Chapter 10
Deontic Logic

Jan Broersen, Dov Gabbay, Andreas Herzig, Emiliano Lorini, John-Jules Meyer, Xavier Parent, and Leendert van der Torre

10.1 Introduction

The chapter is organized in ten sections, each of them presenting a fundamental issue or problem in the area of deontic logic. Section 10.2 is about the issue of norms and truth, i.e., whether norms should have truth value. It opposes deontic logic viewed as a logic of normative propositions to deontic logic viewed as a logic of imperatives. Section 10.3 deals with the problem of contrary-to-duty (CTD) reasoning while Sect. 10.4 is about the problem of normative conflicts, i.e., how a logic of norms can represent conflicting obligations both syntactically and semantically. Section 10.5 focuses on the issue of norm revision relating it to the problem of belief revision as studied in the classical approach of Alchourron, Gardenfors and Makinson (AGM). Sections 10.6 and 10.7 consider two fundamental problems in the logical representation of norms, namely the logical representation of the temporal aspects of norms and the logical representation of norms about

J. Broersen (✉) • J.-J. Meyer
Department of Information and Computing Sciences, Universiteit Utrecht, Utrecht, The Netherlands
e-mail: J.M.Broersen@uu.nl; jj@cs.uu.nl

D. Gabbay
Department of Computer Science, King's College London, London, UK
e-mail: dov.gabbay@kcl.ac.uk

A. Herzig • E. Lorini
Logic, Interaction, Language, and Computation Group, IRIT, Université Paul Sabatier, Toulouse, France
e-mail: herzig@irit.fr; Emiliano.Lorini@irit.fr

X. Parent • L. van der Torre
Individual and Collective Reasoning (ICR) group, University of Luxembourg, Walferdange, Luxembourg
e-mail: xavier.parent@uni.lu; leon.vandertorre@uni.lu

S. Ossowski (ed.), *Agreement Technologies*, Law, Governance and Technology Series 8, DOI 10.1007/978-94-007-5583-3_10,

actions (as opposed to norms about states of affairs). Section 10.8 touches on the issue of the relationship between norms and games and, in particular, the problem of the relationship between norms and agents' preferences and the problem of how norms are created through agreement. In Sect. 10.9 the problem of the representation of permissive norms is discussed, while Sect. 10.10 deals with the issue of the relationship between norms and mental attitudes such as beliefs, knowledge, preferences and intentions. Finally, Sect. 10.11 discusses the classical distinction between regulative rules and constitutive rules.

10.2 Norm Without Truth

The first problem is to reconstruct deontic logic in accordance with the idea that norms are neither true nor false. There are two approaches.

The mainstream approach is to reconstruct deontic logic as a logic of normative propositions. The idea is that, though norms are neither true nor false, one may state that (according to the norms), something ought to be done: the statement "John ought to leave the room" is, then, a true or false description of a normative situation. Such a statement is usually called a normative proposition, as distinguished from a norm. The Input/Output (I/O) framework of Makinson and van der Torre (2000), and the bi-modal system NOBL due to Åqvist (2008), are two different reconstructions of deontic logic as a logic of normative propositions, thus conceived.

The other approach consists of reconstructing deontic logic as a logic of imperatives. This approach is documented in Hansen (2005, 2008), to which the reader is referred for further details.

10.3 Reasoning About Norm Violation

The question of how to deal with violations and obligations resulting from violations is known as the problem of contrary-to-duty (CTD) reasoning (Chisholm 1963). It is of key importance to the analysis of multi-agent systems. Agents are supposed to be autonomous normative entities. So, they must be able to take into account the existence of social norms in their decisions (either to follow or violate the latter norms). Sanctions are also needed to increase the degree of predictibility of the system (Castelfranchi et al. 2000). Since SDL[1] was criticized for not being able to deal with CTD duties, the issue of CTD has not disappeared from the stage of deontic logic. New standards have been developed in order to make deontic logic suitable for application to the analysis of normative multi-agent systems.

[1]SDL stands for "Standard Deontic Logic". This is a misnomer, because it is no longer considered a standard.

These standards are documented in Hansson (1969), Loewer and Belzer (1983), Prakken and Sergot (1997), van der Torre and Tan (1997), Carmo and Jones (2002), Makinson and van der Torre (2001) and Parent (2003).

10.4 Normative Conflicts

There are two main questions here. The first one is: how can deontic logic accommodate possible conflicts between norms? The first systems of deontic logic precluded the possibility of any such conflict. This makes them unsuitable as a tool for analyzing normative reasoning. Different ways to accommodate normative conflicts have been studied over the last 15 years. A comparative study of them can be found in Goble (2007).

The second question is: how can the resolution of conflicts amongst norms be semantically modeled? An intuitively appealing modeling approach involves using a priority relation defined on norms. There have been several proposals to this effect, and the reader is referred to the discussions in Boella and van der Torre (2003), Hansen (2005, 2008), Horty (2007) and Parent (2010, 2011). An open question is whether tools developed for so-called non-monotonic reasoning are suitable for obligations and permissions.

10.5 Revision of a Set of Norms

Alchourrón and Makinson were the first to study the changes of a legal code (Alchourrón and Makinson 1981, 1982). The question is: how to revise a set of regulations or obligations? Does belief revision (as modelled by the so-called AGM model (Alchourrón et al. 1985)) offer a satisfactory framework for norms revision?

Some of the AGM axioms seem to be rational requirements in a legal context, whereas they have been criticized when imposed on belief change operators. An example is the *success* postulate, requiring that a new input must always be accepted in the belief set. It is reasonable to impose such a requirement when we wish to enforce a new norm or obligation. However, it gives rise to irrational behaviors when imposed to a belief set, as observed for instance in Gabbay et al. (2003).

On the other hand, when we turn to a proper representation of norms, like in the input/output logic framework, the AGM principles prove to be too general to deal with the revision of a normative system. For example, one difference between revising a set of propositions and revising a set of regulations (pointed out in Boella et al. 2009) is the following: when a new norm is added, coherence may be restored modifying some of the existing norms, not necessarily retracting any of them.

Another type of change that has been studied in deontic logic is the aggregation of regulations (Booth et al. 2006; Cholvy and Cuppens 1999; Grégoire 2004).

10.6 Time

Most formalisms do not have temporal operators in the object language, nor do they have, in their standard formulation, an interpretation in temporal models. Yet for several scenarios and decisions involving deontic reasoning, the temporal aspect of the reasoning seems crucial, and several researchers have sought to study logics for the interactions between temporal and deontic modalities. The research question is: what is the relation between deontic conditionals and temporal deontic modalities?

Two natural concepts to be considered are 'validity time' and 'reference time' of an obligation, prohibition or permission. The validity time is the point in time where a deontic modality is true (surpassing the issue of Sect. 10.2 here we simply assume normative modalities have truth values relative to some coherent body of norms that is left implicit) and the reference time is the point in time the obligation, prohibition or permission applies to. For instance, we can have the obligation now (validity time) to show up at the dentist's tomorrow (reference time).

Systems dealing with these temporal differences have been studied, for instance, in Åqvist and Hoepelman (1981) and Thomason (1981). Subtleties in expressing deontic temporal statements involving deontic deadlines have been studied in Broersen et al. (2004) and Broersen (2006).

10.7 Action

We often think of deontic modalities as applying to actions instead of states of affairs. The problems arising in this area are the following: how do we combine deontic modalities with action modalities? How do deontic and action modalities interact. Which action formalisms are best suited for a deontic extension?

Two approaches to deontic action logic prominent in the literature are dynamic deontic logic (Meyer 1988) and deontic *stit* logic (Horty 2001). In dynamic deontic logic normative modalities are reduced to dynamic logic action modalities by using violation constants. Prohibition, for instance, is modeled as the dynamic logic conditional assertion that if the action is executed, a violation will occur. In deontic *stit* logic, the perspective on action is different. Where in dynamic logic actions are objects that are given proper names in the object language, in *stit* logic actions derive their identity from the agent(s) executing them and the effect they achieve. This allows for a proper theory of agency, ability and joint ability. In Horty (2001) normativity is introduced in *stit* theory by means of a deontic ideality ordering. But the alternative of violation constants has also been used in the *stit* context (Bartha 1993; Broersen 2011). A perspective that is symmetric to violation constants is taken in Herzig et al. (2011a) where a dynamic logic is introduced that has special constants encoding an agent's permissions to perform actions (and these deontic abilities are opposed to constants encoding an agent's ontic or non-deontic abilities).

10.8 Norm Emergence and Games

To understand why a norm emerges in an agent society, one has to understand in what sense norms are related to the social preferences and abilities of coalitions of agents. This is the setting of game theory (see also Chap. 12). In deontic logic we distinguish between situations where norms are likely to be in-line with the individual preference, like in coordination problems, and situations where norms, once established, are likely to oppose the preferences of individuals.

Broersen et al. (2008) models the dependency of socially optimal norms on the underlying preferences, in the context of Coalition Logic (Pauly 2002). There it is assumed that the reachability of outcomes that are optimal for the whole group gives rise to a social norm saying that sub-groups should not pursue their own best interest if that conflicts with the group's interest. There are close connections with other work in deontic logic (Kooi and Tamminga 2008) that have to be explored. Open questions include the generalization to the fully strategic case (i.e., from normal game forms to extensive game forms), and the connection with logical models for the dynamics of preferences (Liu 2008).

Norm acceptance can also be considered a game played with other agents subject to the same norm. This idea is explored in Ågotnes et al. (2007). An open question is the complexity of deciding whether or not a normative system is a Nash-equilibrium (or 'Nash-implementation'), relative to a set of other normative systems.

Another interesting issue is how norms are created through agreement. Lorini and Longin (2008) and Lorini et al. (2009) have proposed a logical model of collective acceptance assuming that the existence and the dynamics of a norm depend on its acceptance by the members of an institution (e.g. the existence and dynamics of the rules of chess depend on their acceptance by chess players).

10.9 Permissive Norms

For a long time, it was naively assumed that permission can simply be taken as the dual of obligation, just as possibility is the dual of necessity in modal logic. Something is permitted if its negation is not forbidden. Nowadays in deontic logic a more fine-grained notion of permission is used. The notions of explicit permission, dynamic permission,[2] and permission as exception to a pre-existing obligation are also used. One main finding is that these normative concepts can all be given a well-defined semantics in terms of Input/Output logic (Boella and van der Torre 2008; Makinson and van der Torre 2003; Stolpe 1997, 2010). The main open problem concerns their proof-theory, which is still lacking.

[2] A dynamic permission is forward-looking and is like a constitutional right – it sets limits on what can be forbidden.

10.10 Knowledge and Intentions

For a complete logical picture of rational agency, we need to study the interactions of the deontic modalities with other motivational attitudes like desire and intention, and with epistemic attitudes like belief and knowledge.

The 'BOID' architecture (Broersen et al. 2002, 2005) studies the interplay between beliefs, obligations, intentions and desires in the formation of agent goals. One of the issues discussed in the context of BOID is that the interplay between 'internal' motivations and 'external' motivations (originating from norms of the agent's social context), enables one to distinguish between several agent types. For instance, a benevolent agent will give priority to norms, while an egocentric agent will not. The difference between benevolent agents and egocentric agents shows that the main issue here is 'norm acceptance'. Benevolent agents are more willing to internalize, or accept norms than egocentric ones.

In Broersen (2011) the relation between deontic modalities and epistemic modalities is studied in the context of formalizing different modes of acting. Different modes of acting are relevant in a deontic context, since the deontic status of an act depends, for instance, on whether it is performed knowingly, intentionally, deliberately, etc.

10.11 Constitutive Norms

In legal and social theory one encounters various types of norms. First of all there are the regulative norms describing obligations, prohibitions and permissions. But also there are so-called constitutive norms, which make possible basic 'institutional' actions such as the making of contracts, the issuing of fines, the decreeing of divorces. Basically they tell us what *counts as* what for a given institution. An example is that "cars count as vehicles" in a certain institution having to do with traffic. As pointed out in Boella and van der Torre (2006a), constitutive norms have been identified as the key mechanism to normative reasoning in dynamic and uncertain environments, for example to realize agent communication and electronic contracting.

Although the 'count-as' relation "X counts as Y in context C" was already introduced by Searle (1969), the paper by Jones and Sergot (1996) is often credited for having launched the area of logical investigation of constitutive norms. There, the counts-as relation is viewed as expressing the fact that a given action "is a sufficient condition to guarantee that the institution creates some (usually normative) state of affairs". A conditional connective $\Rightarrow_s$ is used to express the "counts-as" connection holding in the context of an institution s. In his thesis (2007) Grossi disentangles various notions of counts-as, such as classificatory, proper classificatory, and constitutive counts-as. He also treats their formal logical representations

and axiomatisations (in modal logic), as well as their formal relations, and as such clarifies and improves upon the seminal work of Jones and Sergot mentioned above (also cf. Grossi et al. 2006).

When defining constitutive norms, the main issue is in defining their relation with regulative norms. To this end, Boella and van der Torre (2006b) use the notion of a logical architecture combining several logics into a more complex logical system, also called logical input/output nets (or lions). Grossi (2007, p. 104) argues that regulative norms may be viewed as a special case of constitutive norms by employing some kind of Anderson's reduction and putting Obligated(p) as "¬ p counts as V", where V stands for a violation atom. An approach combining a logic of action and deontic ability with a counts-as connective is in Herzig et al. (2011b).

It is expected that deontic logic, as a field of study, will increasingly attract the interest of researchers working in computer science, philosophy, legal theory and even cognitive science. Deontic logic is at the center of many new developments in computer science, motivated by the need to describe distributed interacting autonomous systems at higher levels of abstraction. In philosophy, theories of agency and action can only be seriously evaluated if normative aspects in the form of responsibility, blame and excuse are added to the picture. In legal theory there is a tendency towards formalization and automation and this cannot be achieved without input from deontic logic. Finally, in cognitive science, computational models of the mind might find inspiration in the models used to interpret systems of deontic logic, and vice versa.

References

Ågotnes, T., M. Wooldridge, and W. van der Hoek. 2007. Normative system games. In *Proceedings of the sixth international conference on autonomous agents and multiagent systems (AAMAS 2007), IFAMAAS (2007)*, Honolulu, ed. M. Huhns and O. Shehory, 876–883.

Alchourrón, C. E., and D. C. Makinson. 1981. Hierarchies of regulations and their logic. In *New studies in deontic logic*, ed. R. Hilpinen, 125–148. Dordrecht: D. Reidel.

Alchourrón, C. E., and D. C. Makinson. 1982. The logic of theory change: Contraction functions and their associated revision functions. *Theoria* 48: 14–37.

Alchourrón, C., P. Gärdenfors, and D. Makinson. 1985. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic* 50: 510–530.

Åqvist, L. 2008. Alchourrón and Bulygin on deontic logic and the logic of norm-propositions: Axiomatization and representability results. *Logique et Analyse* 51(203): 225–261.

Åqvist, L., and J. Hoepelman. 1981. Some theorems about a tree system of deontic tense logic. In *New studies in deontic logic*, ed. R. Hilpinen, 187–221. Dordrecht: D. Reidel.

Bartha, P. 1993. Conditional obligation, deontic paradoxes, and the logic of agency. *Annals of Mathematics and Artificial Intelligence* 9(1–2): 1–23.

Boella, G., and L. van der Torre. 2003. Permissions and obligations in hierarchical normative systems. In *Proceedings of the eighth international conference on artificial intelligence and law (ICAIL'03)*. Edimburgh, 109–118. ACM Press.

Boella, G., and L. van der Torre. 2006a. Constitutive norms in the design of normative multiagent systems. In *Computational logic in multi-agent systems, 6th international workshop, CLIMA VI*. LNCS, vol. 3900, 303–319. London: Springer.

Boella, G., and L. van der Torre. 2006b. A logical architecture of a normative system. In *Deontic logic and artificial normative systems. Proceedings of the 8th international workshop on deontic logic in computer scicence*, DEON 2006, Utrecht, ed. L. Goble and J. J. C. Meyer. Berlin: Springer.

Boella, G., and L. van der Torre. 2008. Institutions with a hierarchy of authorities in distributed dynamic environments. *Artificial Intelligence and Law* 16(1): 53–71.

Boella, G., G. Pigozzi, and L. van der Torre. 2009. Five guidelines for normative multiagent systems. In *JURIX*, Rotterdam, 21–30.

Booth, R., S. Kaci, and L. van der Torre. 2006. Merging rules: Preliminary version. In *Proceedings of the eleventh international workshop on non-monotonic reasoning (NMR'06)*, Lake District, UK, 2–5 June 2006.

Broersen, J. 2006. Strategic deontic temporal logic as a reduction to ATL, with an application to Chisholm's scenario. In *Proceedings 8th international workshop on deontic logic in computer science (DEON'06)*. Lecture Notes in Computer Science, vol. 4048, ed. L. Goble and J. J. Meyer, 53–68. Berlin: Springer.

Broersen, J. 2011. Deontic epistemic *stit* logic distinguishing modes of mens rea. *Journal of Applied Logic* 9(2): 127–152.

Broersen, J., M. Dastani, J. Hulstijn, and L. van der Torre. 2002. Goal generation in the BOID architecture. *Cognitive Science Quarterly Journal* 2(3–4): 428–447.

Broersen, J., F. Dignum, V. Dignum, and J. J. Meyer. 2004. Designing a deontic logic of deadlines. In *Proceedings 7th international workshop on deontic logic in computer science (DEON'06)*. Lecture Notes in Computer Science, vol. 3065, ed. A. Lomuscio and D. Nute, 43–56. Berlin: Springer.

Broersen, J., M. Dastani, and L. van der Torre. 2005. Beliefs, obligations, intentions and desires as components in an agent architecture. *International Journal of Intelligent Systems* 20(9): 893–920.

Broersen, J., R. Mastop, J. J. C. Meyer, and P. Turrini. 2008. A deontic logic for socially optimal norms. In *Proceedings 9th international workshop on deontic logic in computer science (DEON'08)*. Lecture Notes in Computer Science, vol. 5076, ed. L. v. d. Torre and R. v. d. Meyden, 218–232. Berlin: Springer.

Carmo, J., and A. Jones. 2002. Deontic logic and contrary-to-duties. In *Handbook of philosophical logic*, vol. 8, 2nd ed, ed. D. Gabbay and F. Guenthner, 265–344. Dordrecht: Kluwer Academic.

Castelfranchi, C., F. Dignum, C. M. Jonker, and J. Treur. 2000. Deliberative normative agents: Principles and architecture. In *6th international workshop on intelligent agents VI, agent theories, architectures, and languages (ATAL)*, 364–378. London: Springer.

Chisholm, R. 1963. Contrary-to-duty imperatives and deontic logic. *Analysis* 24(2): 33–36.

Cholvy, L., and F. Cuppens. 1999. Reasoning about norms provided by conflicting regulations. In *Norms, logics and information systems*, ed. P. McNamara and H. Prakken. Amsterdam: IOS.

Gabbay, D. M., G. Pigozzi, and J. Woods. 2003. Controlled revision – an algorithmic approach for belief revision. *Journal of Logic and Computation* 13(1): 3–22.

Goble, L. 2007. Prima facie norms, normative conflicts and dilemmas. In *Handbook of deontic logic and normative systems*, ed. D. Gabbay, J. Horty, R. van der Meyden, and L. van der Torre. London: College Publications.

Grégoire, E. 2004. Fusing legal knowledge. In *Proceedings of the 2004 IEEE international conference on information reuse and integration (IEEE-IRI'2004)*, Las Vegas, 522–529.

Grossi, D. 2007. Designing invisible handcuffs: Formal investigations in institutions and organizations for multi-agent systems. Ph.D. thesis, Utrecht University.

Grossi, D., J. J. C. Meyer, and F. Dignum. 2006. Classificatory aspects of counts-as: An analysis in modal logic. *Journal of Logic and Computation* 16(5): 613–643.

Hansen, J. 2005. Deontic logics for prioritized imperatives. *Artificial Intelligence and Law* 14: 1–34.

Hansen, J. 2008. Prioritized conditional imperatives: Problems and a new proposal. *Journal of Autonomous Agents and Multi-Agent Systems* 17(1): 11–35.

Hansson, B. 1969. An analysis of some deontic logics. *Noûs* 3: 373–398.

Herzig, A., E. Lorini, F. Moisan, and N. Troquard. 2011a. A dynamic logic of normative systems. In *Proceedings of the twenty-second international joint conference on artificial intelligence (IJCAI'11)*. Barcelona: IJCAI/AAAI.

Herzig, A., E. Lorini, and N. Troquard. 2011b. A dynamic logic of institutional actions (regular paper). In *Computational logic in multi-agent systems (CLIMA)*, LNC-S/LNAI, ed. J. Leite and P. Torroni. Berlin: Springer.

Horty, J. F. 2001. *Agency and deontic logic*. Oxford: Oxford University Press.

Horty, J. 2007. Defaults with priorities. *Journal of Philosophical Logic* 36: 367–413.

Jones, A., and M. Sergot. 1996. A formal characterisation of institutionalised power. *Journal of IGPL* 3: 427–443.

Kooi, B., and A. Tamminga. 2008. Moral conflicts between groups of agents. *Journal of Philosophical Logic* 37(1): 1–21.

Liu, F. 2008. Changing for the better: Preference dynamics and agent diversity. Ph.D. thesis. ILLC Dissertation Series, Amsterdam.

Loewer, B., and M. Belzer. 1983. Dyadic deontic detachment. *Synthese* 54(2): 295–318.

Lorini, E., and D. Longin. 2008. A logical account of institutions: From acceptances to norms via legislators. In *Proceedings of the international conference on principles of knowledge representation and reasoning (KR 2008)*, ed. G. Brewka and J. Lang, 38–48. Menlo Park: AAAI.

Lorini, E., D. Longin, B. Gaudou, and A. Herzig. 2009. The logic of acceptance: Grounding institutions on agents' attitudes. *Journal of Logic and Computation* 19(6): 901–940.

Makinson, D., and L. van der Torre. 2000. Input-output logics. *Journal of Philosophical Logic* 29(4): 383–408.

Makinson, D., and L. van der Torre. 2001. Constraints for input-output logics. *Journal of Philosophical Logic* 30(2): 155–185.

Makinson, D., and L. van der Torre. 2003. Permissions from an input-output perspective. *Journal of Philosophical Logic* 32(4): 391–416.

Meyer, J. J. C. 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.

Parent, X. 2003. Remedial interchange, contrary-to-duty obligation and commutation. *Journal of Applied Non-Classical Logics* 13(3/4): 345–375.

Parent, X. 2010. Moral particularism and deontic logic. In *Proceedings of the 10th international workshop on deontic logic (DEON'10)*, pp. 84–96, ed. G. Governatori and G. Sartor. Berlin/Heidelberg: Springer.

Parent, X. 2011. Moral particularism in the light of deontic logic. *Artificial Intelligence and Law* 19: 75–98.

Pauly, M. 2002. A modal logic for coalitional power in games. *Journal of Logic and Computation* 12(1): 149–166.

Prakken, H., and M. Sergot. 1997. Dyadic deontic logic and contrary-to-duty obligation. In *Defeasible deontic logic*, ed. D. Nute, 223–262. Dordrecht: Kluwer Academic.

Searle, J. 1969. *Speech acts. An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Stolpe, A. 1997. Relevance, derogation and permission. In *Defeasible deontic logic*, ed. D. Nute, 98–115. Dordrecht: Kluwer Academic.

Stolpe, A. 2010. A theory of permission based on the notion of derogation. *Journal of Applied Logic* 8(1): 97–113.

Thomason, R. H. 1981. Deontic logic as founded on tense logic. In *New studies in deontic logic*, ed. R. Hilpinen, 165–176. Dordrecht: D. Reidel.

van der Torre, L., and Y. H. Tan. 1997. The many faces of defeasibility in defeasible deontic logic. In *Defeasible deontic logic*, ed. D. Nute, 79–121. Dordrecht: Kluwer Academic.