



## Action Type Deontic Logic

**Bentzen, Martin Mose**

*Published in:*  
Journal of Logic, Language and Information

*Link to article, DOI:*  
[10.1007/s10849-014-9205-0](https://doi.org/10.1007/s10849-014-9205-0)

*Publication date:*  
2014

*Document Version*  
Early version, also known as pre-print

[Link back to DTU Orbit](#)

*Citation (APA):*  
Bentzen, M. M. (2014). Action Type Deontic Logic. *Journal of Logic, Language and Information*, 23(4), 397-414.  
<https://doi.org/10.1007/s10849-014-9205-0>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Action Type Deontic Logic

Martin Mose Bentzen

the date of receipt and acceptance should be inserted later

**Abstract** A new deontic logic, *Action Type Deontic Logic*, is presented. To motivate this logic, a number of benchmark cases are shown, representing inferences a deontic logic should validate. Some of the benchmark cases are singled out for further comments and some formal approaches to deontic reasoning are evaluated with respect to the benchmark cases. After that follows an informal introduction to the ideas behind the formal semantics, focussing on the distinction between action types and action tokens. Then the syntax and semantics of Action Type Deontic Logic is presented and it is shown to meet the benchmarks. Finally, possibilities for further research are indicated. In the appendix, decidability of the satisfiability of formulas is proved via a technique known from monadic First Order Logic.

**Keywords** deontic logic, Ross' paradox, free choice inferences, decidability of satisfiability.

## 1 Introduction

Deontic logics aim at representing correct deontic reasoning, the inferences people ought to make about norms, i.e. what they may, should or must do. Many deontic logics apply deontic operators such as *must* and *may* to propositions, as in the sentence, 'It must be the case that Frank runs', see e.g. Chellas (1980), Hilpinen (1971), Hilpinen (1981), McNamara (2006). However, an old tradition in modern deontic logic going back to von Wright's first paper on the topic, applies deontic operators to action types, as in the simpler sentence, 'Frank must run', see von Wright (1951). Although deviating substantially from von Wright's original proposal, the basic idea of applying deontic operators to action types is maintained here. A set of *acceptable* or *good* action tokens instantiate various action types. An action of a certain type may be done iff there is an action token of that type. An action of a certain type must be

---

Martin Mose Bentzen  
Department of Management Engineering, Technical University of Denmark, Denmark  
E-mail: mmbe@dtu.dk

done iff all action tokens are of that type. Negations, conjunctions, and disjunctions of action types are defined inductively. The disjuncts of disjunctive action types are non-empty in the sense that there are action tokens instantiating each disjunct in the given situation. The aim of this paper is to show that a rather well-behaved deontic logic ensues from these semantic devices. In order to operationalize this claim a number of benchmark cases for a deontic logic will be presented, representing deontic inferences speakers usually make in natural language contexts. I evaluate some well-known deontic logics with regard to these benchmarks. After this, an ontology of action types and action tokens is introduced informally. I then present a new formal language and semantics, *Action Type Deontic Logic*, which validates inferences that coincide with the benchmark cases. Some limitations of the logic and possibilities for further research are discussed. Decidability of satisfiability of formulas of the logic is shown via a variant of well-known technique from monadic First Order Logic.

## 2 Informal deontic reasoning

The cases to be presented below represent deontic inferences most fluent speakers would probably accept in natural language contexts.<sup>1</sup> The benchmark cases are divided into three categories, *consequences* (what one should be able to infer), *non-consequences* (what one should not be able to infer), and *equivalences* (what should mean the same). These are presented in Figures 1, 2, 3. To the best of my knowledge, there is no deontic logic which meets all these benchmarks, that is, validates or proves all of the consequences, does not validate or prove any of the non-consequences and validates or proves all of the equivalences - with the exception of the logic developed here. Most of the examples are in the figures because they play an important role in natural language. However, most of them are also there to highlight specific problems encountered by one or more existing deontic logics.

Many of the inferences in the figures will be highlighted in this section and the next. We will start with Figure 2: 1, 2, known as *Ross' paradox* and introduced by Alf Ross in Ross (1941). Informally, it is exemplified by the inference from 1. to 2. below.

- Example 1*
1. 'You must send the letter.'
  2. 'You must send the letter or burn the letter.'

In standard deontic logic (and many other deontic logics) the paradox crops up in form of the following validity.

---

<sup>1</sup> This claim is based on my own linguistic intuitions, informed by discussions with the people present in various contexts where I have presented this work. I have, however, not tested the claim statistically, hence the qualification 'would probably accept.' Since my default belief is that natural language should be respected as far as possible, I elevate these cases to normative standards or benchmarks that a good deontic theory should meet. Here I appeal to normal linguistic intuitions as a regulative guideline. My view is not that a logical theory should be founded on empirical observations of natural language phenomena. On the contrary, a good theory gives a plausible theoretical explanation of the inferences it validates. However, a theory should provide a very good theoretical reason, if it diverges from natural language, preferably a reason arising within the theory itself. Philosophical logicians are often too eager to sacrifice natural language to save the logical theory they have become accustomed to working with.

	From	Infer
1	You must run or hide	You may run and you may hide
2	You must run	You may run
3	You may run or hide	You may run and you may hide
4	You may run and you may hide	You may run or hide
5	You may run and hide	You may run and you may hide
6	You must run and hide	You must run and you must hide
7	You must run and you must hide	You must run and hide
8		You may run or you may not run
9		It is not the case that you must run and you must not run

Fig. 1 Consequences

	From	Do not Infer
1	You must run	You must run or eat
2	You may run	You may run or eat
3	You must run or hide	You must run or you must hide
4	You may run and you may hide	You may run and hide
5	You may eat or not eat	You may run
5	You may eat	You may eat and hit the waiter

Fig. 2 Non-consequences

		If and only if
1	You must run	It is not the case that you may not run
2	You may run	It is not the case that you must not run

Fig. 3 Equivalences

$$\bigcirc\phi \rightarrow \bigcirc(\phi \vee \psi)$$

In the above sentence,  $\bigcirc$  is read as 'obligatory'. The paradox has a dual twin, where  $\mathbf{P}$  is read as 'permissible'.

$$\mathbf{P}\phi \rightarrow \mathbf{P}(\phi \vee \psi)$$

Ross' paradox brings out differences between deontic and epistemic or alethic modalities. For instance, the following justification of Ross' paradox does not work: since a disjunction gives less information than either disjunct it is pragmatically inappropriate to give a disjunctive permission, when you can give a categorial one. The following example emphasizes this point.

*Example 2* An employer tells an employee that she may work through July. In August he reveals that it was actually the case that she was allowed to work through July or take a vacation. He explains that the reason he did not reveal this, is that he wanted to give her as precise information as possible.

Clearly, the employer's justification is not acceptable. The employer has hidden important information about his employee's rights from her. In my view, this shows that he has not understood the meaning of the word 'may'. Let us continue the example and assume that she in fact did work through July and that the employer strongly believed that this would be the case when he uttered the first permission. This belief was based on his experience with this employee. He thus had a true justified belief that the employee would work through July. Nonetheless, his justification is not acceptable. It is a thesis of this paper that the reason is there are semantic peculiarities of deontic sentences which in turn influence the basic logical inferences that can be made from sentences containing these words, a prime example of which are *free choice inferences*. The following is an example of a free choice inference, see also Figure 1: 1, 3.

*Example 3* 1. 'You may send the letter *or* burn the letter.'  
2. 'You may send the letter *and* you may burn the letter.'

An inference of deontic logic, which is usually considered to be unproblematic is the following, which will be referred to as '*must implies may*', see Figure 1: 1.

*Example 4* From 'you must send the letter' infer 'you may send the letter.'

Or, in general, for any formula  $\phi$ ,

$$\bigcirc\phi \rightarrow \mathbf{P}\phi$$

It is now perhaps clear how we may get problematical reasoning by combining Ross' paradox with the free choice inference, as sketched in the following example.

*Example 5* 1. 'You must send the letter.' (Given.)  
2. 'You must send the letter *or* burn the letter.' (From 1. by Ross' Paradox.)  
3. 'You may send the letter *or* burn the letter.' (From 2. by *must implies may*.)  
4. 'You may send the letter *and* you may burn the letter.' (From 3. by free choice inference.)  
5. 'You may burn the letter.' (From 4. by classical logic.)

This is clearly unacceptable. From any obligation to do something, I get the permission to do anything at all! Now, coming from classical logic, one is inclined to explain Ross' paradox away by denying the non-classical free choice inference. However, as Hans Kamp pointed out already in 1973, we have a strong intuition that a disjunctive

permission *entails* the permission to do either disjunct, see Kamp (1973). Reconsidering the inference above with natural language as a guide, the problematic step does not seem to be from 3. to 4., but rather from 1. to 2.. We should therefore be looking for a logic which blocks Ross' paradox and validates free choice inferences.

### 3 Formal deontic logics

The reason why Ross' paradox crops up in standard deontic logic, indeed in any normal modal logic such as e.g. Horty's stit logic, see Horty (2001), is that a normal modal operator preserves logical consequences, and a disjunction follows from a categorical statement. Of course, the free choice inference is not valid in normal modal logic, but this should be seen as a deficit. Thus any normal modal logic is disqualified from meeting the benchmarks, and we will now look at some alternatives.

A tradition in deontic logic, which does take free choice permissions seriously goes back to Anderson's *logic of prudence* and is now known as a logic of *strong permission*, see e.g. Asher and Bonevac (2005). Anderson employs the notion of a *sanction*, a constant which when true at a world indicates that that world is not morally acceptable. The idea behind the logic of prudence is to "reverse" the modal (box) truth condition, so that  $\mathbf{P}\phi$  is true, if and only if, all worlds which make  $\phi$  true are acceptable. Anderson writes:

States of affairs which entail the denial of the sanction might, for example, be regarded as prudent...

(Anderson; 1966, p. 199)

Anderson's proposal validates free choice inferences and blocks Ross' paradox. However, there are obvious problems. The proposal validates the inferences  $\mathbf{P}\phi \rightarrow \mathbf{P}(\phi \wedge \psi)$  and  $\mathbf{P}(\phi \vee \neg\phi) \rightarrow \mathbf{P}\psi$ , both counter-intuitive in a deontic context, see Figure 2: 5, 6. Moreover, Anderson's strong permission fails to validate the inference  $\mathbf{P}(\phi \wedge \psi) \rightarrow \mathbf{P}\phi$  called *conjunction exploitation*, see Figure 1: 5. In a way, this is quite natural, when we think about Anderson's original idea behind the semantics. From 'it is prudent to jump from a plane and wear a parachute' it does not follow simply that 'it is prudent to jump from a plane' - if you do not wear a parachute, you might get in trouble. However, taken as a representation of permission, the proposal does not capture what we normally mean by by this. For instance, the following conjunction seems problematic.

*Example 6* 'You may invite Beth and invite Smilla, but you may not invite Beth or you may not invite Smilla.' (?)

How can it be that you may invite Beth and Smilla, but there is one of them you may not invite? In other words, it seems that deontic logic should validate conjunction exploitation. The converse, on the other hand should not hold. It does not follow, from the fact that you may take an apple and you may take a pear, that you may take an apple and take a pear, see Figure 2: 4.

A recent update of Anderson's proposal fails to validate conjunction exploitation, see Asher and Bonevac (2005). Moreover, on any such account permissions and obligations fail to be duals - thus we lose the quite natural equivalences in Figure 3: 1,2.

A lot of recent work has been given the heading *dynamic deontic logic*. This work starts with Meyer (1988), and continues with e.g. van der Meyden (1996), Segerberg (2012), see also Meyer and Veltman (2007). One source of inspiration for this work is the following validity of propositional dynamic logic, see e.g. (Goldblatt; 1992, p. 111).

$$[\alpha \cup \beta]\phi \leftrightarrow [\alpha]\phi \wedge [\beta]\phi$$

This is valid because the relation for the modality  $[\alpha \cup \beta]$  is the union of the two relations for  $[\alpha]$  and  $[\beta]$ . Since  $\alpha \cup \beta$  is read ‘do either  $\alpha$  or  $\beta$ ’, it looks a lot like a free choice inference. The sentence may be read ‘after you do either  $\alpha$  or  $\beta$ ,  $\phi$  is true, if and only if, after you do  $\alpha$ ,  $\phi$  is true and after you do  $\beta$ ,  $\phi$  is true’. Even though the idea behind these logics is that modal operators should be applied to action types, Ross’ paradox is not blocked and free choice inferences are not allowed in the first paper mentioned, Meyer (1988). I will not treat it any further here. In van der Meyden (1996), however, Ross’ paradox is blocked and free choice inferences are allowed. Broadly speaking, van der Meyden’s logic is in the tradition of Anderson’s strong permission. An execution of an action is represented as a sequence of states. An action kind or type is a set of such sequences. Some of these sequences or executions are *green* or permitted. A benefit of this framework is of course that one may speak of sequential actions, a topic not treated in this paper. However, in other ways the expressivity of the logic is limited. It is not clear how to talk about doing actions of different kinds simultaneously. This problem is related to the objection to the truth condition of the free choice permission operator stated below. A free choice permission of an action  $a$  holds, when all executions of  $a$  are permitted. Obligation is treated as a dual to a (traditional) weak permission, giving us some of the problems of standard deontic logic for these operators, including Ross’ paradox. Thus, we have a further problem of having two different kinds of permissions, without a well-founded conceptual distinction between the two. A simple objection is that empty actions (with no possible executions) are always permitted, but as van der Meyden points out this can be remedied by a non-emptiness condition. However, there is a more serious objection, which is that the universal style truth condition is too strong to model permission. The problem is not so easy to detect in van der Meyden’s logic because it is not clear how to represent simultaneous actions, or as we might say, action tokens (sequences) instantiating several action types. Intuitively, though, an execution  $s$  is of two types  $\alpha$  and  $\beta$  if it is an element in both sets representing the types. However, if that is a correct way of seeing it, the following problem follows.<sup>2</sup>

*Example 7* There are many different ways an agent may eat a pear. With some of these action executions the agent also kills a person, since the agent eats a pear and kills the person at the same time. However, (we may assume) it is not permitted for the agent to kill a person. Hence this particular action execution cannot be permitted. According to van der Meyden’s semantics, it is not permitted for the agent to eat a pear!

<sup>2</sup> This objection and the following example was presented to me by Frank Veltman.

This seems very counterintuitive, but it is a natural consequence of the requirement that *every single execution of an action* must be green, for the action to be permitted. In view of the discussion about action tokens and action types in the following section, it seems ontologically correct to say that a single action token may instantiate several action types. Further, it should be enough that *some* ways of performing an action type is enough to say that this action type is permitted. However, this *existential* nature of permission is blocked in van der Meyden's semantics. As van der Meyden points out, Meyer's logics do not have this problem, as he has an existential style truth condition for permission. However, in view of the above, it also seems correct to keep van der Meyden's intuition that free choice permissions should be allowed and Ross' paradox blocked.

A related body of work is in the tradition of dynamic epistemic logic, see van Benthem and Liu (2007). Here the focus is on changing the preference relation between worlds dynamically with normative utterances, see also Yamada (2006). This is again related to work on dynamics in logical linguistics, see Veltman (1996). The focus is on the context changing potential of deontic sentences. A classic paper from this perspective is van der Torre and Tan (1999). A recent one is Mastop (2011). However, since the focus in these papers is on performative aspects of deontic sentences, I will not treat them further.

Finally, I mention two proposals more closely related to this paper. The first was developed in Jackson (1985). In that paper, an action is always a member of a set of actions, the available alternative actions. The following truth condition of 'it ought to be that  $A$ ' is given, where  $A$  is an action.

**Definition 1** It ought to be that  $A$  out of  $\{A, A_1, \dots, A_n\}$ , iff what would be the case were  $A$  true is better than what would be the case were  $A_i$  true, for all  $1 \leq i \leq n$ .

Jackson's main intuition that there are a set of available action tokens some of which are permitted is also the main intuition behind the theory presented here. The main difference between Jackson's approach and the one taken here, is perhaps the explicit distinction between action types and action tokens introduced below. On Jackson's proposal, deontic operators are applied to action tokens, whereas they will be applied to action types in the logic to follow.

The second proposal I mention is Zimmermann (2000). In the dynamic deontic logics discussed above, as in dynamic logic in general, there are two symbols for disjunction, one used between sentences and one used between actions. Thus natural language disjunction receives two different formal representations. This is also the approach taken in this paper. However, Zimmermann suggests that disjunction *always* should be interpreted to allow the free choice inference. A disjunction is taken to be a list of epistemic possibilities. An agent may do  $A_1$  or...or  $A_n$ , iff the agent may do all the things on the list  $A_1, \dots, A_n$ .

However, Zimmermann acknowledges that the theory has a problem with deontic *must* as in the following sentence.

*Example 8* 'Mr. X must take a taxi or a bus.'

The only reading his theory can get of the above implies that either Mr. X must take a taxi or Mr. X must take a bus. But the sentence means (interpreted deontically)



that there are two options, each permitted, and that they taken together exhaust all acceptable possibilities. This should imply that Mr. X may do either, but *not* that he must do one of them, see Figure 2: 3.

A solution to this problem is offered in Geurts (2005). However, on both Zimmermann's account and Geurt's revision of it there is no clear distinction between factual contexts, where e.g. disjunction introduction is allowed, and contexts where it not allowed. If we give up classical disjunction in all contexts, we lose inferences that are both common and natural. That this is an issue is recognized by Geurts.

To conclude, none of these approaches get all the inferences in the figures right. Although there are many other deontic logics out there, I will now examine the distinction between action tokens and action types, which is crucial for the semantics to follow.

#### 4 Action types and action tokens

According to von Wright norms apply to *action types* or as he calls them, *acts*, which he takes to cover *act qualifying properties*.

We shall say that theft, murder smoking, etc. are acts. The individual cases that fall under theft, murder, smoking, etc. we shall call act-individuals. It is of acts and not of act-individuals that deontic words are predicated.  
(von Wright; 1951, p. 2)

Peter Geach has criticized deontic logic from moving away from actions and towards applying deontic operators to formulas expressing arbitrary propositions, see Geach (1982). Stit theory is a rather successful attempt at taking Geach's objections seriously, see also (Horty; 2001, p. 4). However, stit theory has been criticized for not representing actions directly.

No author in the Anselm-Kanger-Chellas line up through Belnap - Davidson belongs to a different tradition - has countenanced the existence of actions in logic: action talk, yes; ontology of actions, no.  
(Lindström and Segerberg; 2007, p. 1199)

The authors explicitly say that this remark does not apply to Horty. However, Horty states that there is no way of representing action types in stit theory. Thus stit theory appears to be only suited for treating *action tokens* and their consequences, cf. Horty:

These actions are only action tokens, however - individual concrete actions. There is no such thing as the action type of "opening a window", for example. There are individual, concrete openings of individual windows, but nothing to group them together.  
Horty (2001)

In the philosophy of action the view of treating actions as objects is especially associated with Donald Davidson, see Davidson (1963). As such it is sometimes considered opposed to the modal view on action and agency advocated by e.g. stit theory. Indeed Belnap and his colleagues write,

We delight in the fact that stit does not treat actions primarily as “things” to be counted or named.

(Belnap et al.; 2001, p. vi)

We have one group of philosophers claiming that actions are not to be treated as objects and another group saying they should. In the next section I present the ideas behind the formal logic to follow, where an attempt is made at striking a compromise between the simple naming of action tokens in the tradition of Davidson and the refusal to name actions at all and only talk directly about their consequences as in stit theory.

## 5 The ideas behind Action Type Deontic Logic

I agree with Belnap and colleagues and with von Wright that we do not often talk directly about action *tokens*, especially when it comes to future-directed sentences such as the ones involving obligations and permissions. We do not give names to particular instances of brushing teeth before the act, in fact this seems practically impossible given the many possible ways someone might brush his or her teeth. On the other hand, we talk about action *types* all the time, see e.g. Ross (1930), Hage and Brouwer (2000). It is a thesis of this paper, that by doing so, we indirectly talk about action tokens as well, we implicitly quantify over a set of action tokens. The latter was the part missing in von Wright’s original proposal. Consider the following informal argument. At all times you wish to perform one of the (legally or morally) acceptable actions available to you. If one of these actions instantiates the type of, say, eating an apple, then you may eat an apple. On the other hand, if you may eat an apple, then there has to be an acceptable action of that type. If all of the acceptable actions are of a certain type, e.g. of not-killing, then you must not-kill. On the other hand, if you must not-kill, then it cannot be that there is an acceptable action of the killing type.

In the formal semantics action tokens are seen as primitive objects, as is a set of *basic action types*. Complex action types are then defined inductively. An action token instantiates the action type *not T*, if and only if it does not instantiate *T*. Thus your action is of the type of ‘not going to the beach’, if the action is not of the going to the beach type. An action token may instantiate several types, e.g. the same action token may instantiate ‘going to the beach’, ‘going for a run’, ‘getting some exercise’, ‘meeting a friend’, and so on. Action type conjunctions are thus defined via the various action types the action token instantiates. An action token is of the type *T and R* if and only if, it is of both types. Disjunctions of action types, as conceived here, has the feature of being non-empty. An action token is of the disjunctive action type *S or R*, if and only if, first, it is of either of the types, and second, there are action tokens instantiating both types in the situation. Non-emptiness may here be interpreted as a criterion of relevance or availability. In order for the disjunctive type to be instantiated, there must be tokens available instantiating each disjunct. Now deontic operators **may** and **must** operating on action type terms to form statements can be interpreted as follows. **must**[T] holds if and only if, *every* good action is of type T. **may**[T] holds if and only if *some* good action is of type T.

Let me finish this section by listing a number of idealizing assumptions which will be made in the following. The logic to follow is restricted to one agent in one situation. This means that multi-agent issues are sidestepped, and that sequences of actions are not treated, i.e. the question of doing various things in a certain order cannot be expressed. Further, the set of acceptable actions is *given* in this situation. Depending on the application, this set might represent what is legally, socially, or morally acceptable. However, this implies ignoring the question as to what makes an action “acceptable”, i.e. whether there are universal moral truths, or moral norms depend solely on the values of the community in which they are taken to hold. The logic is propositional in the sense that First Order issues are not treated, e.g. the distinction between *de re* and *de dicto* deontic statements or the question as to whether deontic logic is extensional or intensional. Defeasibility of norms are not treated, and normative dilemmas are not possible in the logic. The possibility of not fulfilling an obligation is also not treated, as it is presupposed that the agent chooses an acceptable action. I will return to some of these idealizing assumptions in the final section. The task now will be to show that even with these restrictions, an interesting deontic logic can be developed.

## 6 Logic

The Action Type Deontic Logic which will be presented is intended to represent the deontic reasoning of a single agent in a specific situation. The idealizing assumptions listed in the previous section will be taken to hold, and knowledge of propositional logic and basic set theory will be assumed.

### 6.1 Syntax

#### 6.1.1 Symbols

The logic is generated from the following symbols.

- Definition 2**
1. A countable set,  $BT$ , of *basic action type symbols*,  $BT = (T_1, T_2, \dots)$ .
  2. The *negation, conjunction and disjunction action type operators*,  $\neg, \cap, \cup$ .
  3. Deontic operators, **must** and **may**.
  4. A countable set,  $P$ , of propositional variables,  $P = (p_1, p_2, \dots)$ .
  5. Propositional operators,  $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ .
  6. Round and square parentheses, left and right,  $(, ), [, ]$ .

#### 6.1.2 Action type terms

An expression of the language is just a string of symbols. We are mainly interested in well-formed action type terms and well-formed formulas. The set,  $ACT$ , of well-formed action type terms is defined inductively as the smallest set satisfying the following conditions:

**Definition 3** 1. For any basic action type term  $T_i \in BT$ ,  $T_i \in ACT$ .  
 2. If  $T, S \in ACT$ , then  $\neg T, (T \cap S), (T \uplus S) \in ACT$ .

### 6.1.3 Formulas

The set of well-formed formulas,  $WFF$ , is defined as the smallest set satisfying the following conditions.

**Definition 4** 1. If  $T \in ACT$ , then  $\mathbf{may}[T] \in WFF$  and  $\mathbf{must}[T] \in WFF$ .  
 2. If  $p_i \in P$ , then  $p_i \in WFF$ .  
 3. If  $\phi, \psi \in WFF$ , then  $\neg\phi, (\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi), (\phi \leftrightarrow \psi) \in WFF$ .

## 6.2 Semantics

### 6.2.1 Models

A *model* for action type deontic logic is a structure  $M = \langle G, V, \theta \rangle$ , where  $G$  (also written  $dom(M)$ ) is a non-empty set called the *domain* of  $M$ . The domain  $G$  represents the set of *acceptable* or *good* action tokens available to an agent. I will use lower case Greek letters,  $\alpha, \beta, \dots$  for members of the domain.<sup>3</sup>  $\theta$  is an interpretation of the propositional variables assigning a truth-value  $t$  or  $f$  to each propositional variable.  $V$  is a valuation function assigning to each basic action type term,  $T_i \in BT$ , a subset of the domain, i.e.  $V : BT \rightarrow \mathcal{P}(G)$ .  $V(T)$  is the set of action tokens of that particular type available to the agent in the situation. In order to interpret the disjunctive action types of the form  $(T \uplus S)$  we use the union of non-empty sets, an operation on sets defined as follows, where  $K$  and  $L$  are sets.

**Definition 5** 1.  $T \uplus S = T \cup S$ , when  $S \neq \emptyset$  and  $T \neq \emptyset$ ,  
 2.  $T \uplus S = \emptyset$ , when  $S = \emptyset$  or  $T = \emptyset$ .

The union of non-empty sets is just ordinary union, when both sets are non-empty. Otherwise, it is defined as the empty set.

$V$  is extended to the entire set of well-formed action type terms as follows, where  $T, S \in ACT$ . In the definiens,  $\setminus$  is set-theoretical complement,  $\cap$  is set-theoretical intersection, and  $\uplus$  is the union of non-empty sets. In the definiendum they are symbols of the formal logic.

**Definition 6** 1.  $V(\neg S) = G \setminus V(S)$ .  
 2.  $V((T \cap S)) = V(T) \cap V(S)$ .  
 3.  $V((T \uplus S)) = V(T) \uplus V(S)$ .

<sup>3</sup> Presumably, the agent also has available a set of actions which are not acceptable, but since we are only concerned with idealized deontic reasoning here, we have no need to model these non-ideal actions. It is an interesting methodological point that we can do without non-ideal actions when modeling deontic reasoning, given the enormous effort which has been put into modeling these in the history of deontic logic. On the other hand, the logic must be extended if we are to handle violations of obligations and contrary-to-duty obligations.

### 6.2.2 Evaluation of formulas

The truth of a formula in a model  $M$  is defined as follows, where  $T \in ACT$  is any action type term,  $p_i \in P$  is any propositional variable, and  $\phi$  and  $\psi$  are any formulas.

**Definition 7** 1.  $M \models \mathbf{may}[T]$  iff for some  $\alpha \in G$ ,  $\alpha \in V(T)$ .

2.  $M \models p_i$  iff  $\theta(p_i) = t$ .

3.  $M \models \neg\phi$  iff  $M \not\models \phi$ .

4.  $M \models \phi \wedge \psi$  iff  $M \models \phi$  and  $M \models \psi$ .

Propositional operators  $\vee$ ,  $\rightarrow$ ,  $\leftrightarrow$  are defined as usual. We define the operator **must** as  $\neg\mathbf{may}[\neg T]$ . **must** $[T]$  then gets the following derived truth condition.

**Fact 1**

$M \models \mathbf{must}[T]$  iff for any  $\alpha \in G$ ,  $\alpha \in V(T)$ .

*Proof* Left to the reader.

Validity, satisfiability, and logical consequence are defined as follows.

**Definition 8** A sentence is valid if it is true in all models. A sentence is satisfiable if it is true in some model. We also say that a set  $\Gamma$  of sentences is true in a model, denoted  $M \models \Gamma$ , if for every  $\phi \in \Gamma$ ,  $M \models \phi$ . We say that a set of sentences  $\Gamma$  is (jointly) satisfiable, if there is a model  $M$ , such that  $M \models \Gamma$ . A sentence  $\phi$  is a logical consequence of a set of sentences  $\Gamma$ , if for any model  $M$  such that  $M \models \Gamma$ , we have  $M \models \phi$ .

### 6.3 Validities

The following validities correspond to ones presented in Figure 1. In the following,  $T$  and  $S$  are in  $ACT$ . I have omitted outer parentheses, when this does not cause any problems.

**Proposition 1** 1.  $\models \mathbf{must}[T \uplus S] \rightarrow (\mathbf{may}[T] \wedge \mathbf{may}[S])$ .

2.  $\models \mathbf{must}[T] \rightarrow \mathbf{may}[T]$ .

3.  $\models \mathbf{may}[T \uplus S] \rightarrow (\mathbf{may}[T] \wedge \mathbf{may}[S])$ .

4.  $\models (\mathbf{may}[T] \wedge \mathbf{may}[S]) \rightarrow \mathbf{may}[T \uplus S]$ .

5.  $\models \mathbf{may}[T \cap S] \rightarrow (\mathbf{may}[T] \wedge \mathbf{may}[S])$ .

6.  $\models \mathbf{must}[T \cap S] \rightarrow (\mathbf{must}[T] \wedge \mathbf{must}[S])$ .

7.  $\models (\mathbf{must}[T] \wedge \mathbf{must}[S]) \rightarrow \mathbf{must}[T \cap S]$ .

8.  $\models \mathbf{may}[T] \vee \mathbf{may}[\neg T]$ .

9.  $\models \neg(\mathbf{must}[T] \wedge \mathbf{must}[\neg T])$ .

*Proof* Left to the reader.

## 6.4 Non-validities

The following non-validities correspond to the ones presented in Figure 2.

- Proposition 2**
1.  $\not\models \mathbf{must}[T] \rightarrow \mathbf{must}[T \uplus S]$
  2.  $\not\models \mathbf{may}[T] \rightarrow \mathbf{may}[T \uplus S]$ .
  3.  $\not\models \mathbf{must}[T \uplus S] \rightarrow (\mathbf{must}[T] \vee \mathbf{must}[S])$ .
  4.  $\not\models (\mathbf{may}[T] \wedge \mathbf{may}[S]) \rightarrow \mathbf{may}[T \cap S]$ .
  5.  $\not\models \mathbf{may}[T \uplus \neg T] \rightarrow \mathbf{may}[S]$ .

*Proof* The proof consists of a countermodel to the inferences,  $M$ . Let  $\text{dom}(M) = \{a, b\}$ . Let  $V(T_1) = \{a, b\}$ ,  $V(T_2) = \emptyset$ ,  $V(T_3) = \{a\}$ ,  $V(T_4) = \{b\}$ , and  $V(T_i) = V(T_2)$ , for  $i > 4$ . Set  $\theta(p_i) = f$  for any  $p_i \in P$ .

- 1.+ 2. We have  $M \models \mathbf{must}[T_1]$ , and  $M \models \mathbf{may}[T_1]$ , but neither  $M \models \mathbf{must}[T_1 \uplus T_2]$ , or  $M \models \mathbf{may}[T_1 \uplus T_2]$ .
3. We have  $M \models \mathbf{must}[T_3 \uplus T_4]$ , but neither  $M \models \mathbf{must}[T_3]$  or  $M \models \mathbf{must}[T_4]$ .
4. We have  $M \models \mathbf{may}[T_3] \wedge \mathbf{may}[T_4]$ , but not  $M \models \mathbf{may}[T_3 \cap T_4]$ .
5. We have  $M \models \mathbf{may}[T_3 \uplus \neg T_3]$ , but not  $M \models \mathbf{may}[T_2]$ .

## 6.5 Equivalences

Finally, here are logical equivalences corresponding to the ones presented in Figure 3 (plus a few extra ones).

- Proposition 3**
1.  $\models \mathbf{must}[T] \leftrightarrow \neg \mathbf{may}[\neg T]$ .
  2.  $\models \mathbf{may}[T] \leftrightarrow \neg \mathbf{must}[\neg T]$ .
  3.  $\models \mathbf{may}[T \cap S] \leftrightarrow \mathbf{may}[(S \cap T)]$ .
  4.  $\models \mathbf{may}[T \uplus S] \leftrightarrow \mathbf{may}[S \uplus T]$ .
  5.  $\models \mathbf{must}[T \cap S] \leftrightarrow \mathbf{must}[T \cap S]$ .
  6.  $\models \mathbf{must}[T \uplus S] \leftrightarrow \mathbf{must}[S \uplus T]$ .

*Proof* 1. By definition of  $\mathbf{must}$ . 2. *Left to right.* Assume  $M \models \mathbf{may}[T]$  and  $M \not\models \neg \mathbf{must}[\neg T]$ . Then there is  $\alpha \in V(T)$ , but since  $M \not\models \neg \mathbf{must}[\neg T]$  implies  $M \models \mathbf{must}[\neg T]$ ,  $\alpha \in V(\neg T)$ , i.e.  $G \setminus V(T)$ , contradiction.

*Right to left.* Assume  $M \models \neg \mathbf{must}[\neg T]$  and  $M \not\models \mathbf{may}[T]$ . Then  $M \not\models \mathbf{must}[\neg T]$ . Hence it is not the case for every  $\alpha \in G$  that  $\alpha \in G \setminus V(T)$ . Thus, for some  $\alpha$ ,  $\alpha \in V(T)$ . However, since  $M \not\models \mathbf{may}[T]$ , there is no  $\alpha \in V(T)$ , contradiction.

3.  $M \models \mathbf{may}[(T \cap S)]$  iff there is  $\alpha \in G$ , s.t.  $\alpha \in V(T)$  and  $\alpha \in V(S)$  iff there is  $\alpha \in G$ , s.t.  $\alpha \in V(S)$  and  $\alpha \in V(T)$  iff  $M \models \mathbf{may}[(S \cap T)]$ .

The rest of the equivalences are similar.

## 6.6 Decidability

It is perhaps intuitively apparent that Action Type Deontic Logic is not very expressive, and hence satisfiability (and thus validity) of formulas ought to be decidable. In particular, the deontic operators are like quantifiers and action types seem like

monadic predicates. In fact, we can obtain the following theorem by a technique used to establish decidability for monadic First Order Logic.

**Theorem 2** *Satisfiability for formulas of Action Type Deontic Logic is decidable.*

Theorem 2 is proved in the appendix.

## 7 Limitations and further research

The **must** and **may** operators are meant to represent deontic modalities as they occur in natural language. As in von Wright's original system (and in natural language, except for emphasis, as in the sentence, 'you must, must, must come tonight!') iterations of these operators are not meaningful. Sequential action types of the conditional type, e.g. 'first brushing your teeth, then flossing', cannot be expressed within the semantics. An obvious avenue of further research would be to extend the system to handle sequences of actions, looking to dynamic logic and temporal logic for inspiration, see e.g. van der Meyden (1996), Goldblatt (1992). In the single situation context presented here, we can still express conditionals with propositional 'triggers' as in the following example.

- Example 9*
1. 'If it rains, you must use an umbrella.' ( $p \rightarrow \mathbf{must}[T]$ )
  2. 'You must use an umbrella only if it rains.' ( $\mathbf{must}[T] \rightarrow p$ )
  3. 'If you must hand in a paper, you must correct it first.' ( $\mathbf{must}[H] \rightarrow \mathbf{must}[C]$ )
  4. 'If it rains, you must *not* use an umbrella.' ( $p \rightarrow \mathbf{must}[-T]$ )
  5. 'If it does not rain, it is not the case that you must use an umbrella.' ( $\neg p \rightarrow \neg \mathbf{must}[T]$ )

In the third sentence, intonation can be used to emphasize that deontic must is meant in the antecedent 'you must hand in a paper'. In the fourth sentence the scope of the negation is 'use an umbrella', i.e. it is an internal action type negation. In the fifth sentence the scope of the negation in the consequent is 'you must use an umbrella'. In many contexts, this sentence could be paraphrased 'If it does not rain, you need not use an umbrella'. These considerations are similar for **may**.

There are many interesting First Order issues in deontic logic, which cannot be treated within the semantics thus far. It would be nice to be able to express for instance the following.

*Example 10*

The Prime Minister of Denmark must be appointed by the Monarch.

However, this would require research into First Order Action Type Deontic Logic, taking inspiration from First Order Modal Logic, see e.g. Fitting and Mendelsohn (1998), Bräuner and Ghilardi (2007), Goble (1996). Another obvious limitation concerns contrary-to-duty obligations, see e.g. Prior (1954), Chisholm (1964), Forrester (1984), Prakken and Sergot (1997), and defeasible obligations. With regard to the first, the semantics could be extended with a set of non-acceptable actions and the valuation of action types suitably extended. This would enable us to say that an agent has violated a norm. Another extension would be to *defeasible* norms, e.g. taking

W.D. Ross' famous theory on *prima facie* vs. *all things considered* duties as a starting point, as well the work done in this area, see Ross (1930), Nute (1997), Bringsjord and Taylor (2012). One possibility is putting a preference ordering on the action tokens, in order to represent a more fine-grained theory about acceptable actions, see e.g. Horty (2001).

The logic presented does not cover the inference from 'You may run or you may hide' to 'you may hide'.<sup>4</sup> The technical reason why it is not validated is that in this sentence the deontic disjuncts are within the scope of a sentential disjunction which is classical in the current system. However, I have no doubt that many people will in fact accept this inference, which could be seen as indicating a certain spillage from the deontic context into the surrounding sentential context. Dealing with this phenomenon is beyond the scope of the present system. A solution would require that we give up classical disjunction in further contexts than the ones within the scope of deontic operators. However, I maintain that there are also purely factual contexts where disjunction is classical and where disjunction introduction should be permitted. The system presented here presents one compromise between these two facts, but there are undoubtedly other ways to cut the cake and further research is required.

Finally, I mention the need to extend the logic to handle multi-agent situations. This might enable us to shed new light on the distinction between individual agency and group agency, see e.g. Belnap et al. (2001), Horty (2001).

Despite these limitations, Action Type Deontic Logic solves a good deal of the problems found in the literature on deontic logic in a natural way. It is intended to serve as a foundation for further research into extensions and applications for instance in semantics, see e.g. Portner (2009), formal decision theory, see e.g. [reference to author's own work removed], machine law, see e.g. Prakken (1993), and robot ethics, see e.g. Arkin (2009), Abney (2012), Bringsjord and Taylor (2012).

## 8 Appendix

I now state the definitions and lemmas needed to prove theorem 2 stating that Action Type Deontic Logic.

**Definition 9** Let  $M = \langle G, V, \theta \rangle$  and  $M' = \langle G', V', \theta' \rangle$  be models. Define a *strong homomorphism* from  $M$  to  $M'$  to be a function  $h$  from  $dom(M)$  to  $dom(M')$ , such that.

1. For any  $p_i \in P$ ,  $\theta(p_i) = \theta'(p_i)$ .
2. For any  $T_i \in BT$ ,  $\alpha \in G$  ( $\alpha \in V(T_i)$  iff  $h(\alpha) \in V'(T_i)$ ).

For the following lemma to go through we need a *surjective* strong homomorphism. Let  $M$  and  $M'$  be models and  $h$  a surjective, strong homomorphism from  $M$  to  $M'$ .

**Lemma 1** 1. For any  $T \in ACT$ ,  $\alpha \in G$  ( $\alpha \in V(T)$  iff  $h(\alpha) \in V'(T)$ ).  
2. For any  $\phi \in WFF$ ,  $M \models \phi$  iff  $M' \models \phi$

Part 2 of the above lemma shows that satisfiability is invariant under surjective, strong homomorphisms.

<sup>4</sup> I would like to thank an anonymous reviewer for pointing this out.



*Proof* 1. We define the *complexity* of an action term as the number of action type terms in the action type term. The proof is by induction on the complexity of action terms.

1. Induction basis. Let  $T_i \in BT$ ,  $\alpha \in G$ . We have  $\alpha \in V(T_i)$  iff  $h(\alpha) \in V'(T_i)$  by definition of surjective, strong homomorphism.
2. Suppose the lemma holds for any action type term of complexity lower than  $n$ . Now suppose  $T$  is an action type term with complexity  $n$ . The proof divides into several cases.
  - (a)  $T$  is of the form  $-S$ . *Left to right*. If  $\alpha \in V(-S)$ , then  $\alpha \in G \setminus V(S)$ , and so  $\alpha \notin V(S)$ . Then  $h(\alpha) \notin V'(S)$  (by induction hypothesis), so  $h(\alpha) \in V'(-S)$ . *Right to left*. If  $h(\alpha) \in V'(-S)$ , then  $h(\alpha) \in G' \setminus V'(S)$ , and so  $h(\alpha) \notin V'(S)$ . Then  $\alpha \notin V(S)$  (by induction hypothesis), so  $\alpha \in V(-S)$ .
  - (b)  $T$  is of the form  $(S \cap R)$ . *Left to right*. If  $\alpha \in V(S \cap R)$ , then  $\alpha \in V(S)$  and  $\alpha \in V(R)$ . Hence, by induction hypothesis,  $h(\alpha) \in V'(S)$  and  $h(\alpha) \in V'(R)$ . Hence  $h(\alpha) \in V'(S \cap R)$ . *Right to left*. If  $h(\alpha) \in V'(S \cap R)$ , then  $h(\alpha) \in V'(S)$  and  $h(\alpha) \in V'(R)$ . Hence, by induction hypothesis,  $\alpha \in V(S)$  and  $\alpha \in V(R)$ . Hence  $\alpha \in V(S \cap R)$ .
  - (c)  $T$  is of the form  $(S \uplus R)$ . *Left to right* If  $\alpha \in V(S \uplus R)$ , then 1)  $\alpha \in V(S)$  or 2)  $\alpha \in V(R)$ . Assume 1. By induction hypothesis  $h(\alpha) \in V'(S)$ . Since  $V(R) \neq \emptyset$  by definition of  $\uplus$ , there is  $\beta \in V(R)$ . By induction hypothesis  $h(\beta) \in V'(R)$ . Hence  $h(\alpha) \in V'(S \uplus R)$ . The other case is similar. *Right to left*. If  $h(\alpha) \in V(S \uplus R)$ , then 1)  $h(\alpha) \in V'(S)$  or 2)  $h(\alpha) \in V'(R)$ . Assume 1. By induction hypothesis  $\alpha \in V(S)$ .  $V'(R) \neq \emptyset$  by definition of  $\uplus$ . Let  $\delta \in V'(R)$ . Since  $h$  is surjective, there is  $\beta \in G$  such that  $h(\beta) = \delta$ . By induction hypothesis  $\beta \in V(R)$ . Hence,  $\alpha \in V(S \uplus R)$ . The other case is similar.

2. We define the *complexity* of a formula as the number of propositional connectives in the formula. The proof is by induction on the complexity of formulas.

1. There are two basic cases, according to whether  $\phi$  is a propositional variable  $p_i \in P$  or  $\phi$  is of the form  $\mathbf{may}[T]$  for some  $T \in ACT$ .  $M \models p_i$  iff  $M' \models p_i$  is obvious from definition of surjective, strong homomorphism. Assume  $M \models \mathbf{may}[T]$ . Then there is  $\alpha \in V(T)$ . By part 1 of this lemma,  $h(\alpha) \in V'(T)$ . Hence  $M' \models \mathbf{may}[T]$ . Assume  $M' \models \mathbf{may}[T]$ . Then there is  $\beta \in V'(T)$ . Since  $h$  is surjective, there is  $\delta \in G$ , such that  $h(\delta) = \beta$ . Hence by part 1 of this lemma,  $\delta \in V(T)$ . Therefore,  $M \models \mathbf{may}[T]$ .
2. Suppose the lemma holds for all formulas of complexity lower than  $n$ . Now, let  $\phi$  be a formula of complexity  $n$ . We consider two cases.
  - (a) Suppose  $\phi$  is of the form  $\neg\psi$ . *Left to right*. If  $M \models \neg\psi$ , then  $M \not\models \psi$  and hence  $M' \not\models \psi$  (by induction hypothesis) and so  $M' \models \neg\psi$ , i.e.  $M' \models \phi$ . *Right to left*. If  $M' \models \neg\psi$ , then  $M' \not\models \psi$  and hence  $M \not\models \psi$  (by induction hypothesis) and so  $M \models \neg\psi$ .
  - (b) Suppose  $\phi$  is of the form  $\psi \wedge \chi$ . *Left to right*. If  $M \models \psi \wedge \chi$ , then  $M \models \psi$  and  $M \models \chi$ . Hence  $M' \models \psi$  and  $M' \models \chi$  (by induction hypothesis). So,  $M' \models \psi \wedge \chi$  i.e.  $M' \models \phi$ . *Right to left*. If  $M' \models \psi \wedge \chi$ , then  $M' \models \psi$  and  $M' \models \chi$ . Hence  $M \models \psi$  and  $M \models \chi$  (by induction hypothesis). So,  $M \models \psi \wedge \chi$ .

Now we show decidability of satisfiability with a technique used for showing this for monadic First Order Logic, see e.g. Jeffrey et al. (1997). We prove the following lemma.

**Lemma 2** *If a formula  $\phi$  is satisfiable, then it is satisfiable in a model with a domain of size at most  $2^n$ , where  $n$  is the number of basic action type terms occurring in  $\phi$ .*

*Proof* Let  $\phi$  be a formula and  $M = \langle G, V, \theta \rangle$  a model such that  $M \models \phi$ . Let  $S = \{s_1, \dots, s_n\}$  be the basic action types occurring in  $\phi$ , and  $Q = \{q_1, \dots, q_m\}$  be the propositional variables occurring in  $\phi$ . We first derive the model  $M' = \langle G', V', \theta' \rangle$  from  $M$  as follows.

1.  $dom(M') = dom(M)$ .
2. For any  $T_i \in BT$ , if  $T_i \in S$ , then  $V'(T_i) = V(T_i)$ . Otherwise  $V'(T_i) = \emptyset$ .
3. For any  $p_i \in P$ , if  $p_i \in Q$ , then  $\theta'(p_i) = T(p_i)$ . Otherwise  $\theta'(p_i) = f$ .

In a sense,  $M'$  is  $M$  stripped of all information not pertaining to  $\phi$ . Since the truth of  $\phi$  only depends on its subformulas, an easy inductive argument shows that  $M' \models \phi$ .

Let the *signature* of an  $\alpha \in dom(M')$  be a sequence  $(j_1, \dots, j_n)$  of 0's and 1's, such that, for each  $s_k \in S$

1.  $j_i = 1$ , if  $\alpha \in V(s_k)$ .
2.  $j_i = 0$ , if  $\alpha \notin V(s_k)$ .

There are at most  $2^n$  signatures. Call  $\alpha$  and  $\beta$  *similar*, if they have the same signature. Obviously, similarity is an equivalence relation. For an element  $\alpha \in G'$ , we use the notation  $|\alpha|$  for the equivalence class  $C$  s.t.  $\alpha \in C$ . Now we define  $N = \langle F, V^N, \theta^N \rangle$  as follows. The domain of  $N$  consists of exactly one element from each of the equivalence classes induced by the similarity relation. Hence  $dom(N)$  is of size at most  $2^n$ . Further, we define:

**Definition 10** 1. For each  $T_i \in BT, \beta \in dom(N)$  ( $\beta \in V^N(T_i)$  iff  $\beta \in V'(T_i)$ ).  
2. For each  $p_i \in P, \theta^N(p_i) = \theta'(p_i)$ .

We now define a function  $h$ , from  $dom(M')$  to  $dom(N)$  as follows. For each  $\alpha \in G', h(\alpha) = \delta$ , where  $\delta \in dom(N)$ , and  $|\alpha| = |\delta|$ . Clearly, there is exactly one such  $\delta$ , hence  $h$  is a surjective function. Further, by Definition 10  $h$  is a strong homomorphism. Hence by Lemma 1,  $N \models \phi$ .

Theorem 2 follows directly from Lemma 2.

**Acknowledgements** I would like to thank Frank Veltman, Jens Ulrik Hansen, Klaus Frovin Jørgensen, Carola Eschenbach, and Patrick Blackburn for helpful suggestions regarding this work.

## References

Abney, K. (2012). Robotics, ethical theory, and metaethics: A guide for the perplexed, in P. Lina, K. Abney and G. A. Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics*, MIT Press, pp. 35–52.

- Anderson, A. R. (1966). The formal analysis of normative systems, in N. Rescher (ed.), *The Logic of Decisions and Actions*, University of Pittsburgh Press.
- Arkin, R. (2009). *Governing Lethal Behavior in Autonomous Robots*, CRC Press.
- Asher, N. and Bonevac, D. (2005). Free choice permission is strong permission, *Synthese* **145**: 303–323.
- Belnap, N., Perloff, M. and Xu, M. (2001). *Facing the Future*, Oxford University Press.
- Bräuner, T. and Ghilardi, S. (2007). First-order modal logic, in P. Blackburn, J. van Benthem and F. Wolter (eds), *Handbook of Modal Logic*, Elsevier, pp. 549–620.
- Bringsjord, S. and Taylor, J. (2012). The divine-command approach to robot ethics, in P. Lina, K. Abney and G. A. Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics*, MIT Press, pp. 85–108.
- Chellas, B. F. (1980). *Modal Logic*, Cambridge University Press.
- Chisholm, R. M. (1964). Contrary-to-duty imperatives and deontic logic, *Analysis* **24**: 33–36.
- Davidson, D. (1963). Actions, reasons, and causes, *The Journal of Philosophy* **60**.
- Fitting, M. and Mendelsohn, R. L. (1998). *First-Order Modal Logic*, Kluwer Academic Publishers.
- Forrester, J. (1984). Gentle murder, or the adverbial samaritan, *The Journal of Philosophy* **81**: 193–197.
- Geach, P. (1982). Whatever happened to deontic logic, *Philosophia* **11**: 1–12.
- Geurts, B. (2005). Entertaining alternatives: disjunctions as modals, *Natural Language Semantics* **13**: 383–410.
- Goble, L. (1996). ‘ought’ and extensionality, *Noûs* **30**: 330–355.
- Goldblatt, R. (1992). *Logics of Time and Computation*, Lecture Notes, second edn, Center for the Study of Language and Information.
- Hage, J. and Brouwer, B. (2000). Action types and act tokens in deontic logic of the ought-to-do type, *DEON 00*, pp. 187–201.
- Hilpinen, R. (ed.) (1971). *Deontic Logic: Introductory and Systematic Readings*, D. Reidel Publishing Company.
- Hilpinen, R. (ed.) (1981). *New Studies in Deontic Logic*, D. Reidel Publishing Company.
- Horty, J. F. (2001). *Agency and Deontic Logic*, Oxford University Press.
- Jackson, F. (1985). On the semantics and logic of obligation, *Mind* pp. 177–195.
- Jeffrey, B. and Burgess (1997). *Logic and Computability*, Cambridge University Press.
- Kamp, H. (1973). Free choice permission, *Proceedings of the Aristotelian Society* **74**: 57–74.
- Lindström, S. and Segerberg, K. (2007). Modal logic and philosophy, in P. Blackburn, J. van Benthem and F. Wolter (eds), *Handbook of Modal Logic*, Elsevier, pp. 1150–1214.
- Mastop, R. (2011). Norm performatives and deontic logic, *European Journal of Analytic Philosophy* **7**.
- McNamara, P. (2006). Deontic logic, *Stanford Encyclopedia of Philosophy*.
- Meyer, J.-J. (1988). A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic, *Notre Dame Journal of Formal Logic* **29**: 109–136.

- Meyer, J.-J. and Veltman, F. (2007). Intelligent agents and common-sense reasoning, in P. Blackburn, J. van Benthem and F. Wolter (eds), *Handbook of Modal Logic*, Elsevier, pp. 991–1029.
- Nute, D. (ed.) (1997). *Defeasible Deontic Logic*, Kluwer Academic Publishers.
- Portner, P. (2009). *Modality*, Oxford University Press.
- Prakken, H. (1993). *Logical Tools for Modelling Legal Argument*, Vrije Universiteit, Amsterdam.
- Prakken, H. and Sergot, M. (1997). Dyadic deontic logic and contrary-to-duty obligations, in D. Nute (ed.), *Defeasible Deontic Logic*, Kluwer Academic Publishers, pp. 223–262.
- Prior, A. N. (1954). The paradoxes of derived obligation, *Mind* **63**: 64–65.
- Ross, A. (1941). Imperatives and logic, *Theoria* **7**: 53–71.
- Ross, W. D. (1930). *The Right and the Good*, Clarendon Press.
- Segeberg, K. (2012). D $\Delta$ L: a dynamic deontic logic, *Synthese* **185**: 1–17.
- van Benthem, J. and Liu, F. (2007). Dynamic logic of preference upgrade, *Journal of Applied Non-Classical Logics* **17**: 129–155.
- van der Meyden, R. (1996). The dynamic logic of permission, *Journal of Logic and Computation* **6**: 465–479.
- van der Torre, L. W. and Tan, Y.-H. (1999). An update semantics for deontic reasoning, in P. McNamara and H. Prakken (eds), *Norms, Logics and Information Systems*, IOS Press.
- Veltman, F. (1996). Defaults in update semantics, *Journal of Philosophical Logic* **25**: 221–261.
- von Wright, G. H. (1951). Deontic logic, *Mind* **60**: 1–15.
- Yamada, T. (2006). Acts of commanding and changing obligations, *Proceedings of the Seventh International Workshop on Computational Logic in Multi-Agent Systems (CLIMA VII)*, pp. 1–19.
- Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility, *Natural Language Semantics* **8**: 255–290.