphi$  iff there is an  $h'$  such that  $Alt(h, h')$  and  $\mathcal{M}, h' \models \varphi$ .
- $\mathcal{M}, h \models P\varphi$  iff  $\|\varphi\| \in n_P(h)$ .
- $\mathcal{M}, h \models O\varphi$  iff  $\|\varphi\| \in n_O(h)$ .

*Let  $\mathbf{K}$  be the class of deontic models/frames. Then  $\mathbf{K} \models \varphi$  is the usual notion of validity on a class of models/frames defined as truth in all states of all models.*

<sup>9</sup> Keep in mind that in neighborhood semantics  $n_O(h) = \emptyset$  is a very different condition from  $\emptyset \in n_O(h)$ .

<sup>10</sup> The semantics below thus allow for cases where  $\mathcal{M}, h \models O\varphi$  and  $\mathcal{M}, h' \models \neg O\varphi$ , even though  $Alt(h, h')$ . Call *uniform* frames where such cases do not arise. Failure of uniformity can be seen as an oddity. In this case there is a simple additional frame condition/axiom that enforces uniformity. But such odd cases can be seen as what decision-theorists call “moral hazard” or “state-act dependence.” These are important for decision and game theory, witness the debate between evidential and causal decision theory, and recent work on correlations in games (Brandenburger & Friedenberg, 2008). We show an example of this in Section 4.2. So we keep this possibility open in the core system.

**3.2. Basic properties.** Validities in deontic models already unveil interesting principles relating deontic notions to each other and to alethic modalities. First, as expected from (Ought-Perm), obligation implies permission, but not the other way around. Furthermore, this logic validates neither of the two directions of the usual duality between obligation and permission. Finally, the logic makes room for normative indifference.

**OBSERVATION 3.3.**  $O\varphi \rightarrow P\varphi$  is valid on the class of deontic frames, but not its converse. Furthermore, neither  $O\varphi \rightarrow \neg P\neg\varphi$  nor the converse implication are valid on that class of frames.

*Proof.* The first claim is a trivial consequence of the constraints on  $n_O$ . Here is a counter-model to the validity of  $P\varphi \rightarrow O\varphi$  and  $O\varphi \rightarrow \neg P\neg\varphi$ . Let  $\mathcal{M}$  be a deontic model, with  $H = \{h_1\}$ ,  $Alt = H^2$ , the neighborhood functions  $n_P(h_1) = \{\{h_1\}, \emptyset\}$ ,  $n_O(h_1) = \{\{h_1\}\}$ , and  $V(p) = \{h_1\}$ . This gives us  $\mathcal{M}, h_1 \models O\varphi \wedge P\neg\varphi$  and  $\mathcal{M}, h_1 \models P\neg\varphi \wedge \neg O\neg\varphi$ , and thereby falsifying  $\models O\varphi \rightarrow \neg P\neg\varphi$  and  $\models P\varphi \rightarrow O\varphi$ , respectively. A counter-model for  $\neg P\neg\varphi \rightarrow O\varphi$  can easily be constructed by modifying  $\mathcal{M}$  such that  $n_O(h_1) = n_P(h_1) = \emptyset$ .  $\square$

It is worth pausing to notice the plausibility of the counter-examples to both directions of the dual when obligations are weakest permitted action types. First consider  $\neg P\neg\chi \rightarrow O\chi$ . Take our second example from section 2.3 again, and let  $\varphi$  stand for “you take off your shoes”. Of course, you do not have an obligation to take off your shoes, i.e.  $\neg O\varphi$  (they’re your shoes after all). Does this mean that  $\neg\varphi$  is permitted? Of course not. The open reading of permission requires every  $\neg\varphi$ -token to be OK in order for  $P\neg\varphi$  to be true. But there are many ways in which you do not take of your shoes that are not OK (e.g. the ones where you walk away and do not save the child). So under (Weakest-Perm)  $O$  and  $P$  are not dual, and rightly so.

Plausible counter-examples for the converse are also easy to come by. Obviously, it is consistent for an impossible action  $\psi \wedge \neg\psi$  to be permitted. Given the open reading, this is even very plausible: Since there are no tokens of an impossible action type, every token of that very type is OK. At the same time, this by no means excludes the trivial action  $\psi \vee \neg\psi$  from being obligatory. So this gives us  $O\varphi \wedge P\neg\varphi$  (for at least one  $\varphi$ ), and thereby falsifies the converse direction. The following example illustrates this phenomenon: Suppose that Olivier is on a stretch of the German Autobahn where there is no speed limit, the only thing he can do is drive, and he ought do so. But his car being what it is, he can’t drive faster than 90 km/h. So the situation is the following: as far as driving rules are concerned Olivier can do whatever he wants. The only obligatory action type is the trivial one: driving. But given the limitation of his car the only action tokens available to him are those where he is driving at most at 90 km/h. Driving faster is an impossible action type. Let’s assume for simplicity that there is nothing else that Olivier ought to do. So he ought to drive within the limits of his car’s capacity: let  $\varphi$  be “drive at any speed up to 90km/h”. Then  $O_{Olivier}\varphi$ . But since there is no speed limit  $\neg\varphi$ , to be read “drive above 90km/h”, is permitted too. It’s just that this is not an alternative for Olivier right now.

In the present logic neighborhood semantics has the welcome *side effect* of making the logic immune to some of the deontic paradoxes in their classical formulations. This really is a side effect. Recall that  $P$  and  $O$  are non-normal to avoid trivialization. But this has the further welcome consequence of invalidating three of the paradoxical implications in Standard Deontic Logic, despite the fact that the logic validates the “ $K$ -axiom”, both for  $P$  and for  $O$ .

OBSERVATION 3.4 (Three paradoxes). *None of the following formulae is valid in deontic frames:*<sup>11</sup>

- $O\neg\varphi \rightarrow O(\varphi \rightarrow \psi)$  (contrary to duty paradox)
- $O\varphi \rightarrow O(\varphi \vee \psi)$  (Ross' paradox)
- $(O\varphi \wedge \Box(\varphi \rightarrow \psi)) \rightarrow P\psi$  (Good Samaritan paradox)

*Proof.* The following model  $\mathcal{M}$  provides a counter-example to all three. Take  $H = \{h_1, h_2\}$ , the neighborhood functions  $n_P(h_1) = \{\{h_1\}, \emptyset\}$ ,  $n_P(h_2) = \{\{h_2\}, \emptyset\}$ ,  $n_O(h_1) = \{\{h_1\}\}$ ,  $n_O(h_2) = \{\{h_2\}\}$  and  $V(p) = \{h_1\}$  and  $V(q) = \{h_2\}$ . We obtain  $\mathcal{M}, h_1 \models Op \wedge \neg O(p \vee q)$ , falsifying Ross's paradox;  $\mathcal{M}, h_2 \models O\neg p \wedge \neg O(\neg p \vee \neg q)$ , falsifying Contrary to Duty, and  $\mathcal{M}, h_1 \models Op \wedge \Box(p \rightarrow (p \vee \neg p)) \wedge \neg P(p \vee \neg p)$ , falsifying Good Samaritan. □

Obligations are unique in this logic, up to equivalence. This is reflected by the following validity.

OBSERVATION 3.5. *The following is valid in deontic frames:*

$$O\varphi \wedge O\psi \rightarrow \Box(\varphi \leftrightarrow \psi)$$

*Proof.* Let  $\mathcal{F}$  be a deontic frame and  $h \in H$ . Suppose both  $\|\varphi\|$  and  $\|\psi\|$  are in  $n_O(h)$ . By (Ought-Perm) we know that both are also in  $n_P(h)$ , and by (Weakest-Perm) it follows that  $\|\varphi\| = \|\psi\|$ . □

OBSERVATION 3.6. *The following principles are valid in deontic frames.*

$$P(\varphi \rightarrow \psi) \rightarrow (P\varphi \rightarrow P\psi)$$

$$O(\varphi \rightarrow \psi) \rightarrow (O\varphi \rightarrow O\psi)$$

*Proof.* The validity of the first follows directly from Open Reading and the fact that  $\varphi \rightarrow \psi$  is equivalent to  $\neg\varphi \vee \psi$ . For the second, suppose  $\mathcal{M}, h \models O(\varphi \rightarrow \psi) \wedge O\varphi$ . By Observation 3.5 we have  $\|\varphi\| = \|\neg\varphi \vee \psi\|$ . By Ought-Can we also know that  $\|\varphi\| \neq \emptyset$ . But then since  $\|\neg\varphi\| \subseteq \|\neg\varphi \vee \psi\| = \|\varphi\|$ , it must be that  $\|\neg\varphi\| = \emptyset$ . This means that  $\|\neg\varphi \vee \psi\| = \|\psi\|$ , and so that  $\mathcal{M}, h \models O\psi$ . □

Observations 3.5 and 3.6 highlight the fact that the distribution of  $O$  over implication hinges on an assumption *about the alternatives*. Obligations as weakest permission thus give rise to an unusual interplay between deontic and alethic modalities. This shouldn't come as a surprise. In our interpretation an action type is obligatory only if it is the weakest possible permitted *feasible* action type.

Call an action type  $\varphi$  a *matter of indifference* when neither  $\varphi$  nor its negation are obligatory (McNamara, 2014). Deontic models can capture this form of normative indifference. In fact they can capture an even stronger form of indifference. They allow for cases where the normative code is silent on a given type  $\varphi$ . Neither  $\varphi$  nor its negation are obligatory. It could be that neither of them are permitted either.

OBSERVATION 3.7. *The following is not valid in deontic frames:*

$$O\varphi \vee O\neg\varphi \vee P\varphi \vee P\neg\varphi$$

<sup>11</sup> Note that under this formulation Contrary to Duty is logically equivalent to Ross' paradox. We nonetheless include both here to highlight the fact that they are not valid.

*Proof.* Take a deontic model where  $H = \{h_1, h_2, h_3\}$ ,  $V(p) = \{h_1\}$ ,  $V(q) = \{h_3\}$ ,  $n_O(h_1) = V(q)$ ,  $n_P(h_1) = \{\emptyset, V(q)\}$ . At  $h_1$  the formula above is false for  $\|\varphi\| = V(p)$ .  $\square$

Situations like this are plausible. Consider the case of “sharing” music in the early days of the internet. This was a relatively new action type. It took time for it to be included in national copyright laws. Until then sharing was not only a matter of indifference in the sense that it was neither obligatory nor forbidden; neither sharing nor not sharing was explicitly permitted.

**3.3. Axiomatization.** The set of valid formulas on deontic frames is completely axiomatizable by the axioms and rules in Table 1.

Table 1. <i>The axiom system 5HD. Here D is either O or P</i>	
1. All propositional tautologies and S5 for $\diamond$ , together with:	
<b>(Univ)</b> $\Box(\varphi \leftrightarrow \psi) \rightarrow (D\varphi \leftrightarrow D\psi)$	
2. For P:	
<b>(FCP)</b> $P(\psi \vee \varphi) \rightarrow P\varphi \wedge P\psi$	
3. Interaction axioms:	
<b>(Ought-Perm)</b> $O\varphi \rightarrow P\varphi$	
<b>(Ought-Can)</b> $O\varphi \rightarrow \diamond\varphi$	
<b>(Weakest-Perm)</b> $O\varphi \rightarrow (P\psi \rightarrow \Box(\psi \rightarrow \varphi))$	
4. Rules: Modus Ponens and	
	$\text{(NEC)} \frac{\vdash \varphi}{\vdash \Box\varphi}$

The reader accustomed to neighborhood semantics might wonder why the usual rule “E” has been omitted.

$$\text{(E)} \frac{\vdash \varphi \leftrightarrow \psi}{\vdash D\varphi \leftrightarrow D\psi}$$

This rule is easily derivable in 5HD: Apply (NEC) to  $\vdash \varphi \leftrightarrow \psi$ . Then  $\vdash D\varphi \leftrightarrow D\psi$  follows from (Univ) by Modus Ponens. We now turn to completeness.<sup>12</sup>

**THEOREM 3.8.** *The logic 5HD (Table 1) is sound and complete with respect to the class of deontic frames.*

*Proof.* For soundness the only non-standard axiom is (Weakest-Perm). From  $\mathcal{M}, h \models O\varphi$  we know that  $\|\varphi\| \in n_O(h)$ , but then for any  $\psi$  such that  $P\psi$  is true at  $h$ , the Weakest-Perm frame condition gives us that  $\|\psi\| \subseteq \|\varphi\|$ , and so that  $\Box(\psi \rightarrow \varphi)$  is also true at  $h$ .

We now turn to completeness. A routine argument shows that every 5HD-consistent set of formulas can be extended to a maximally consistent set (MCS). Take such a maximally consistent set  $\Gamma$ . Write  $H^\Delta$  for the set of all maximally consistent sets  $\Delta$  such that for all formulas  $\psi$  of the form  $\Box\varphi$ , we have that  $\psi \in \Gamma$  iff  $\psi \in \Delta$ .  $H^\Delta$  will be the domain of the model which will be built.

<sup>12</sup> We use standard techniques (maximal consistency, canonical models, etc.) from modal logic, see Blackburn *et al.* (2002, pp. 190ff) for details.

DEFINITION 3.9 (Neighborhood tools). *The proof set  $|\varphi|$  of a formula  $\varphi$  of  $\mathcal{L}_D$  is defined as  $\{\Delta \in H^\Lambda : \varphi \in \Delta\}$ . A neighborhood function  $n^D$  on  $H^\Lambda$  is canonical for 5HD if for all  $\varphi$ ,  $|\varphi| \in n^D(\Delta)$  iff  $D\varphi \in \Delta$ , for  $D$  either  $O$  or  $P$ . A minimal  $D$ -neighborhood function  $n_{min}^D : H^\Lambda \rightarrow 2^{2^{H^\Lambda}}$  for a modality  $D$  is defined as:*

$$n_{min}^D(\Delta) = \{|\varphi| : D\varphi \in \Delta\}$$

Let  $n_{min}^O$  be the minimal function for  $O$ , and  $n_{\downarrow min}^P$  the closure under subsets of the minimal function for  $P$ .

FACT.  $n_{\downarrow min}^P$  and  $n_{min}^O$  are well-defined and canonical.

*Proof.* Canonicity is trivial for  $n_{min}^O$  and for the right to left direction of  $n_{\downarrow min}^P$ . For the left to right, the only case to check is whether  $|\varphi|$  could have been added to  $n_{\downarrow min}^P(\Delta)$  by closing under subsets, without  $P\varphi$  being in  $\Delta$ . This cannot happen. Suppose that  $|\varphi| \in n_{\downarrow min}^P(\Delta)$  and  $|\varphi| \subset |\psi|$  for  $P\psi \in \Delta$ . We know that  $\psi$  is equivalent to  $(\psi \wedge \varphi) \vee (\psi \wedge \neg\varphi)$ . Using E and FCP, we get that  $P(\varphi \wedge \psi) \in \Delta$ . But since  $|\varphi| \subset |\psi|$  we can show using Lemma 3.12 that  $\Box((\varphi \wedge \psi) \leftrightarrow \varphi) \in \Delta$ , and so that  $P\varphi \in \Delta$  using E again. Well-definedness follows from Lemma 3.12, using (Univ) and the usual properties of MCSs. □

DEFINITION 3.10 (Canonical model). *The canonical model  $\mathcal{M}^\Lambda = \langle H^\Lambda, Alt^\Lambda, n_{min}^O, n_{\downarrow min}^P, V^\Lambda \rangle$  is defined as usual:  $H^\Lambda$ ,  $n_{\downarrow min}^P$  and  $n_{min}^O$  are as above.  $Alt^\Lambda$  is the canonical relation for  $\diamond$ .  $V^\Gamma$  is the canonical valuation:  $V(p) = \{\Gamma : p \in \Gamma\}$ .*

The Existence Lemma for  $\diamond$  is routine.

LEMMA 3.11 (Existence lemma for  $O$  and  $P$ ). *If  $P\varphi \in \Delta$  then  $|\varphi| \in n_{\downarrow min}^P(\Delta)$ , and similarly for  $O$ .*

*Proof.* For  $O$  and  $P$  it follows trivially from the definitions of  $n_{\downarrow min}^P$  and  $n_{min}^O$ . □

LEMMA 3.12.  $|\varphi| \subseteq |\psi|$  if and only if  $\Box(\varphi \rightarrow \psi) \in \Gamma$ , for all  $\Gamma \in H^\Lambda$

*Proof.* The “if” direction follows directly from the T axiom. For the “only if”, suppose there is  $\Gamma$  such that  $\Box(\varphi \rightarrow \psi) \notin \Gamma$ . Then  $\diamond(\varphi \wedge \neg\psi) \in \Gamma$  because  $\Gamma$  is an MCS. But then the Existence Lemma for  $\diamond$  gives us a  $\Delta \in H^\Lambda$  such that  $\varphi \wedge \neg\psi \in \Delta$ , and thus  $|\varphi| \not\subseteq |\psi|$ . □

LEMMA 3.13 (Truth lemma). *For all  $\varphi \in \Gamma$ ,  $\mathcal{M}^\Lambda, \Delta \models \varphi$  iff  $\varphi \in \Delta$ .*

*Proof.* As usual, by induction on  $\varphi$ . All cases are standard. □

All that remains to be shown is that  $\mathcal{M}^\Lambda$  is a deontic model.

LEMMA 3.14. *For all  $\Delta \in H^\Lambda$ , if  $|\varphi| \in n_{min}^O(\Delta)$  then for all  $Y \in n_{\downarrow min}^P(\Delta)$ ,  $Y \subseteq |\varphi|$ .*

*Proof.* Suppose  $|\varphi| \in n_{min}^O(\Delta)$ . Take  $Y \in n_{\downarrow min}^P(\Delta)$  arbitrary. By the definition of  $n_{\downarrow min}^P$ , either  $Y = |\psi|$  or  $Y \subsetneq |\psi|$  for some  $P\psi \in \Delta$ . From Weakest-Perm we thus can conclude that  $\Box(\psi \rightarrow \varphi) \in \Delta$ . But then Lemma 3.12 gives us that  $Y \subseteq |\psi| \subseteq |\varphi|$ , as wanted. □

CLAIM 3.15.  $\mathcal{M}^\Lambda$  is a deontic model.

*Proof of Claim.* It follows from the construction of the canonical model that  $Alt^\Delta$  is the universal relation.<sup>13</sup> Given Lemma 3.14, there are only three things that still need to be checked. First, that if  $X \in n_{min}^O(\Delta)$  then  $X \in n_{min}^P(\Delta)$ . This follows directly from Ought-Perm. Second, that if  $X \in n_{min}^O(\Delta)$  then there is a  $\Delta' \in X$  such that  $Alt^\Delta(\Delta, \Delta')$ . This follows directly from (Ought-Can) and the Existence Lemma. Finally, we need to show that if  $X \cup Y \in n_{min}^P(\Delta)$  then both  $X, Y \in n_{min}^P(\Delta)$ . This follows directly from FCP, the properties of MCS, and the fact that  $n_{min}^P$  is closed under subsets.  $\square$

This concludes the proof of Theorem 1.  $\square$

**§4. Extensions and applications.** In this section we first consider extensions of the core system where being the weakest permitted action type is sufficient for it to be obligated. Then we expand on the last example presented in section 2.3, namely recommendations and permissions stemming from decision- and game-theoretic rationality notions.

**4.1. Extensions to “if and only if”.** In certain cases, being the weakest permitted action-type is sufficient for it being obligatory (c.f. Section 2.2). Technically, this amounts to requiring the following:

$$\text{If } X \in n_P(h) \text{ and } Y \subseteq X \text{ for all } Y \in n_P(h) \text{ then } X \in n_O(h) \quad (\text{Conv})$$

This “converse direction” conjures a pattern of second-order quantification that is different from the one modal logicians are used to (Blackburn *et al.*, 2002, Sec. 3.2). These are generally of the form:

$$\forall P_1, \dots, P_n \varphi(P_1, \dots, P_n)$$

(Conv) is not of that form.<sup>14</sup> Slightly abusing our notation for the sake of illustration, this condition can be read as:

$$\forall S(\forall Q(P(Q) \rightarrow \Box(Q \rightarrow S)) \rightarrow OS)$$

which is equivalent to

$$\forall S \exists Q((P(Q) \rightarrow \Box(Q \rightarrow S)) \rightarrow OS)$$

So to axiomatize this frame condition one needs more powerful tools. Here we use the following infinitary inference rule. Let  $E = \langle p_1, p_2, \dots \rangle$  be an enumeration of all atomic propositions in PROP, then:

$$\frac{\vdash Pp_1 \rightarrow \Box(p_1 \rightarrow \varphi) \quad \vdash Pp_2 \rightarrow \Box(p_2 \rightarrow \varphi) \quad \dots \quad \text{for all } p_i \in E}{\vdash P\varphi \rightarrow O\varphi} \quad (\text{R-Conv})$$

<sup>13</sup> We restricted  $H^\Delta$  to MCS that have the same “Box” theory, and it is a standard observation that this makes  $Alt^\Delta$  the universal relation in that model, and hence  $\Box$  a universal modality in that model. Take any two MCS  $\Gamma$  and  $\Delta$ . We have to show that  $Alt^\Delta(\Gamma, \Delta)$ . Take any  $\phi \in \Delta$ . By T we get  $\Diamond\phi \in \Delta$ , and so  $\Box\neg\phi \notin \Delta$ . By construction for all formulas of the form  $\Box\psi, \Box\psi \in \Delta$  iff  $\Box\psi \in \Gamma$ . So  $\Box\neg\phi \notin \Gamma$ , as required.

<sup>14</sup> Observe that (Weakest-Perm) is, after some quantifier shifting. Slightly abusing our notation, its condition corresponds to:

$$\forall S(O(S) \rightarrow \forall Q(P(Q) \rightarrow \Box(Q \rightarrow S)))$$

which is equivalent to

$$\forall S \forall Q(O(S) \rightarrow (P(Q) \rightarrow \Box(Q \rightarrow S)))$$

This equivalence is at the root of the completeness result just presented.



This rule is sufficient to ensure weak completeness with respect to the intended class of frames. See below. We do not know whether it is strongly complete. In any case the resulting logic is not compact. Take the set of formulas above the line of (R-Conv), together with the negation of the conclusion of that rule. This whole set is inconsistent, but none of its finite subsets is.

**THEOREM 4.1.** *Let  $5HD^+$  be the logic  $5HD$  (page 818) extended with (R-Conv).  $5HD^+$  is sound and weakly complete with respect to the class of uniform deontic frames that satisfy (Conv).*

*Proof.* We show first that (R-Conv) preserves validity. Suppose that for a given frame  $\mathcal{F}$  there is a valuation  $V$  and a  $h$  such that  $\mathcal{F}, V, h \models P\varphi \wedge \neg O\varphi$ . By (Conv) we know that there is a  $Y \in n_P(h)$  such that  $Y \not\subseteq \|\varphi\|$ . So there is a  $h' \in Y$  such that  $h' \notin \|\varphi\|$ . Take an atom  $p$  that does not occur in  $\varphi$  and a valuation  $V'$  that is identical to  $V$  except that  $V(p) = Y$ . This gives us  $\mathcal{F}, V', h \models Pp \wedge \diamond(p \wedge \neg\varphi)$ , as required.

Now for completeness. In the presence of the (R-Conv) we need to re-do the Lindenbaum Lemma.

**LEMMA 4.2.** *Every  $5HD^+$ -consistent finite set of formulas  $\Sigma$  can be extended to an MCS  $\Sigma^*$ .*

*Proof.* Let  $S = \langle \varphi_0, \dots \rangle$  be an enumeration of formulas of  $\mathcal{L}_D$ . Then  $\Sigma^*$  is inductively defined as follows. Throughout the induction, let  $At_n$  be the set of all atoms in PROP that are not in  $\Sigma_n \cup \{\varphi_n\}$ . Since  $\Sigma$  is finite  $At_n$  is countable infinite, for all  $n$ .

1.  $\Sigma_0 = \Sigma$ .
2.  $\Sigma_{n+1}$ . Take  $\varphi_n \in S$ . There are two cases.
  - (a)  $\Sigma_n \not\vdash \neg\varphi_n$ . What to do then depends on the shape of  $\varphi_n$ .
    - i If  $\varphi_n = P\psi \rightarrow \Box(\psi \rightarrow \varphi)$  then fix
 
$$\Sigma_{n+1} = \Sigma_n \cup \{\varphi_n\} \cup \{p^*\} \cup \{\chi^*\}$$
 where  $\chi^* = (\neg O\varphi \wedge P\varphi \wedge p^*) \rightarrow (Pp^* \wedge \diamond(p^* \wedge \neg\varphi))$  and  $p^* \in At_n$ .
      - ii O/w then fix
 
$$\Sigma_{n+1} = \Sigma_n \cup \{\varphi_n\}$$
    - (b)  $\Sigma_{n+1} := \Sigma_n \cup \neg\varphi_n$  otherwise.

Set  $\Sigma^* = \bigcup_{n < \omega} \Sigma_n$ .

**CLAIM 4.3.** *For all  $n \in \mathbb{N}$ ,  $\Sigma_n$  is consistent.*

*Proof.*  $\Sigma$  is consistent by assumption. Consider  $\Sigma_{n+1}$ . The only non-standard case is 2(a), and the inconsistency must come from the fact that  $\Sigma_n \vdash \neg\chi^*$ . But unless  $\Sigma_n$  is already inconsistent, this can't be. For in that case we would have  $\Sigma_n \vdash p^*$ . But this can't be since  $p^*$  does not occur in  $\Sigma_n$ . □

**CLAIM 4.4.**  *$\Sigma^*$  is an MCS.*

*Proof.* Maximality follows directly by construction. So now for consistency. Every finite subset of  $\Sigma^*$  is consistent, by the previous claim. So the only inconsistency that can arise is by taking  $\bigcup_{n < \omega} \Sigma_n$ , and this could only come from (R-Conv), which uses an infinite set of premises. So suppose that  $\Sigma^*$  is not consistent. This means that there is a formula  $\varphi$  such

that  $Pp \rightarrow \Box(p \rightarrow \varphi) \in \Sigma^*$  for all  $p \in \text{PROP}$ , but that  $\Sigma^*$  also contains  $P\varphi \wedge \neg O\varphi$ . By construction it must be that  $P\varphi \wedge \neg O\varphi$  has been added at some stage  $\Sigma_n$ . Now consider an atom  $q \in At_n$ . Then the formula  $Pq \rightarrow \Box(q \rightarrow \varphi)$  must have been added at some later stage  $m > n$ , and with it both  $p^*$  and  $(\neg O\varphi \wedge P\varphi \wedge p^*) \rightarrow (Pp^* \wedge \Diamond(p^* \wedge \neg\varphi))$  for some  $p^* \in At_m$ . But then by the choice of this  $p^*$ ,  $\Sigma_m \vdash Pp^* \wedge \Diamond(p^* \wedge \neg\varphi)$  and thus the formula  $Pp^* \rightarrow \Box(p^* \rightarrow \varphi)$  could not have been added later on, against our assumption.  $\square$

This proves the Lemma.  $\square$

From there we can use the same canonical model construction as in the proof of Theorem 3.8. All that remains to be shown is that this model satisfies (Conv). Take an MCS  $\Gamma$  and suppose that, first,  $X \in n_P(\Gamma)$ . This means that  $X = |\varphi|$  for some  $P\varphi \in \Gamma$ . Suppose also that  $Y \subseteq X$  for all  $Y \in n_P(\Gamma)$ . Take an arbitrary  $p \in \text{PROP}$ . If  $Pp \notin \Gamma$  then  $Pp \rightarrow \Box(p \rightarrow \varphi) \in \Gamma$ . Suppose then that  $Pp \in \Gamma$ . By Lemma 3.12 we get  $\Box(p \rightarrow \varphi) \in \Gamma$ , and so we have  $Pp \rightarrow \Box(p \rightarrow \varphi) \in \Gamma$  in that case too. But then since MCSs are closed under inference rules, we get that  $O\varphi \in \Gamma$ , which is sufficient to prove the theorem.  $\square$

**4.2. Applications to decision and game theory.** In this section we apply the logic of obligations as weakest permissions to norms of practical rationality in decision and game theory. We provide first a philosophical argument to the effect that this reading of deontic modals is the correct one for such norms. A welcome consequence of this view is that the logic of practical recommendations in decision and game theory is immune to the classical deontic paradoxes. We illustrate this by two well-known examples, and sketch how to represent practical recommendations in deontic models. This section is mainly philosophical, and repeats some of the arguments from Section 2, but it builds on the technical apparatus developed in the meanwhile.

First some terminology.<sup>15</sup> By a *decision-theoretic choice rule* we mean a rule that outputs a set of actions for each agent and each decision problem. Maximization of expected utility is the classical example. Solution concepts in game theory can also be seen in this abstract way. The main difference is that, in general, solution concepts output sets of outcomes/strategy profiles for each game. These sets of outcomes need not correspond to strategies for each player. For equilibrium concepts this is not the case, while for iterated elimination of strictly dominated strategies it is. The classical normative interpretation of choice rules and solution concepts is that they pinpoint the set of rational actions and/or outcomes, write it R, and that agents ought to play rational actions or outcomes (OR).

The general argument on page 813 can now be applied to the present case. The claim is that the deontic logic of practical recommendations stemming from choice rules and solution concepts is the logic of obligation as weakest permission.

Let us start with obligations. Take a choice rule or a solution concept, and suppose that it partitions the set of available actions or possible outcomes into two sets, the rational (R) and irrational ones ( $\neg R$ ).<sup>16</sup> Under the normative interpretation, the agent(s) ought to choose rationally, so we have (OR). Choosing an irrational action is not OK. So we are under reading (a') (Section 2.1.2). Now we argue that R is the unique obligatory action type. The crux of this argument is that, *on rationality grounds alone*, no specific rational action in

<sup>15</sup> We remain informal on the relevant decision and game-theoretical notions. See e.g. Rasmusen (2007) for the formal definitions.

<sup>16</sup> Most if not all standard choice rules and solution concepts are like that. There are exceptions, c.f. Cubitt & Sugden (2011), but they do not compromise the generality of our argument.

R is “more rational” than any other. The same holds for action types that imply R. So there is no logically stronger action type that is obligatory. What about logically weaker types than R? The same argument as in Section 2.2 can be run here. We assume that obligation implies permission. So under the open reading of permissions, for which we will argue presently, if there is a  $\varphi$  such that  $R \subset \varphi$  such that  $O\varphi$  then there is an irrational but nonetheless rationally permitted alternative, a direct contradiction. So R is the unique action type which is obligatory under the choice rule or solution concept at hand.

Now for permissions. By the same argument as above R is permitted, and nothing logically weaker is. Now we claim that the right reading of this permission is the open reading. Again, the key point is that, on the ground of rationality alone, all actions in R are equally recommendable. So it cannot be the case that some are rationally allowed while others are not. The agent can thus pick any of them, or any action type that implies R. But then R must be the logically weakest permission the agent has. In other words, the only thing that a choice rule or a solution concept prescribes is to play what it pinpoints as a rational strategy. Within these limits, the agent has free choice permission. She may rationally choose any action.

Let us take two examples to illustrate this. First, one from decision theory. Consider the “Medical Newcomb’s problem” illustrated in Table 2.<sup>17</sup> In that story smoking is positive evidence that you have the Cancer Gene. But according to standard causal decision theory this type of evidential correlation should be ignored. You either have the gene or you do not, whatever you do. So smoking is the only rational thing to do. But if there are many ways to smoke, in that situation rationality does not provide any way to distinguish between them. So they are all rationally permitted. For similar reasons rationality recommends not choosing any action that would prevent you from smoking. So smoking is the logically weakest permitted action-type, and this is indeed what one ought to do.

Table 2. *Smoking and the cancer gene*

	No Cancer Gene	Cancer Gene
Not Smoking	2	-1
Smoking	4	0

One can make this more precise using deontic frames. Suppose we have a given set of action tokens  $\tau_1, \dots, \tau_n$ . This is going to be our set of alternatives  $H$ . Suppose that  $H$  can be partitioned into  $S$  and  $\bar{S}$ , the smoking and the non-smoking tokens. For all  $h \in H$ , fix  $n_P(h)$  to be the set of all subsets of  $S$ , and  $n_O(h) = \{S\}$ . This is a deontic frame, and an easy check reveals that it provides counter-examples to the three deontic paradoxes mentioned in Observation 3.4 (page 817).

Now for a game-theoretic example. Take Matching Pennies (Table 3). This is a typical example where there are no pure-strategy Nash equilibria. Now suppose that Ann (row) and Bob (column) are playing this game, and that their decisions are positively correlated

<sup>17</sup> See Weirich (2012) for further references and more details.

Table 3. *The matching pennies game*

	T	H
T	(-1,1)	(1,-1)
H	(1,-1)	(-1,1)

with one another. Say they are twins.<sup>18</sup> Ann choosing *T* is strong evidence for her that Bob will do so, and similarly for *H* and vice-versa for Bob. This puts each of them up against a decision problem where there is no rati-fiable or deliberatively stable action (Jeffrey, 1965; Joyce, 2012). If Ann chooses *T* then Bob will choose *T* too, so Ann's best response should have been *H*. But if she chooses *H* then Bob will do the same, and so her best response brings her back to choosing *T*. If Ann chooses an action *A* in that case, she ought to have chosen otherwise ( $\neg A$ ), and this obligation is the weakest permitted action she has.

This sort of example can also be captured in deontic frames. Take again  $H = \tau_1, \dots, \tau_n$  to be the set of action tokens available to, say, Ann, and suppose that this set can be partitioned into *Head* and *Tail*, the "playing heads" and "playing tails" tokens. The conundrum in which Ann finds herself corresponds to the fact that if  $h \in \text{Head}$  then  $n_P(h) = \{X : X \subseteq \text{Tail}\}$  and  $n_O(h) = \{\text{Tail}\}$ . And vice-versa if  $h \in \text{Tail}$ :  $n_P(h) = \{X : X \subseteq \text{Head}\}$  and  $n_O(h) = \{\text{Head}\}$ . Again, these sets of obligations and permissions are not prone to the deontic paradoxes mentioned above.

It should be clear that the construction in these two examples generalizes. All that it requires is a partition of *H* into rational and irrational strategies according to a given choice rule. So given a decision problem, be it decision or game theoretic, and a choice rule, one can construct a deontic frame that reflects rational obligations and permissions. Together with the philosophical argument given above, this means that the general deontic logic of rational recommendations in game and decision theory is the non-normal logic that we have studied in this paper, with the welcome corollary that it avoids most of the known deontic paradoxes.<sup>19</sup>

**§5. Conclusion.** This paper is about the logic of a non-standard understanding of deontic modals, where an action type is obligatory (if and) only if it is permitted and no logically weaker action type is permitted. We argued that this reading is philosophically plausible, and has interesting mathematical properties. In the core system, being the weakest permission is only necessary for an action type to be obligatory. We showed, among other things, that this system gives rise to an unusual interplay between deontic and alethic modalities and that it avoids some of the classical deontic paradoxes. We then studied an extension of this system, where being the weakest permission is also sufficient for an action type to be obligatory, and have shown that this understanding of obligations and permissions applies naturally to practical norms in decision and game theory.

There exist some congenial proposals in the deontic logic literature, starting from van Benthem (1979). The work of Trypuz & Kulicki (2011), Trypuz & Kulicki (2013), Trypuz & Kulicki (2009) and Czelakowski (1997) are prime examples. The first step to

<sup>18</sup> We use this specific example because it shows that some "non-uniform" frames are philosophically interesting. See footnote 10 (page 815) for the definition of uniformity.

<sup>19</sup> This point is developed in more detail in Roy *et al.* (2014).

establish a formal connection has been made in Dong (2014). Under a different interpretation, the present logic could be used to study the relation between *prima facie* and “all things considered” obligations (Donagan, 1984; Brink, 1994; Horty, 2012). The former would be the *Ps* in our system, the latter the *Os*. The conjecture is that the present logic, minus Free Choice, would come close to the “disjunctive account” (Horty, 2012) of all things considered obligations. Under that reading, all things considered obligations would arise from the requirement that even the weakest *prima facie* obligations do not conflict with it. Finally, it would be interesting to add a temporal component to the present logic, and study its relation with *stit* theory.

Some technical questions are still left open here. We have not studied the complexity of the core or the extended system. Some advances have been made in the study of the proof theory of a system close to 5HD (Gratzl, 2013; Lellmann, 2014), but for now it is still open whether 5HD has a proof system that admits Cut. We also left open the question of a strongly complete axiomatization of the “converse direction”. Finally, the interplay between deontic logic and game theory is an active area of research. Some steps towards the connection of our proposal with this literature have been made in Roy *et al.* (2014), but at the time of writing this is still very much an ongoing project.

**§6. Acknowledgments.** This research has been supported by the Alexander von Humboldt Foundation and by an LMU Research Fellowship as part of the LMU Academic Career Program and the Excellence Initiative. We would like to thank the two anonymous referees of this journal for extremely useful comments on earlier versions of this paper. We would further like to thank H. Leitgeb, G. Schurz, H. van Ditmarsch, J. van Benthem, O. Foisch, R. Trypuz, P. Kulicki, H. Dong, audiences at various seminars in Munich and Salzburg, and participants of Trends in Logic XI (Bochum, 2012) and Non-classical epistemic logic (Prague, 2012).

#### BIBLIOGRAPHY

- Anglberger, A., Dong, H., & Roy, O. (2014). Open reading without free choice. In Cariani, F., Grossi, D., Meheus, J., and Parent, X., editors. *Deontic Logic and Normative Systems*. Heidelberg: Springer, pp. 19–32.
- Asher, N., & Bonevac, D. (2005). Free choice permission is strong permission. *Synthese*, **145**(3), 303–323.
- Belnap, N. D., Perloff, M., & Xu, M. (2001). *Facing the Future: Agents and Choices in our Indeterminist World*. New York: Oxford University Press on Demand.
- van Benthem, J. (1979). Minimal deontic logics. *Bulletin of the Section of Logic*, **8**(1), 36–41.
- Blackburn, P., De Rijke, M., & Venema, Y. (2002). *Modal Logic*. Cambridge, UK: Cambridge University Press.
- Brandenburger, A., & Friedenberg, A. (2008). Intrinsic correlation in games. *Journal of Economic Theory*, **141**(1), 28–67.
- Brink, D. (1994). Moral conflict and its structure. *The Philosophical Review*, **103**(2), 215–247.
- Broersen, J. (2004). Action negation and alternative reductions for dynamic deontic logics. *Journal of Applied Logic*, **2**, 153–168.
- Brown, F. M. (1987). *The Frame Problem in Artificial Intelligence*. Los Altos: Morgan Kaufmann Publishers Inc.
- Cubitt, R. P., & Sugden, R. (2011). Common reasoning in games: A lewisian analysis of common knowledge of rationality. Technical report, CeDEx discussion paper series.

- Czelakowski, J. (1997). Action and deontology. In Ejerhed, E., and Lindström, S., editors. *Logic, Action and Cognition*. New York: Springer, pp. 47–87.
- Donagan, A. (1984). Consistency in rationalist moral systems. *The Journal of Philosophy*, **81**(6), 291–309.
- Dong, H. (2014). A survey between 5wp and dal. Unpublished manuscript, University of Bayreuth.
- Gratzl, N. (2013). Sequent calculi for multi-modal logic with interaction. In D. Grossi, O. Roy, and H. Huang, editors. *Logic, Rationality, and Interaction*. Heidelberg: Springer, pp. 124–134.
- Hansson, S. (2013). The varieties of permissions. In Gabbay, D., Horty, J., Parent, X., van der Meyden, R., and van der Torre, L., editors. *Handbook of Deontic Logic and Normative Systems*. London: College Publications.
- Horty, J. (2012). *Reasons as Defaults*. Oxford, New York: Oxford University Press.
- Jeffrey, R. (1965). *The Logic of Decision*. New-York: McGraw-Hill.
- Joyce, J. M. (2012). Regret and instability in causal decision theory. *Synthese*, **187**(1), 123–145.
- Kamp, H. (1973). Free choice permission. *Proceedings of the Aristotelian Society*, **74**, 57–74.
- Lellmann, B. (2014). Axioms vs hypersequent rules with context restrictions: Theory and applications. In Demri, S., Kapur, D., and Weidenbach, C., editors. *Automated Reasoning*, Volume 8562 of Lecture Notes in Computer Science. Springer International Publishing, pp. 307–321. [http://dx.doi.org/10.1007/978-3-319-08587-6\\_23](http://dx.doi.org/10.1007/978-3-319-08587-6_23).
- Lewis, D. (1979). A problem about permission. In Saarinen, E. et al., editors. *Essays in honour of Jaakko Hintikka*. Dordrecht: Reidel, pp. 163–175.
- Makinson, D. (1984). Stenius' approach to disjunctive permission. *Theoria*, **50**(2–3), 138–147.
- McCarty, J. (1980). Circumscription – a form of nonmonotonic reasoning. *Artificial Intelligence*, **13**(1–2), 27–39.
- McNamara, P. (2014). Deontic logic. In Zalta, E. N., editor. *The Stanford Encyclopedia of Philosophy* (Winter 2014 edition). <http://plato.stanford.edu/archives/win2014/entries/logic-deontic/>.
- Rasmusen, E. (2007). *Games and Information* (fourth Edition). Oxford: Blackwell.
- Roy, O., Anglberger, A., & Gratzl, N. (2014). The logic of best action from a deontic perspective. In Baltag, A., and Smets, S., editors. *Johan F.A.K. van Benthem on Logical and Informational Dynamics*. Berlin: Springer, pp. 657–676.
- Singer, P. (2011). *Practical Ethics* (Third ed.). Cambridge: Cambridge University Press.
- Trypuz, R., & Kulicki, P. (2009). A systematics of deontic action logics based on boolean algebra. *Logic and Logical Philosophy*, **18**(3–4), 253–270.
- Trypuz, R., & Kulicki, P. (2011). A norm-giver meets deontic action logic. *Logic and Logical Philosophy*, **20**, 59–72.
- Trypuz, R., & Kulicki, P. (2013). On deontic action logics based on boolean algebra. *Journal of Logic and Computation*.
- Weirich, P. (2012). Causal decision theory. In Zalta, E. N., editor. *The Stanford Encyclopedia of Philosophy* (Winter 2012 edition). <http://plato.stanford.edu/archives/win2012/entries/decision-causal/>.
- Wright, G. H. v. (1963). *Norm and Action - A Logical Enquiry*. London: Routledge.
- Wright, G. H. v. (1968). *An Essay in Deontic Logic and the General Theory of Action*. Amsterdam: North-Holland Publishing Company.



MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY  
LMU MUNICH  
FAKULTÄT FÜR PHILOSOPHIE, WISSENSCHAFTSTHEORIE  
UND RELIGIONSWISSENSCHAFT  
MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY  
GESCHWISTER-SCHOLL-PLATZ 1  
D-80539 MÜNCHEN  
*E-mail:* albert@anglberger.org

MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY  
LMU MUNICH  
FAKULTÄT FÜR PHILOSOPHIE, WISSENSCHAFTSTHEORIE  
UND RELIGIONSWISSENSCHAFT  
MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY  
GESCHWISTER-SCHOLL-PLATZ 1  
D-80539 MÜNCHEN  
*E-mail:* Norbert.Gratzl@lrz.uni-muenchen.de

UNIVERSITÄT BAYREUTH  
DEPARTMENT OF PHILOSOPHY  
D-95440 BAYREUTH  
*E-mail:* Olivier.Roy@uni-bayreuth.de