Priority Structures in Deontic Logic

Johan van Benthem[1][2], Davide Grossi[3], Fenrong Liu[4]

[1] Institute for Logic, Language and Computation, University of Amsterdam
Science Park 904, 1098 XH, Amsterdam, The Netherlands
j.vanbenthem@uva.nl
[2] Department of Philosophy, Stanford University
Stanford, California, USA
[3] Department of Computer Science, University of Liverpool
d.grossi@liverpool.ac.uk
[4] Department of Philosophy, Tsinghua University
100084, Beijing, China
fenrong@tsinghua.edu.cn

Abstract. This paper proposes a systematic application of recent developments in the logic of preference to a number of topics in deontic logic. The key junction is the well known Hansson conditional for dyadic obligations. These conditionals are generalized by pairing them with reasoning about syntactic priority structures. The resulting two-level approach to obligations is tested first against standard scenarios of contrary-to-duty obligations, leading also to a generalization for the Kanger-Anderson reduction of deontic logic. Next, the priority framework is applied to model two intuitively different sorts of deontic dynamics of obligations, based on information changes and on genuine normative events. In this two-level setting, we also offer a new take on the vexed issue of strong permission. Finally, the priority framework is shown to provide a unifying setting for the study of operations on prioritized norm structure as such, in particular, adding or deleting individual norms, and even merging whole norm systems in different manners.

Key words: Deontic logic, preference logic, modal logic, dynamic logic, contrary-to-duty, norm change.

1 Introduction

There is a long-standing meta-ethical intuition that deontic notions of obligation, permission, and prohibition involve a normative "ideality ordering". This ordering ranks possible situations or courses of action according to how well they meet moral demands. Thus, [34] writes (as quoted in [52, p. 6]):

" [...] to assert that a certain line of conduct is [...] absolutely right or obligatory, is obviously to assert that more good or less evil will exist in the world, if it is adopted, than if anything else be done instead."

Accordingly, deontic logic has long considered models involving betterness ordering of worlds or states, going back at least to [20]. There, statements of dyadic obligation: "it ought to be the case that φ under condition ψ" ($\mathbf{O}(\varphi \mid \psi)$) were interpreted in terms of a binary relation $s \preceq t$ between states s, t:

$$\mathcal{M}, s \models \mathbf{O}(\varphi \mid \psi) \Longleftrightarrow \max_{\preceq}(\llbracket \psi \rrbracket_{\mathcal{M}}) \subseteq \llbracket \varphi \rrbracket_{\mathcal{M}} \tag{1}$$

Here $\mathcal{M}$ is a standard modal model on an ordering frame $\mathcal{F} = (S, \preceq)$, $\llbracket . \rrbracket_{\mathcal{M}}$ is the truth-set function of $\mathcal{M}$, and max is a 'selection function' picking out the $\preceq$-maximal elements in any given set.

This framework is quite flexible: depending on the properties of the relation $\preceq$, different logics can be obtained. [20] starts with a $\preceq$ which is only reflexive, moving eventually to total pre-orders that are reflexive, transitive and connected. Each step imposes stronger restrictions on how we make comparisons between worlds. Another type of variation makes the ordering $\preceq$ dependent on the world of evaluation. [5] This semantics has sparked a literature that continued well into the 1990s (cf. [49]), though it has also been criticized for less than ideal fit with intuitive deontic reasoning. By now, maximality-based ordering semantics has emerged in many branches of philosophical logic, including conditional logic, doxastic logic, and defeasible inference, [31].

In the present paper, we extend maximality-based ordering models and their applications to deontic logic by tapping into recent developments of this semantics in the area of preference logic (in particular, [28, 16, 29]).

Outline of the paper. Section 2 defines priority structures of syntactic criteria and shows how these relate to ideality or betterness orderings on states. Some technical results are in the Appendix. Section 3 looks at the simple modal logics induced by priority structures and relates them to Hansson-style deontic logics. This yields a two-level perspective on reasoning about ideality orderings whose usefulness we illustrate by three applications to deontic logic. Section 4 applies priority structures to the benchmark problem of contrary-to-duty obligations (cf.[38]) and the Kanger-Anderson reduction. Sections 5 and 6 explain how our setting supports a variety of dynamic events that change obligations. Using techniques from [28, 29]—inspired by normative actions such as commands—we show how betterness dynamics on states and on prioritized syntactic criteria work in harmony, delivering a novel analysis of several deontic phenomena. This brings deontic logic closer to current dynamic logics of belief and preference change [44]. Finally, Section 7 identifies further lines of research that now open up. Throughout the paper we make use of and elaborate on examples taken from the literature in deontic logic.

---

[5] [27] is an overview of various moves in the early literature on dyadic obligation.

2   Priorities and betterness

2.1   Reasons for ordering

The betterness relation between situations that grounds our obligations (or preference in general) often stems from an explicit code for what is right or wrong. Consider this rousing quote from St. Paul's First Letter to the Corinthians:

"It is good for a man not to touch a woman. But if they cannot contain, let them marry: for it is better to marry than to burn." (cf. [52, p. 6])

This passage identifies three properties of states: most ideal is that men do not touch women, less ideal is the more permissive property of men either not touching women or marrying them, and most permissive, and, least ideal in terms of possible punishment, is 'none of the above', leading to the trivial property $\top$. This is not a scenario of deontic reasoning, but of an equally important process of norm giving. We can formalize the latter as a sequence of relevant propositions:

$$(\neg t \vee m \vee \neg m) \prec (\neg t \vee m) \prec \neg t \tag{2}$$

We will call structures of this kind 'priority sequences' of [28], and in their general format of 'priority graphs', they will be the focus of this paper.[6]

2.2   Priority graphs and derived betterness ordering

Definition 1 (P-graphs). *Let $\mathcal{L}(\mathbf{P})$ be a propositional language built on the set of atoms $\mathbf{P}$. A P-graph is a tuple $\mathcal{G} = \langle \Phi, \prec \rangle$ such that:*

- *$\Phi \subset \mathcal{L}(\mathbf{P})$ with $|\Phi| < \omega$;*
- *$\prec$ is a strict order on Φ : property ψ is strictly better than φ; also, for all propositions $\varphi, \psi \in \Phi$: if $\varphi \prec \psi$ then φ logically implies ψ.*

Intuitively, a P-graph is a finite graph of formulae from a propositional language, where each formula logically implies its immediate successor in the order. In what follows, we will often drop the index referring to the current language.

P-graphs where $\prec$ is a strict linear order are called *P-sequences*, they are referred to as sequences $\langle \varphi_1, \ldots, \varphi_n \rangle$ of propositional formulae and are denoted by letter $\mathcal{S}$. Given a P-graph $\mathcal{G} = \langle \Phi, \prec \rangle$ we denote by $S_\mathcal{G}$ the set of maximal P-sequences (i.e., $\prec$-chains) included in the order $\prec$.

Here is how a given P-graph determines a betterness ordering on states:

Definition 2 (State betterness from P-graphs). *Let $\mathcal{G} = \langle \Phi, \prec \rangle$ be a P-graph, S a non-empty set of states and $\mathcal{I} : \mathbf{P} \longrightarrow 2^S$ a valuation function. The betterness relation $\preceq_\mathcal{G} \subseteq S^2$ is defined as follows:*

$$s \preceq_\mathcal{G} s' := \forall \varphi \in \Phi : s \in [\![\varphi]\!]_\mathcal{I} \Rightarrow s' \in [\![\varphi]\!]_\mathcal{I}. \tag{3}$$

The function outputting this pre-order is called sub *(from 'subsumption').*

---

[6] One can also construe the order of properties differently, with not-touching on top, and marrying as second best. We return to options in extracting priorities below.
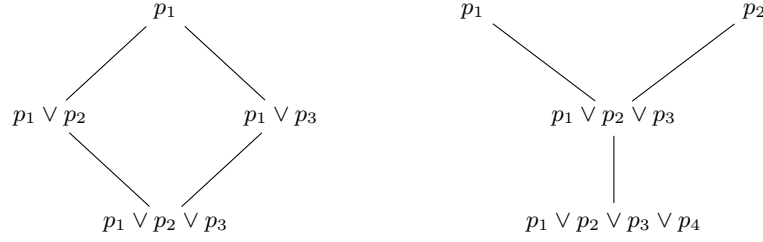
Fig. 1. Hasse diagrams of P-graphs

Definition 2 orders states in S according to which elements of the P-graph they satisfy. If a state satisfies a property in the graph, it also satisfies all $\prec$-worse properties in the graph. Here are some useful properties:

Fact 1 (Basic properties of $\preceq_{\mathcal{G}}$) *Let $\mathcal{G} = \langle \Phi, \prec \rangle$ be a P-graph. For any valuation $\mathcal{I} : \mathbf{P} \longrightarrow 2^S$ it holds that:*

1. *The relation $\preceq_{\mathcal{G}}$ is a pre-order whose strict part $\prec_{\mathcal{G}}$ is upward well-founded;*
2. *If $\varphi_i \prec \varphi_j$, then for all worlds $s \in [\![\varphi_i]\!]_{\mathcal{I}}, s' \in [\![\varphi_j]\!]_{\mathcal{I}}: s \preceq_{\mathcal{G}} s';$*
3. *If $\varphi_i \prec \varphi_j$, then for all worlds $s \in [\![\varphi_i \wedge \neg\varphi_j]\!]_{\mathcal{I}}, s' \in [\![\varphi_j]\!]_{\mathcal{I}}: s \prec_{\mathcal{G}} s'.$*

Remark 1 (From inclusion to priority as primitive). We chose to make P-graphs consist of propositions ordered by (proper) inclusion (Definition 9). These induce a total preorder on states (Definition 2). An alternative definition (see [30]) allows unrelated properties in the graph. An ordering on worlds is then induced by the following map `lex` (from 'lexicographic').:

$$
\begin{aligned}
s \preceq_{\mathcal{G}}^{\texttt{lex}} s' \Longleftrightarrow \forall \varphi \in \Phi : [s \in [\![\varphi]\!]_{\mathcal{I}} \Rightarrow s' \in [\![\varphi]\!]_{\mathcal{I}} \\
\text{or } \exists \varphi' : [\varphi \prec \varphi' \text{ and } s \notin [\![\varphi']\!]_{\mathcal{I}} \text{ and } s' \in [\![\varphi']\!]_{\mathcal{I}}]].
\end{aligned}
\tag{4}
$$

State s is better than s' if and only if at the first φ in the sequence where they differ qua properties, s has φ while s' lacks it. This formulation can be proven equivalent to ours, cf. the Appendix to this paper.

3 Reasoning about betterness

Semantic models, even enriched with syntactic priority structure, do not yet tell us how agents reason about their obligations. We now present a simple logic capturing reasoning with betterness orderings induced by priority graphs.

3.1 A logic of pre-orders

The definitions that follow here are standard in current modal logics of preference, and we refer to [10, 16] for details and further motivation:

Language and semantics. The basic modal language of preference $\mathcal{L}(\mathsf{U}, \preceq)$ is built from a countable set $\mathbf{P}$ of atoms according to the following inductive syntax:

$$\mathcal{L}(\mathsf{U}, \preceq) : \varphi ::= p \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi \mid [\preceq]\varphi \mid [\mathsf{U}]\varphi$$

Here $p \in \mathbf{P}$. Existential modalities and Boolean operators are defined as usual. The intended meaning shows in the semantic structures of Section 2:

Definition 3 (Models). *A model for $\mathcal{L}(\mathsf{U}, \preceq)$ on the set of atoms $\mathbf{P}$ is a tuple $\mathcal{M} = \langle S, \preceq, \mathcal{I} \rangle$ where:*

  i) S is a non-empty set of states,
  ii) $\preceq$ is a preorder over S ("at least as good as"), and
5 *iii) $\mathcal{I} : \mathbf{P} \longrightarrow 2^S$.*

A pointed model is one with a distinguished world s: $(\mathcal{M}, s)$. We also the strict suborder $\prec$ ("strictly better than") of $\preceq$ as usual: $s \prec s'$ iff $s \preceq s'$ and $s' \not\preceq s$.

Here are the key clauses for our central modal operators:

Definition 4 (Satisfaction). *Let $\mathcal{M} = \langle S, \preceq, \mathcal{I} \rangle$ be a model. Truth for a for-*
10 *mula $\varphi \in \mathcal{L}(\mathsf{U}, \preceq)$ in a pointed model $(\mathcal{M}, s)$ is defined inductively:*

$$\mathcal{M}, s \models p \Longleftrightarrow w \in \mathcal{I}(p)$$
$$\mathcal{M}, s \models [\preceq]\varphi \Longleftrightarrow \forall s' \in S \text{ s.t. } s \preceq s' : \mathcal{M}, s' \models \varphi$$
$$\mathcal{M}, s \models [\mathsf{U}]\varphi \Longleftrightarrow \forall s' \in S : \mathcal{M}, s' \models \varphi$$

A useful notation is this: $[\![\varphi]\!]_{\mathcal{M}}$ denotes the truth-set of φ in $\mathcal{M}$.

$[\mathsf{U}]$-formulae define global properties of a model, $[\preceq]$-formulae local properties true in all states at least as good as the current state.

Axiomatics. A complete axiomatic proof calculus for our system consists of the
15 standard modal logic **S4** for betterness, **S5** axioms for the universal modality, and one inclusion axiom **Incl**. This logic is known to be sound and strongly complete for pre-orders. Its uses go back to [10].

    Having formulated the basic logic of preferential reasoning, let us now investigate our models a little more closely for their fine-structure.

20 *Models from P-graphs.* Given a P-graph $\mathcal{G}$ and a propositional valuation function $\mathcal{I}$, Definition 2 yields models of the above type (Definition 3) $\mathcal{M} = \langle S, \preceq_{\mathcal{G}}, \mathcal{I} \rangle$ where $\preceq_{\mathcal{G}}$ is the total pre-order derived from $\mathcal{G}$.[7]

---

[7] Similar structures occur in other areas. Each P-graph plus valuation induces a Lewis system of spheres in conditional logic, or a doxastic Grove model. The extra syntax has analogies with default logic, and entrenchment for beliefs [35].

Expressive power: defining semantic 'best'. Our modal language can define various notions used in our semantic analysis of Section 2:

Fact 2 *On total pre-orders $\preceq$ with an upward well-founded strict part $\prec$, conditional obligation in the sense of [20] is defined as follows:*

$$\mathbf{O}_{\preceq}(\varphi \mid \psi) := [\mathsf{U}](\psi \to \langle \preceq \rangle(\psi \wedge [\preceq](\psi \to \varphi))) \tag{5}$$

This equivalence may be understood as follows. On well-founded ordering models, all maximal states satisfy φ iff, for every world, there is a better world all of whose better alternatives satisfy φ, i.e., $[\mathsf{U}]\langle \preceq \rangle[\preceq]\varphi$. Relativizing in the usual manner to a formula ψ yields Formula (5).

Thus, on the right kind of models, modal logics can also deal with global conditional operators. Moreover, it is easy to add expressive power to this semantics and axiomatics. For instance, one can add modalities for strict betterness (cf. [46]) and extend both the model and the axiomatization to deal with these.

3.2 Adding priorities: fitting in syntactic 'best'

Our modal logic talks about semantic ordering. What about the syntactic priority structure of Section 2? [30]) discusses enriched logics for such 'two-level structures', but we will only show the harmony between the two levels.

Syntactic 'best' operator. The definition of maximality in Formula (5) relies on an underlying order. In terms of P-graphs $\mathcal{G}$, the matching intuition is that the best states under condition ψ are those that belong to the most ideal properties in the P-graph compatible with ψ, in the sense of non-empty intersection of their truth-sets. Call this more syntactic notion $\mathbf{O}_{\mathcal{G}}(. \mid .)$. It says: 'the best properties of $\mathcal{G}$ that are compatible with ψ all imply φ'.

Given a graph $\mathcal{G}$, this is definable in our modal language as follows:

$$\mathbf{O}_{\mathcal{G}}(\varphi \mid \psi) := [\mathsf{U}]\left(\left(\bigvee_{\langle \varphi_1, \ldots, \varphi_n \rangle \in S_{\mathcal{G}}} \bigwedge_{1 \leq i \leq n} (\langle \mathsf{U} \rangle(\varphi_i \wedge \psi) \to (\varphi_i \wedge \psi)) \right) \to \varphi \right) \tag{6}$$

Here, for $\langle \varphi_1, \ldots, \varphi_n \rangle \in S_{\mathcal{G}}$, $\bigwedge_{1 \leq i \leq n} (\langle \mathsf{U} \rangle(\varphi_i \wedge \psi) \to (\varphi_i \wedge \psi))$ denotes the maximal element of the P-sequence $\langle \varphi_1, \ldots, \varphi_n \rangle$ that has a non-empty intersection with ψ. Then $\bigvee_{\langle \varphi_1, \ldots, \varphi_n \rangle \in S_{\mathcal{G}}} \bigwedge_{1 \leq i \leq n} (\langle \mathsf{U} \rangle(\varphi_i \wedge \psi) \to (\varphi_i \wedge \psi))$ takes the union of these, and Formula (6) states that this is subsumed by φ.

It is important to stress that the definition in Formula (6) makes $\mathbf{O}_{\mathcal{G}}$ a non-monotonic operator, even though its syntactic form, at first inspection, seems to suggest the contrary: $[\mathsf{U}](\xi(\psi) \to \varphi)$ where $\xi(\psi)$ is the syntactic schema which, applied to ψ, denotes the maximal ψ-elements in $\mathcal{G}$. Now this syntactic structure makes antecedent strengthening for formulae $[\mathsf{U}](\xi(\psi) \to \varphi)$ unsound. The reason is that $\xi(\psi \wedge \psi')$ (i.e., the union of the non-empty maximal properties of $\mathcal{G}$ compatible with $\psi \wedge \psi'$) does not logically imply the union $\xi(\psi)$.

Correspondence of syntactic and semantic 'best'. The earlier operator $\mathbf{O}_{\preceq}$ defined in Formula 5 and our new operator $\mathbf{O}_{\mathcal{G}}$ are in complete harmony:

Theorem 1 (Correspondence). *Let* $\mathcal{G} = \langle \Phi, \prec \rangle$ *be a P-graph,* $\mathcal{M}_{\mathcal{G}}$ *a model derived by Definition 2 from* $\mathcal{G}$*,* $\mathcal{I}$ *a valuation function and* s *a state:*

$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}_{\preceq}(\varphi \mid \psi) \Longleftrightarrow \mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}_{\mathcal{G}}(\varphi \mid \psi).$$

Proof. This claim is proven by this series of equivalent statements:

$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}_{\preceq}(\varphi \mid \psi)$$
$$\forall s' \in Max_{\preceq_{\mathcal{G}}}(\llbracket \psi \rrbracket_{\mathcal{M}}) : \mathcal{M}_{\mathcal{G}}, s' \models \varphi$$
$$\forall s' \in \left\llbracket \bigvee_{\langle \Phi_1, \ldots, \varphi_n \rangle \in S_{\mathcal{G}}} \bigwedge_{1 \leq i \leq n} (\langle \mathsf{U} \rangle (\varphi_i \wedge \psi) \rightarrow (\varphi_i \wedge \psi)) \right\rrbracket : \mathcal{M}_{\mathcal{G}}, s' \models \varphi$$
$$\mathcal{M}_{\mathcal{G}}, s \models [\mathsf{U}] \bigwedge_{1 \leq i \leq n} (\neg \langle \mathsf{U} \rangle (\psi \wedge \varphi_{i-1}) \wedge \langle \mathsf{U} \rangle (\psi \wedge \varphi_i) \rightarrow [\mathsf{U}](\psi \wedge \varphi_i \rightarrow \varphi))$$
$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}_{\mathcal{G}}(\varphi \mid \psi)$$

The first equivalence holds by Formula 5 and the fact that the strict part of $\preceq_{\mathcal{G}}$ is upward well-founded (Fact 1). The second equivalence uses the fact that a state is in $Max_{\preceq_{\mathcal{G}}}$ iff it is in some truth-set of a maximal element in $S_{\mathcal{G}}$ (recall Formula (6)). The fourth and fifth equivalences hold by Definition 4) and Formula 6.   □

This correspondence adds an interesting layer to the **S4** logic of preference. It relates conditional maximality statements induced by priority structure $\mathbf{O}_{\preceq}(\varphi \mid \psi)$ to subsumption statements $[\mathsf{U}](\xi(\psi) \rightarrow \varphi)$. The latter capture a natural way of reasoning about obligations: first, use the priority structure to find best properties compatible with the given conditions, then apply a simple subsumption check. Theorem 1 guarantees the equivalence of this explicit method with standard ideality semantics. Given this correspondence, we will often write a deontic modality $\mathbf{O}$ to denote either $\mathbf{O}_{\preceq}$ or $\mathbf{O}_{\mathcal{G}}$.

The rest of the paper takes this two-level view of betterness to deontic logic, where priority-based reasoning arguably plays a central role in the way we commonly conceptualize the notion of obligation.

4   First application: contrary-to-duty reasoning

The section applies P-graphs to classic topics in deontic logic: contrary-to-duty obligations and the so-called Kanger-Anderson reduction of deontic logic.

4.1   Introducing priorities: contrary-to-duty obligations

Syntactic priorities make sense beyond preference. In deontic settings, the 'letter of the law' may be as important as its effects on situations. Consider two classic cases of contrary-to-duty obligations (CTDs):
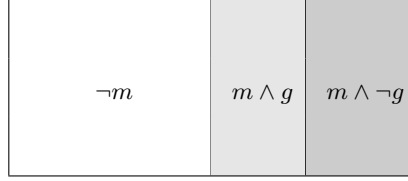
Fig. 2. Gentle Murder as an ordered partition of a state space in clusters of equally preferred states, going from more to less ideal from left to right.

Example 1 (Gentle murder). Here is our first example ([15, p. 194]):

> "Let us suppose a legal system which forbids all kinds of murder, but which considers murdering violently to be a worse crime than murdering gently. [...] The system then captures its views about murder by means of a number of rules, including these two:
> 1. It is obligatory under the law that Smith not murder Jones.
> 2. It is obligatory that, if Smith murders Jones, Smith murders Jones gently."

The scenario mentions two classes of states: those in which Smith does not murder Jones, represented by the formula $\neg m$; and those in which either Smith does not murder Jones or he does murders Jones, but gently, i.e., $\neg m \vee (m \wedge g)$. We thus have a P-sequence $(\neg m \vee (m \wedge g)) \prec \neg m$. The induced betterness relation orders worlds in three disjoint clusters, shown in Figure 1: the most ideal states satisfy $\neg m$, strictly worse but not worst are the states satisfying $m \wedge g$ and strictly worst are the states with $m \wedge \neg g$.

Our second illustration adds a few touches to the preceding one cf. [3]):

Example 2 (The Chisholm scenario).

1. It ought to be that Smith refrains from robbing Jones.
2. Smith robs Jones.
3. If Smith robs Jones, he ought to be punished for robbery.
4. It ought to be that if Smith refrains from robbing Jones he is not punished for robbery.

Again, three items specify a priority sequence: $\neg r \vee (r \wedge p) \prec (\neg r \wedge \neg p)$ (where r stands for "Smith robs Jones" and p for "Smith is punished"). It is most ideal that Smith refrains to rob Jones and, at the same time, that Smith is not punished; and that the most ideal states under the assumption that Smith robs Jones are states in which Smith is punished. The main difference with Example 1 is the factual statement "Smith robs Jones", whose informational point will be our cue for our dynamic treatment in Section 5.

It is worth making a methodological point about our formalization of the above deontic scenarios. Our use of P-graphs emphasizes a non-inferential aspect of deontic logic. Obligations and permissions are handled as usual given their induced normative ordering. But equally important are the criteria themselves that generate the normative order on which our judgments of obligation are based. This criterial structure supports many specific deontic inferences, and hence it may be considered a more permanent part of what drives an agent's behavior. In other words, our formalization should not just aim for extracting obligation and permission operators in a text, but also for cues as to the normative priorities—and what we should maintain generally, is not just a record of the reasoning: the text also helps us maintain a record of the relevant normative priorities.

4.2   CTDs and priority structures

The literature on deontic logic has used CTDs as benchmarks to evaluate logics of conditional obligation with respect to the sort of reasoning they support (see [37]). We think of P-graphs as the natural formalization of a CTD: some norms are given, and obligations are computed going 'down the line'. What is the difference here with standard semantic models? A mere betterness order on worlds has, so to speak, forgotten its origins—whereas now we have these still available as reasons for the ordering, and relevant exception zones.

Theorem 1 guarantees that conditional obligations can be captured by simple subsumptions on a priority graph. Instantiating the theorem with Example 1 gives the following representation of the CTD in the Gentle murder scenario (recall Formula (6)):

$$\mathbf{O}(g \mid m) \Longleftrightarrow [\mathsf{U}]((\langle\mathsf{U}\rangle(m \wedge \neg m) \rightarrow (\neg m \wedge m) \wedge (\langle\mathsf{U}\rangle(m \wedge g) \rightarrow (m \wedge g)) \rightarrow g)$$
$$\Longleftrightarrow [\mathsf{U}]((m \wedge g) \rightarrow g)$$

Similarly, the conditional obligation in the Chisholm Example 2 become:

$$\mathbf{O}(p \mid r) \Longleftrightarrow [\mathsf{U}]((r \wedge p) \rightarrow p)$$
$$\mathbf{O}(\neg p \mid \neg r) \Longleftrightarrow [\mathsf{U}]((\neg p \wedge \neg r) \rightarrow \neg p)$$

for the P-sequence $\langle(\neg p \wedge \neg r), \neg r \vee (r \wedge p)\rangle$. Thus, reasoning about conditional obligation becomes simple analysis of norms.

We are not claiming a totally new analysis of CTDs. Ideal versus 'subideal' zones occur in [24, 42], and even the classic [20] says this:

"The problem of conditional obligation is what happens if somebody nevertheless performs a forbidden act. Ideal worlds are excluded. But, it may be the case that among the still achievable worlds, some are better than others. There should then be an obligation to make the best out of the sad circumstances."

Chains of properties are used in [52], [4][8] and [17][9]. Our contribution is doing this in a general format (P-graphs) that has proven to work in other areas, such as plausibility-based belief [28]. In particular, as we shall see, priority-based CTDs can model the deontic phenomenon of "factual defeasibility" [50]: new information about the world may shift our obligations.

4.3   The Kanger-Anderson reduction and P-sequences

5   We now connect our analysis to a proposal from the founding period of deontic logic. Anderson [1] and Kanger [25] reduced deontic $\mathbf{O}$-formulae to alethic modal $\Box$-formulae with a constant for violation V or ideality I:

$$\mathbf{O}\varphi := \Box(\neg\varphi \to \mathsf{V}) \tag{7}$$

$$\mathbf{O}\varphi := \Box(\mathsf{I} \to \varphi). \tag{8}$$

This reductionist view has been criticized in the deontic logic literature by observing that it cannot accommodate a satisfactory representation of CTDs. We
10 will show briefly how P-graphs, properly specialized, offer a natural extension to Anderson's and Kanger's proposals that does deal with CTDs along the lines they advocated.

Definition 5 (KA-sequences). *Let* $\{\mathsf{I}_1,\ldots,\mathsf{I}_n\} \subseteq \mathbf{P}$. *A Kanger-Anderson sequence ("KA-sequence") for* $\mathcal{L}(\mathbf{P})$ *is a sequence defined as follows:*

$$\left\langle \bigvee_{1 \le j \le i} \mathsf{I}_j \right\rangle_{1 \le i \le n}$$

15 So, KA-sequences are tuples $\langle \mathsf{I}_1, \mathsf{I}_1 \vee \mathsf{I}_2, \ldots, \mathsf{I}_1 \vee \ldots \vee \mathsf{I}_n \rangle$ which are built by using ideality atoms to construct n layers spanning from the most to the least ideal.

Theorem 1 specializes to KA-sequences in an interesting way:

---

[8] [4] uses a special variant of P-sequences $\langle q_1,\ldots,q_n \rangle$ where each q_i and q_j—called there *systematic frame constants*—are pairwise disjoint. So q_1 states are strictly better than q_2 and so on. Note that his is a special case of the sort of graphs we mentioned in Remark 1, and it is equivalent to our P-sequences. The interested reader is referred to Theorem 4 in the Appendix.

[9] [17] proposes an approach to CTDs where a Gentzen calculus is developed for handling formulae of the type $\varphi_1 \otimes \ldots \otimes \varphi_n$ with $\otimes$ being a connective representing a reparatory obligation in a CTD structure: φ_1 ought to be the case but if φ_1 is violated then φ_2 ought to be the case and so on up to φ_n. Unlike this proof-theoretic approach, our approach is geared towards semantics and aims at connecting such CTD structures to modal logics interpreted on orders.

Corollary 1 (Obligations from better to worse). *Let $\mathcal{G}$ be a KA-sequence. For any model $\mathcal{M}_{\mathcal{G}}$, state s, and $1 \leq i < n$ it holds that:*

$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}(\varphi \mid \top) \Longleftrightarrow \mathcal{M}_{\mathcal{G}}, s \models [\mathsf{U}](\mathsf{I}_1 \to \varphi) \qquad (9)$$

$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}(\varphi \mid \mathsf{I}_1) \Longleftrightarrow \mathcal{M}_{\mathcal{G}}, s \models [\mathsf{U}](\mathsf{I}_1 \to \varphi) \qquad (10)$$

$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}(\varphi \mid \neg \bigvee_{1 \leq j \leq i} \mathsf{I}_j) \Longleftrightarrow \mathcal{M}_{\mathcal{G}}, s \models [\mathsf{U}](\bigvee_{1 \leq j \leq i+1} \mathsf{I}_j \to \varphi) \qquad (11)$$

$$\mathcal{M}_{\mathcal{G}}, s \models \mathbf{O}(\varphi \mid \neg \bigvee_{1 \leq j \leq n} \mathsf{I}_j) \Longleftrightarrow \mathcal{M}_{\mathcal{G}}, s \models [\mathsf{U}](\neg \bigvee_{1 \leq j \leq n} \mathsf{I}_j \to \varphi) \qquad (12)$$

Formula 9 says that an unconditional obligation $\mathbf{O}(\varphi \mid \top)$ is what the most ideal states dictate. The corollary shows how obligations changes as we move from most to least ideal circumstances. In most ideal states, where I_1 holds, what ought to be the case is what already is the case (Formula 10). Formula 11 states that, if the ith element has been violated, what ought to be is what follows from the $(i+1)^{th}$ element in the sequence. And in the least ideal states, where I_n is false, what ought to be the case is again what is already the case (Formula 12). Tis generalizes the Anderson-Kanger reduction to CTD reasoning. [10]

5 Second application: information dynamics in deontic settings

So far we have proposed a more richly structured priority-based model for deontic notions under static circumstances. But deontic reasoning is crucially also about changes. To deal with this, we use some basic methods from dynamic-epistemic logic (DEL), a current framework for dealing with actions that change agents' information about possible worlds, or their evaluation of these worlds.

5.1 Logical information dynamics

We just state some basics, much more information can be found in [44]. Epistemic logic describes what agents know on the basis of their current semantic information. Public announcement logic (PAL), the simplest case of DEL, combines epistemic logic with one dynamic event, namely, the *announcement* of new 'hard information' expressed in some proposition φ. The corresponding action $!\varphi$ transforms a current epistemic model $(\mathcal{M}, s)$ into its submodel $(\mathcal{M}|\varphi, s)$ where all worlds that did not satisfy φ have been eliminated. This reflects the basic intuition of information gain, both in science and in common sense, as shrinking one's current epistemic range of uncertainty. The typical new dynamic formula $[!\varphi]\psi$ of this system says that "after announcing the true proposition φ, formula ψ holds". Here is its semantics:

$$\mathcal{M}, s \models [!\varphi]\psi \Longleftrightarrow \text{ if } \mathcal{M}, s \models \varphi \text{ then } \mathcal{M}|\varphi, s \models \psi. \qquad (13)$$

---

[10] A variant of Corollary 1 occurs in the preference logic of [4], with special propositions $\langle q_1, \ldots, q_n \rangle$. See earlier footnote 4.

PAL has been used as a pilot example for the analysis of a variety of epistemic changes capturing the many subtleties involved in, e.g., belief revision [43], but also of deontic changes [6]. Agents must constantly cope with changes in information because they learn more.

In particular, deontic dynamics under hard information extends our modal axiom system for betterness reasoning in Section 3 with further axioms for the dynamic modalities $[!\varphi]$. A typical reduction axiom of the latter kind is this:

$$[!\varphi]\,[\,\preceq\,]\psi \leftrightarrow (\varphi \to [\,\preceq\,][!\varphi]\psi) \tag{14}$$

That is, ψ holds in all better states after the announcement of φ if and only if either φ is not the case or all better states are such that, after the announcement of φ, they satisfy ψ.[11] Given the definition of conditional obligation by Formula 5, one can then derive the following reduction axiom:

$$[!\varphi]\,\mathbf{O}(\psi \mid \chi) \leftrightarrow (\varphi \to \mathbf{O}([!\varphi]\psi \mid (\varphi \wedge [!\varphi]\chi))) \tag{15}$$

As a special case for unconditional obligation, setting χ to the always true proposition $\top$, we get the reduction axiom:

$$[!\varphi]\mathbf{O}\psi \leftrightarrow (\varphi \to \mathbf{O}([!\varphi]\psi \mid \varphi)) \tag{16}$$

Factual formulas Keeping track of the dynamic modalities in this way is crucial here since complex modal statements may change their truth values when a model changes. However, our subsequent illustrations will usually be about purely propositional *factual formulas*, whose truth values cannot change under update of models, and then some simplifications will arise.

We will now show how this simple dynamic logic allows us to model some basic deontic phenomena involved with information change. Dynamic-epistemic methods can also model more properly normative changes in how agents evaluate situations or worlds, but we postpone these till later on.

5.2   Information dynamics in conditional obligations

We have seen that obligations are typically conditional, so changes in circumstances determine changes in what ought to be the case. Semantically, this means that maximally ideal states change under different circumstances, while syntactically, this means that properties in the priority structure that are incompatible with current circumstances can be disregarded. This is often called *deontic detachment*: conditional obligations remain stable, but what changes is what follows from them under different circumstances. We will show that our structured models naturally give us these two faces of deontic dynamics.

We start by revisiting an earlier scenario:

---

[11] It is not difficult to see that Formula 14 is valid on pre-orders. Proofs can be found in [44, 29].

Example 3 (The Chisholm scenario: a dynamic perspective). The Chisholm scenario consisted of three normative statements:

- It ought to be that Smith refrains from robbing Jones.
- If Smith robs Jones, he ought to be punished for robbery.
- It ought to be that if Smith refrains from robbing Jones he is not punished for robbery.

plus a factual one:

- Smith robs Jones.

Now intuitively, there is a difference in function here. The normative statements seem global guides to behavior, but the scenario suggests a dynamic reading of the factual statement. One way of doing so is by taking the robbery to be an event that takes place, changing the world.[12] But in this paper, we rather want to model deontic deliberation, where relevant facts become known, and enter the reasoning. In this line, the acquisition of the factual statement is a dynamic-epistemic event where the information that "Smith robs Jones" becomes available. This triggers precise normative consequences, namely that Smith ought to be punished. Thus, a conditional obligation is 'in suspended animation' until we get the hard information that its antecedent obtains. Then the model changes, and the conditional obligation becomes an absolute one. An illustration is the following dynamic formula matching Example 2):

$$(\mathbf{O}(\neg r \mid \top) \wedge \mathbf{O}(p \mid r)) \to [!r]\mathbf{O}(p \mid \top).^{13} \tag{17}$$

Intuitively: if it ought to be the case that $\neg r$, that p if r, then, after it is learnt that r is the case, it ought to be the case that p. The formula is true in all betterness **S4** models derived by the priority sequence of the Chisholm scenario.

Formula 17 is a special case of the following validity of our logic, that holds for factual formulas:

$$\mathbf{O}(\varphi \mid \psi) \to [!\psi]([\mathsf{U}]\psi \wedge \mathbf{O}(\varphi \mid \top))^{14} \tag{18}$$

If we interpret the universal modality $[\mathsf{U}]$ as an epistemic operator, as not uncommon in epistemic logic, Formula 18 formalizes the following intuitive principle: "If it ought to be the case that φ under condition that ψ then, if it is announced that ψ is the case, it is known that ψ and it ought to be (unconditionally) the case that φ". This suggests a natural form of conditional obligations taken in an epistemic sense: I should do something if I *know* the antecedent to be the case. A number of telling cases of such epistemic notions behind deontic logic were proposed in [36].

---

[12] Modeling physical events would take us to temporal deontic logics, [51, 23].

[13] Often, following standard deontic notation, we will write just $\mathbf{O}\varphi$ for $\mathbf{O}(\varphi \mid \top)$.

[14] The restriction to factual formulas is useful, but not essential: cf. [29].

Remark 2 (Deontic detachment). Turning from concrete scenarios to general reasoning, our dynamic logic also illuminates deontic discussions about putative inference principles. So-called *deontic detachment* has the form

$$(\mathbf{O}(\varphi \mid \psi) \wedge \psi) \to \mathbf{O}\varphi. \tag{19}$$

Our betterness semantics of Section 3 clearly does not validate this principle.[15] However, the following is valid:

$$\mathbf{O}(\varphi \mid \psi) \to [!\psi]\mathbf{O}\varphi \tag{20}$$

for factual ψ, φ. We feel that the usual justifications of detachment are really about this dynamic version, where ψ is not just the case, but it is *learnt that ψ* is the case. A static principle that does come close to this is

$$(\mathbf{O}(\varphi \mid \psi) \wedge [\mathsf{U}]\psi) \to \mathbf{O}\varphi. \tag{21}$$

where $[\mathsf{U}]$ is again interpreted as an epistemic operator. We leave its derivation in our logic to the reader.

Our dynamic logic is a style of deontic reasoning mixing effects of informational events with unpacking of duties. This extends the repertoire of principles that deontic logic is about. Here is one more illustration. It seems reasonable to assume that conditional obligations represent global norms for behavior that are stable under information increase. But this cannot be true, since the following equivalence is clearly non-valid:[16]

$$\mathbf{O}(\varphi \mid \psi) \to [!\chi]\mathbf{O}(\varphi \mid \psi) \tag{22}$$

Shifting to other areas of the priority structure can lead to different conditional obligations, as one more instance of their non-monotonicity. But what can be derived is, for instance, the special case $\mathbf{O}(\varphi \mid \psi) \to [!\psi]\mathbf{O}(\varphi \mid \psi)$ following from Formula (21).

Remark 3 (Non-monotonicity of conditional obligations). In the context of the above discussion, it is fitting to spend a few words about one of the key critiques that has been moved in the deontic logic literature against the formalization of conditional obligations *à la* Hansson. Conditionals $\mathbf{O}(\varphi \mid \psi)$ are clearly non-monotonic. From our epistemic standpoint, this means that if $\mathbf{O}(\varphi \mid \psi)$, by learning $\psi \wedge \chi$ one cannot conclude that $\mathbf{O}(\varphi)$, i.e.: $\mathbf{O}(\varphi \mid \psi) \to [!(\psi \wedge \chi)]\mathbf{O}\varphi$ is not valid. A neatly formulated critique to this aspect of our conditionals can be

---

[15] A simple countermodel is $\mathcal{M} = \langle \{s, s'\}, \preceq, \mathcal{I} \rangle$ with $s \preceq s'$ and $\mathcal{I}(p) = \mathcal{I}(q) = s$. Then $\mathcal{M}, s \models q \wedge \mathbf{O}(p \mid q)$ but $\mathcal{M}, s \not\models \mathbf{O}(p \mid \top)$, as the maximal state s' falsifies p. Detachment makes sense in special conditional or epistemic settings that require assumptions of "centering" for worlds around the current one.

[16] A simple countermodel is $\mathcal{M} = \langle \{s, s'\}, \preceq, \mathcal{I} \rangle$ with $s \preceq s'$ and $\mathcal{I}(p) = s'$ and $\mathcal{I}(q) = s$. Then $\mathcal{M}, s \models q \wedge \mathbf{O}(p \mid p \vee q)$ but $\mathcal{M}, s \not\models [!q]\mathbf{O}(p \mid (p \vee q))$, as in $\mathcal{M}|_q$ the maximal states satisfy $\neg p$.

found in [22], where it is argued that although monotonicity is not a desirable principle for conditional obligation, some limited form of monotonicity should still be available to draw "safe" monotonic conclusions in the logic, e.g., via some non-standard consequence relation.

The key point is made by using these two statements: 1) You ought to put your napkin on your lap; 2) If you are served asparagus, you ought to eat it with your fingers. In our logic this would be represented by $\mathbf{O}(n \mid \top)$ and $\mathbf{O}(f \mid a)$. Now the criticism of [22] is that the logic does not support the inference to $\mathbf{O}(n \mid a)$, i.e., if you are served asparagus, you ought to put your napkin on your lap. This is correct, however, it must be observed that to enable the inference it suffices to assume that there exist maximally ideal states where you are served asparagus or, in other words, it is not incompatible with ideality (it is permitted) that you are served asparagus. We argue this is a rather innocuous and intuitive assumption (and left unspecified in [22]). By $\mathbf{O}(n \mid \top)$ and $\mathbf{O}(f \mid a)$, this assumption allows you to infer that, if you learn that you are served asparagus, then it ought to be the case you put your napkin on your lap:

$$(\mathbf{O}\chi \wedge \mathbf{O}(\varphi \mid \psi) \wedge \neg\mathbf{O}\neg\psi) \to [!\psi]\mathbf{O}\chi$$

We hope this brief discussion further strengthens our claim that conditional obligations *à la* Hansson are a solid ground on which to base the investigations proposed in this paper and a versatile logical set up allowing natural and insightful interfaces with other logics such as the epistemic one dealt with in this section.

5.3   Information and priority

Up to now, our discussion of information dynamics was about conditional obligations at the base level of deontic ideality relations. But information dynamics also shows at the level of priority graphs. To conclude this section, we define an operation on P-graphs that matches the $[!\varphi]$ relational update modalities.

Definition 6 (P-graph restriction). *Let* $\mathcal{G} = \langle \Phi, \prec \rangle$ *be a P-graph, and* ψ *a formula. The restriction of* $\mathcal{G}$ *by* ψ *is the graph* $\mathcal{G}^\psi = \langle \Phi^\psi, \prec^\psi \rangle$ *where:*

$-\ \Phi^\psi = \{\varphi \wedge \psi \mid \varphi \in \Phi\}$;
$-\ \prec^\psi = \{(\varphi \wedge \psi, \varphi' \wedge \psi) \mid \varphi \prec \varphi'\}$.

The restriction of a P-graph $\mathcal{G}$ by ψ simply intersects the elements of the original graph with ψ and keeps the original order:

Theorem 2 (Harmony of P-graph restriction). *The following diagram commutes for all P-graphs* $\mathcal{G}$*, propositional formula* φ *and valuation* $\mathcal{I}$*:*

$$
\begin{array}{ccc}
\mathcal{G} & \longrightarrow & \mathcal{G}^\psi \\[2pt]
\scriptstyle\mathtt{sub}\big\downarrow & & \big\downarrow\scriptstyle\mathtt{sub} \\[2pt]
\langle S, \preceq_{\mathcal{G}}, \mathcal{I} \rangle & \xrightarrow{\ !\psi\ } & \langle \llbracket \psi \rrbracket, \preceq_{\mathcal{G}^\psi}, \mathcal{I}|_\psi \rangle
\end{array}
$$

Proof. Let $s \preceq_{\mathcal{G}} s'$ and $s, s' \in [\![\psi]\!]$. By Definition 2, for all $\varphi \in \Phi$ if $s \in [\![\varphi]\!]$ then $s' \in [\![\varphi]\!]$. Hence, if $s \in [\![\varphi \wedge \psi]\!]$ then $s' \in [\![\varphi \wedge \psi]\!]$, so $s \preceq_{\mathcal{G}^{\psi}} s'$. The other direction is similar.   □

In other words, the order $\preceq_{\mathcal{G}}$ obtained by update of φ is the same order obtained by graph restriction, and the information dynamics at our two levels lives in harmony. As a further illustration, here is an instance of Theorem 1 in this setting of information dynamics:

$$\mathcal{M}_{\mathcal{G}}|_{\psi}, s \models \mathbf{O}_{\preceq}(\varphi \mid \xi) \Longleftrightarrow \mathcal{M}_{\mathcal{G}^{\psi}}, s \models \mathbf{O}_{\mathcal{G}}(\varphi \mid \psi).$$

This relates conditional obligation based on a maximality statement about $\preceq_{\mathcal{G}}$ after an announcement ψ (i.e., in model $\mathcal{M}_{\mathcal{G}}|_{\psi}$) with the conditional obligation based on a subsumption statement $([\mathsf{U}](\xi(\psi) \to \varphi))$ about the restricted priority graph $\mathcal{G}|_{\psi}$.[17]

This gives us a first example of how the two-level perspective we have developed in Section 3 and applied to static scenarios in Section 4 has a natural dynamic extension. The next section will push this line further analyzing the sort of dynamics that arises from genuinely normative updates.

6   Third application: deontic dynamics proper

Deontically relevant events are of many kinds. Some of them are purely informational, as we have seen already. Others are about changes in evaluation of worlds, whether at base level, or at the level of priority structure. We will look at some examples of the latter kind, and then show how to model them at various levels that stay in harmony.

6.1   Generating priorities in concrete scenarios

To get some examples, we add a dynamic twist to our earlier treatment of CTDs. This time, consider how their normative component arises:

Example 4 (Postfixing norms). Classic deontic scenarios come with a normative priority order, but the latter has usually been created. For instance, start the earlier Gentle Murder scenario with the P-sequence

$$\langle \neg m \rangle$$

By Definition 2, this generates a total pre-order with all $\neg m$ states above all m states: "It is obligatory under the law that Smith not murder Jones". Suppose this is the given deontic situation. Now, a lawgiver, or a person with moral authority comes in, and introduces the sub-ideal obligation to kill gently: "it is

---

[17] A more radical view germane to this paper would make informational events syntactic in graphs, in the spirit of the evidence dynamics of [48].

obligatory that, if Smith murders Jones, Smith murders Jones gently"? This can be done by postfixing the original sequence with the property $\neg m \vee g$:

$$\langle \neg m, \neg m \vee g \rangle$$

to obtain the sequence encountered in Example 1.

Placing the new moral priority last leaves the original prohibition on murder intact. There may also be cases where we want to do the opposite.

Example 5 (Prefixing norms). Recall the P-sequence of the gentle murder case:

$$\langle \neg m, \neg m \vee g \rangle$$

Now we want to introduce a stronger norm than just "It is obligatory under the law that Smith not murder Jones", like "It is obligatory under the law that Smith not murder Jones and that Smith not be aggressive against Jones". Also we do not want to introduce any further change in the priority ordering. This can be achieved syntactically by prefixing the P-sequence with the property $\neg m \wedge \neg a$, where a stands for "Smith is aggressive against Jones":

$$\langle (\neg m \wedge \neg a), \neg m, \neg m \vee g \rangle$$

These example suggest the following definition of general pre- and post-fixing operations on P-graphs:

Definition 7 (Prefixing and postfixing in P-graphs). *Let $\mathcal{G} = \langle \Phi, \prec \rangle$ be a P-graph, and φ a propositional formula:*

- *the prefixing of $\mathcal{G}$ by φ yields the graph $\varphi; \mathcal{G}$ where a new maximal element $\varphi \wedge \bigwedge \max(\mathcal{G})$ is added to $\mathcal{G}$, consisting of the conjunction of φ with the conjunction of the maximal of $\mathcal{G}$;*
- *the postfixing of $\mathcal{G}$ by φ yields the graph $\mathcal{G}; \varphi$ where a new minimal element $\varphi \vee \bigvee \min(\mathcal{G})$ is added to $\mathcal{G}$, consisting of the disjunction of φ with the disjunction of the minimal elements of $\mathcal{G}$.*

These operations are illustrated in Figure 3. Observe that the constraints defining P-graphs—namely the logical dependencies of the elements of the graph—are respected by the definition, and that the above examples are special cases of the definition with $\varphi := g$ (Example 4) and, respectively, $\varphi := \neg a$ (Example 5).

6.2 Logic of base level betterness change

Modifying priority structure has concrete effects on betterness order of states induced in the manner of Section 2. These base-level changes can also be studied directly, by means of relation transformers in dynamic-epistemic style. Here is a typical example that has been used widely in the literature, from belief revision to learning theory. It may be viewed as a 'strong command' in favor of realizing some proposition φ:
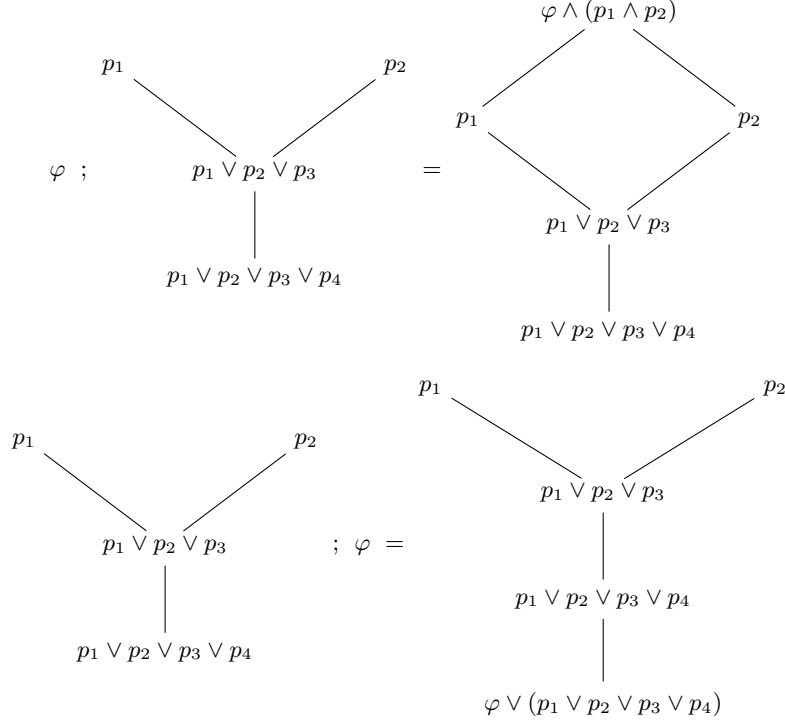
$$\varphi \wedge (p_1 \wedge p_2)$$

$$p_1 \qquad\qquad\qquad p_2$$

$$p_1 \qquad\qquad\qquad\qquad p_2$$

$$\varphi \;\; ; \qquad\qquad p_1 \vee p_2 \vee p_3 \qquad = \qquad p_1 \vee p_2 \vee p_3$$

$$p_1 \vee p_2 \vee p_3 \vee p_4 \qquad\qquad p_1 \vee p_2 \vee p_3 \vee p_4$$

$$p_1 \qquad\qquad\qquad\qquad p_2$$

$$p_1 \qquad\qquad\qquad p_2 \qquad\qquad p_1 \vee p_2 \vee p_3$$

$$p_1 \vee p_2 \vee p_3 \qquad ; \;\; \varphi \;=$$

$$p_1 \vee p_2 \vee p_3 \vee p_4$$

$$p_1 \vee p_2 \vee p_3 \vee p_4$$

$$\varphi \vee (p_1 \vee p_2 \vee p_3 \vee p_4)$$

Fig. 3. Hasse diagrams illustrating the pre- and post-fixing by φ of the P-graph on the right-hand side of Figure 1.

Definition 8 (Radical upgrade). *Let $\mathcal{M} = \langle S, \preceq, \mathcal{I} \rangle$ be a model and φ be a propositional formula. A radical upgrade $\Uparrow \varphi$ yields model $\mathcal{M}_{\Uparrow\varphi} = \langle S, \preceq_{\Uparrow\varphi}, \mathcal{I} \rangle$ where $\preceq^{\Uparrow\varphi}$ is equal to:*

$$\{(s,t) \in \preceq \mid s, t \in [\![\varphi]\!]\} \cup \{(s,t) \in \preceq \mid s, t \notin [\![\varphi]\!]\} \cup \{(s,t) \in S^2 \mid s \notin [\![\varphi]\!] \text{ and } t \in [\![\varphi]\!]\}$$

In other words, a radical upgrade $\Uparrow \varphi$ changes the current order $\preceq$ to that of a new model $\mathcal{M}_{\Uparrow\varphi}$ where all φ-states become better than all $\neg\varphi$-states, while, within those two zones, the old ordering remains.[18]

It is easy to see that radical upgrade preserves pre-orders and is thus well-defined for our semantics. Many other operators are definable in this format, including weaker commands or just 'suggestions': cf. [54]. This variety in normative force is deontically relevant, since agents are exposed to many normative cues, from direct orders to mere hints. It will not be our main topic here, however, and radical upgrade will do for making most of our points.

---

[18] A more compact formulation of $\preceq^{\Uparrow\varphi}$ can be given by the following regular expression:

$$\preceq^{\Uparrow\varphi} = (?\varphi; \preceq; ?\varphi) \cup (?\neg\varphi; \preceq; ?\neg\varphi) \cup (?\neg\varphi; S^2; ?\varphi).$$

These operations do not change the domains of models, like the earlier public announcements of hard information: they change normative order. Still, reasoning with such actions falls within the scope of dynamic-epistemic logic, and [29] shows how they support reduction axioms for absolute and conditional obligations. We merely display an example for changing order below and the axiom for obligation can be derived from it (cf. Formula (14)):

$$[\Uparrow \varphi][\preceq]\psi \leftrightarrow (\varphi \wedge [\preceq](\varphi \to [\Uparrow \varphi]\psi))$$
$$\vee(\neg\varphi \wedge [\preceq](\neg\varphi \to [\Uparrow \varphi]\psi)) \wedge [\mathsf{U}](\varphi \to [\Uparrow]\psi) \tag{23}$$

As with public announcement logics, this allows for much finer analysis of deontic reasoning, now looking also at how one should act in the presence of a much richer variety of normative events that can change one's obligations.

6.3   Harmony

We have seen that deontic dynamics can be located both at the level of P-graphs and at the level of their underlying states. Both represent natural ways of thinking about deontic changes. What is the relation between the two levels? Like for the static and information dynamics cases, we obtain an harmony theorem:

Theorem 3 (Harmony of P-graph pre-\post-fixing). *The following diagram commutes for all P-graphs* $\mathcal{G}$, *propositional formulae* φ *and valuations* $\mathcal{I}$:

$$
\begin{array}{ccc}
\mathcal{G} & \xrightarrow{\ \star\varphi\ } & \mathcal{G} \star \varphi \\[2pt]
{\scriptstyle\texttt{sub}}\Big\downarrow & & \Big\downarrow{\scriptstyle\texttt{sub}} \\[2pt]
\langle S, \preceq_{\mathcal{G}}, \mathcal{I}\rangle & \xrightarrow{\ \Uparrow f_\star(\varphi)\ } & \langle S, \preceq_{\mathcal{G}\star\varphi}, \mathcal{I}\rangle
\end{array}
$$

where $\mathcal{G} \star \varphi$ *denotes either the pre-fixing* $\varphi;\mathcal{G}$ *or the post-fixing* $\mathcal{G};\varphi$ *of* $\mathcal{G}$ *by* φ *and* $f_\star(\varphi)$ *denotes accordingly* $\varphi \wedge \bigwedge \max(\mathcal{G})$ *or* $\varphi \vee \bigvee \min(\mathcal{G})$.

Proof (Sketch). Consider the case of post-fixing $\mathcal{G};\varphi$. We have two cases: 1) $[\![\varphi \vee \bigvee \min(\mathcal{G})]\!] = [\![\bigvee \min(\mathcal{G})]\!]$, in which case clearly $\preceq_{\mathcal{G}}=\preceq_{\mathcal{G}\star\varphi}$ and $= \preceq_{\mathcal{G}\star\varphi}=\preceq_{\Uparrow\varphi}$. 2) $[\![\varphi \vee \bigvee \min(\mathcal{G})]\!] \supset [\![\bigvee \min(\mathcal{G})]\!]$. By Definitions 2 and 7, the total pre-order $\preceq_{\mathcal{G}\star\varphi}$ consists therefore of the same equivalence classes of $\preceq_{\mathcal{G}}$ plus two classes splitting the states that do not satisfy any of the worse properties in $\mathcal{G}$: one consisting of $\varphi \wedge \neg \bigvee \min(\mathcal{G})$-states and one (the bottom one) of $\neg\varphi \vee \bigvee \min(\mathcal{G})$-states. Now consider $\preceq_{\Uparrow\varphi}$. Since $\varphi \vee \bigvee \min(\mathcal{G})$ is weaker than all the properties in $\mathcal{G}$, it will not affect the ordering $\preceq_{\mathcal{G}}$ except for splitting in two equivalence classes the class of worst elements (i.e., the set of $\neg \bigvee \min(\mathcal{G})$-states). These two classes are: one, the $\neg \bigvee \min(\mathcal{G})$-states which satisfy φ; and two the $\neg \bigvee \min(\mathcal{G})$-states which do not satisfy φ. The case for pre-fixing is similar.   □

Thus, we have seen how both informational and normative events can be represented in our priority framework, while also showing how the effects of these are in harmony with a natural base-level dynamics of betterness change on modal state models. As we have seen in a few examples, this offers a much richer view of what deontic scenarios actually involve, and how their normative structures are constructed and modified.

6.4   Case study: strong permission and CTDs

We conclude with one more concrete example of how a dynamic priority setting suggests new takes on old problems in deontic logic. The distinction between weak and strong permission dates back to [53]:

"An act will be said to be permitted in the weak sense if it is not forbidden; and it will be said to be permitted in the strong sense if it is not forbidden but subject to norm. [...] Weak permission is not an independent norm-character. Weak permissions are not prescriptions or norms at all. Strong permission only is a norm-character." [53, p. 86]

Thus, the weak permission "it is permitted that φ" amounts to mere absence of the prohibition "it is forbidden that φ", which is definable as $\neg\mathbf{O}\neg\varphi$. But—and that is the quote's claim—this is not the case for strong permission. How can we do justice to this distinction?[19] Defining strong permission is listed in [19] as one of ten major 'philosophical problems' in deontic logic.

Strong permission does not seem to be one single deontic act. On the just noted epistemic analogy, it could be viewed as opening up a new relevant possibility that had not been considered previously. It may also be viewed as allowing *actions* rather than endorsing possibilities, making comparisons with propositional "may" less immediate. We have nothing to say about these senses, even though they look congenial to a dynamic perspective. But the Von Wright quote in terms of "subject to norm" also highlights another sense of strong permission as related to normative behavior, more in line with legal theory [39, p. 120]:

"Telling me what I am permitted to do provides no guide to conduct unless the permission is taken as an exception to a norm of obligation [...]. Norms of permission have the normative function only of indicating, within some system, what are the exceptions from the norms of the obligation of the system."

Viewing strong permission as exception-making raises the issue what "exception" means. On a radical view, it restricts the action of some earlier norm, on a less radical view, it introduces a new subnorm:

---

[19] Strong permission does not stand alone, and it has analogues in other areas, such as strong epistemic "may" highlighting an epistemic possibility. We will not go into all natural language meanings of this term.

Example 6 (Killing in self-defense). Let us start again with the gentle murder scenario, slightly rephrased:

It is obligatory that Smith not kill Jones.

Like in the examples at the beginning of this section, we want to change this norm, now by a permission stating that:

Smith is allowed to kill Jones, provided he does that in self-defense.

We abbreviate "killing" with k and "killing in self-defense" with d.

There are two ways of interpreting the effects of such a strong permission:

1. The permission gives up the validity of $\mathbf{O}(\neg k \mid \top)$ for that of the weaker obligation $\mathbf{O}(k \to d \mid \top)$. Syntactically, this removes $\neg k$ from the P-sequence and substitutes $k \to d$. In this case, a strong permission repeals an earlier obligation, and then introduces a weaker one widening the field of permissibility [21]. This is close to an act of "derogation" in the law:
   "The difference between weak and strong permission becomes clear when thinking about the function of permissive norms. [...] A permissive norm is necessary when we have to repeal a preceding imperative norm or to derogate to it. That is to abolish a part of it [...]." [7, p. 891-892][20]
2. The strong permission does not modify $\mathbf{O}(\neg k \mid \top)$, but introduces a CTD stating that, in case Smith kills Jones, he should do that in self-defense: $\mathbf{O}(d \mid k)$. Example 7 will show how this can be analyzed as an instance of Theorem 3.

Both options have been object of investigations in the deontic logic literature (e.g., [11, 5] for the first option and [32, 40] for the second). Of importance here is that both options can be dealt with in the sort of dynamic logics of preference change we have introduced in this section.

Example 7 (Two-level dynamics of strong permission). Following Example 6 consider a P-sequence $\langle \neg k \rangle$, whose derived model $\mathcal{M}$ validates $\mathbf{O}\neg k$. Enacting a strong permission to kill if acting in self-defense can be modeled by introducing a modified CTD that makes killing satisfying some specified condition (here, self-defense) better than killing when those conditions are violated: $\neg k \vee (k \wedge d)$, i.e., $k \to d$. What we obtain then is an instance of the earlier harmony theorem in its post-fixing format, for any valuation $\mathcal{I}$:

$$
\begin{array}{ccc}
\langle \neg k \rangle & \xrightarrow{\ \star\varphi\ } & \langle \neg k \rangle; \langle k \to d \rangle \\
\Big\downarrow {\scriptstyle \mathtt{sub}} & & \Big\downarrow {\scriptstyle \mathtt{sub}} \\
\langle S, \preceq_{\langle \neg k \rangle}, \mathcal{I} \rangle & \xrightarrow{\ \Uparrow f_\star(\varphi)\ } & \langle S, \preceq_{\langle \neg k \rangle;\langle k \to d \rangle}, \mathcal{I} \rangle
\end{array}
$$

---

[20] Quoted in the paper on permission and obligation [8].

In other words, strong permission in our exception sense may be modeled by inserting a predicate in a P-graph at some specified position. Through the earlier dynamic logics, this will induce a logic of strong permission that can be checked against prior inferential intuitions — though we will not pursue this line of assessment here.

Our treatment by no means exhausts all deontic views on strong permission[21], but it does emphasize a link with CTDs and their dynamics which, to the best of our knowledge, had not yet been touched upon. Our claim here is that the process of incremental specification of a CTD sequence via betterness update, be it syntactic or semantic, can legitimately be viewed as the enactment of strong permissions. Such permissions are refinements of existing obligations or, to say it otherwise, exceptions to existing obligations that do not reject such obligations altogether, but rather specify conditions under which a violation of such obligations is tolerable. Notice also that, against the intuitions that might be dictated by the natural language formulation of the notion, this interpretation of strong permission has actually more to do with **O**-statements of obligation (allbeit of a CTD type) rather than with $\neg\mathbf{O}\neg$-statements of (weak) permission.

6.5   Generalizing deontic dynamics: graph composition

We conclude this section with an observation which, we believe, should set the stage for the future development of the set of tools presented in this section. We have studied how formulae can update P-graphs and their associated pre-orders. However, a formula φ is nothing but a special P-graph $\mathcal{G}' = \langle\varphi\rangle$ consisting of one single property. So what we have seen up till now is how to update a given graph $\mathcal{G}$ by the special graph $\langle\varphi\rangle$. The natural question arises then of how two graphs $\mathcal{G}$ and $\mathcal{G}'$ can be combined in general and which pre-orders do they induce.

This is not just of technical interest, but it is interesting from the point of view of deontic logic as an abstract setting from which to study how norms—viewed as P-graphs—can be combined to form new norms. Reverting to our running example, here is a concrete illustration of what we have in mind.

Example 8 (Quick murder). Let us assume now there are two normative sources. According to the first one:

1. It is obligatory under the law that Smith not murder Jones.
2. It is obligatory that, if Smith murders Jones, Smith murders Jones gently."

According to the second one:

1. It is obligatory under the law that Smith not murder Jones.
2. It is obligatory that, if Smith murders Jones, Smith murders Jones *quickly.*"

We want to merge the two CTDs and obtain a P-graph which contains the normative information of both components. Assuming the alphabet $\{m, g, q\}$

---

[21] For instance, [12, 9, 32] pursue a view where strong permissions set boundaries to any prohibitions that normative authorities could possibly enact.

with the obvious intuitive interpretation, we can model this scenario by means
of two P-sequences $\mathcal{S}_g$ with $\neg m \succ \neg m \vee (m \wedge g)(= m \rightarrow g)$, and $\mathcal{S}_q$ with
$\neg m \succ \neg m \vee (m \wedge q)(= m \rightarrow q)$. Formally, the graph operation we have in
mind consisting of taking the disjoint union of $\mathcal{S}_g$ and $\mathcal{S}_q$—also called *parallel
composition* [29], which we denote by symbol "$||$". The resulting pre-order should
be such that the best states are $\neg m$-states and the sub-ideal states are split into
two incomparable classes, the class of $m \wedge g$-states and the class of $m \wedge q$-states
(see Figure 4). This is nothing but the intersection of the betterness (total)
pre-orders of the two P-sequences.

The example instantiates a general result, from [2], relating the parallel composi-
tion of P-graphs as disjoint union and the intersection of their derived pre-orders:

Fact 3 (Harmony of parallel composition of P-graphs) *Let $\mathcal{G} = \langle G, \prec \rangle$
and $\mathcal{G}' = \langle G', \prec' \rangle$ be two P-graphs. The following diagram commutes:*

$$
\begin{array}{ccc}
\mathcal{G} & \xrightarrow{\ ||\mathcal{G}'\ } & \mathcal{G}||\mathcal{G}' \\
{\scriptstyle\mathtt{sub}}\Big\downarrow & & \Big\downarrow{\scriptstyle\mathtt{sub}} \\
\langle S, \preceq_{\mathcal{G}} \rangle & \xrightarrow{\ \cap\ \ \preceq_{\mathcal{G}'}\ } & \langle S, \preceq_{\mathcal{G}||\mathcal{G}'} \rangle
\end{array}
$$

Proof. The proof is given by the following equivalences obtained by iterated
application of Definition 2, for any valuation:

$$
s \preceq_{\mathcal{G}||\mathcal{G}'} s' \iff \forall \varphi \in G \cup G' : s \in [\![\varphi]\!] \Rightarrow s' \in [\![\varphi]\!]
$$
$$
\iff \forall \varphi \in G : s \in [\![\varphi]\!] \Rightarrow s' \in [\![\varphi]\!] \text{ and } \forall \varphi \in G' : s \in [\![\varphi]\!] \Rightarrow s' \in [\![\varphi]\!]
$$
$$
\iff s \preceq_{\mathcal{G}} s' \text{ and } s \preceq_{\mathcal{G}'} s'.
$$

This completes the proof.   □

Similarly, the operation generalizing our earlier insertions of new predicates
in priority sequences is the operation of sequential composition of two priority
graphs $(\mathcal{G}; \mathcal{G}')$. Defining this is a bit more tricky in our setting. To generalize
Theorem 3, we must modify the two predicate inclusion orders in the right way.
Here, for convenience, we will follow another road, taking the original sequential
composition of graphs as defined in [2], putting all nodes in $\mathcal{G}'$ behind all those
of $\mathcal{G}$ in the priority order (see also our earlier Formula (4)). [22]

At the level of betterness relations, [2] shows that the relation $\preceq_{\mathcal{G};\mathcal{G}'}$ is defin-
able as a priority composition : $\preceq_{\mathcal{G}} * \preceq'_{\mathcal{G}}$:

$$
s \preceq_{\mathcal{G};\mathcal{G}'} s' \iff (s \preceq_{\mathcal{G}} s' \cap s \preceq_{\mathcal{G}'} s') \cup s \prec_{\mathcal{G}} s'.
$$

These observations make the following graph change diagram commute:

---

[22] The techniques of our Appendix allow us to effectively transform $\mathcal{G}; \mathcal{G}'$ defined in
this way to our earlier inclusion format.

$$\neg m \qquad \qquad \neg m \qquad\qquad\qquad \neg m \qquad \neg m$$

$$\Big| \qquad || \qquad \Big| \qquad = \qquad \Big| \qquad \Big|$$

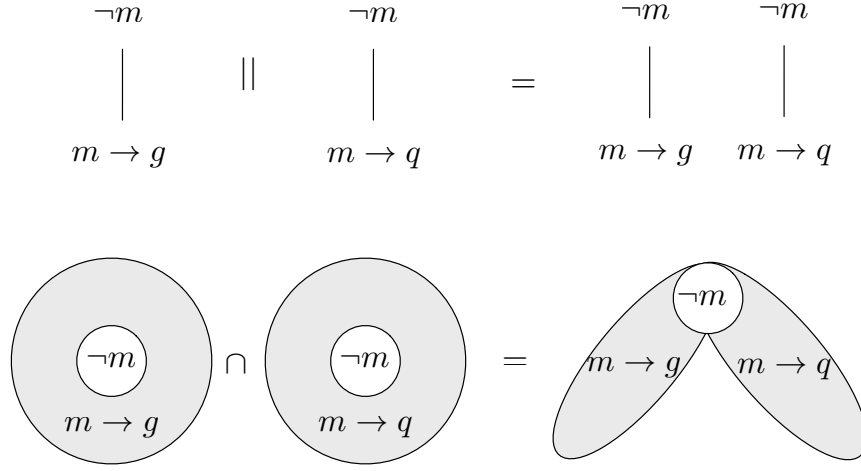$$m \rightarrow g \qquad\qquad m \rightarrow q \qquad\qquad\qquad m \rightarrow g \quad m \rightarrow q$$

Fig. 4. Merging of P-sequences in Example 8. The merging is depicted both in the form of parallel composition of P-sequences (top) and as intersection of total pre-orders in the form of spheres (bottom).

Fact 4 (Harmony of sequential composition of P-graphs) *Let $\mathcal{G} = \langle G, \prec \rangle$ and $\mathcal{G}' = \langle G', \prec' \rangle$ be two P-graphs. The following diagram commutes:*

$$
\begin{array}{ccc}
\mathcal{G} & \xrightarrow{\ ;\mathcal{G}'\ } & \mathcal{G};\mathcal{G}' \\
{\scriptstyle\texttt{lex}}\Big\downarrow & & \Big\downarrow{\scriptstyle\texttt{lex}} \\
\langle S, \preceq_{\mathcal{G}} \rangle & \xrightarrow{\ *\preceq'_{\mathcal{G}}\ } & \langle S, \preceq_{\mathcal{G};\mathcal{G}'} \rangle
\end{array}
$$

This fact explains and generalizes all our earlier results about pre-fixing and post-fixing single propositions to normative structures.

Up to now, we have seen two major composition operations on priority graphs behind the specific examples discussed in this paper. [2] shows why this is a felicitous mathematical choice when analyzing social choice and related phenomena. However, our observations are clearly just the beginning of a more general account of operations that construct and modify the structure of normative criteria.

7   Conclusion

In this paper, we have shown how deontic scenarios can be mined for more structure than just deontic inferences. Equally crucial is normative structure, represented in priority graphs, and the dynamics of informational and deontic events

that change our current obligations. We have shown how this view can be implemented by merging ideas from graph representations of criteria for preference with dynamic epistemic logics of informational events. The result is a framework for representing obligations that fits with current trends in other areas of logic of agency, while also throwing fresh light on old issues in the literature. Moreover, adopting our framework brings many relevant new phenomena into the scope of deontic logic, such as norm change and general calculus of normative code.

While our examples will have shown the flavor of the style of analysis that we have proposed, many topics remain for further research. We list a few here:

- Analyzing the rich linguistic repertoire of commands and suggestions uttered by agents with different deontic roles, thus connecting to *speech act theories*,. (cf. [54].)
- Extending our analysis from obligations in terms of propositions to obligations among *actions* as the primary carriers of moral qualifications. (see [13] on action vs. deontology.) Related to this, we would like to extend our analysis from the *deliberative stance* in this paper, based on handling propositions about a fixed current world that we get to know better, to the *action stance* of a changing world where agents have to choose actions, which is closer to first-person experienced agency. (cf. [33] about deontic transitions)
- Enriching the information dynamics in our treatment to beliefs and plausibility orders. Priority order in this case becomes like 'entrenchment', and one could have dual priority graphs for plausibility and betterness. One can then extend our analysis to games as a richer setting of preference entangled with belief, and strategic interaction. (cf. [26, 41] and [18].)
- Adding social agents and their obligations, with the appropriate priority graphs for social knowledge and action, linking to the original motivation of priority graphs in [16] as models for social choice. We can further connect our dynamics of local informational or deontic events to long term deontic phenomena in agency over time. (cf. [14, 45] on temporal 'protocols', and [23] on deontic STIT logic.)
- Exploring detailed legal argumentation as a natural test for the richer deontic modeling apparatus proposed here and developing our graph calculus to deal with a richer repertoire of natural operations of norm merge and construction of moral codes.

These extensions would tie deontic logic firmly to current studies of agency and games, increasing its range and impact.

References

1. A.R. Anderson. The formal analysis of normative concepts. *American Sociological Review*, 22:9–17, 1957.
2. H. Andreka, M. Ryan, and P. Schobbens. Operators and laws for combining preference relations. *Journal of Logic and Computation*, 12:13–53, 2002.
3. L. Åqvist. Good samaritans, contrary-to-duty imperatives and epistemic obligations. *Nous*, 1:361–379, 1967.

4. L. Åqvist. Systematic frame constants in defeasible deontic logic. In D. Nute, editor, *Defeasible Deontic Logic*, Synthese Library, pages 59–78. Kluwer Academic Publishers, 1997.

5. G. Aucher, D. Grossi, A. Herzig, and E. Lorini. Dynamic context logic. In X. He, J. Horty, and E. Pacuit, editors, *Proceedings of LORI 2009*, volume 5834 of *LNAI*. Springer, 2009.

6. J. van Benthem and F. Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logic*, 17:157–182, 2007.

7. N. Bobbio. Norma. In *Enciclopedia Einaudi*, volume 9, pages 876–907. Einaudi, Torino, Italy, 1980.

8. G. Boella and L. van der Torre. Permissions and obligations in hierarchical normative systems. In *Proceedings of ICAIL'03*, pages 109–118, New York, 2003. ACM Press.

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

10. C. Boutilier. Conditional logics of normality: A modal approach. *Artificial Intelligence*, 68:87–154, 1994.

11. M. Brown. Conditional obligation and positive permission for agents in time. *Nordic Journal of Philosophical Logic*, 5(2):83–112, 2000.

12. E. Bulygin. Permissive norms and normative systems. In A. Martino and F. Socci Natali, editors, *Automated analysis of legal texts*, pages 211–218. Publishing company, Amsterdam, 1986.

13. J. Czelakowski. Action and deontology. In E. Ejerhed and S.Lindstrom, editors, *Logic, Action and Cognition-Essays in Philosophical Logic*, pages 47–87. Kluwer Academic Publishers, 1997.

14. R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge*. Cambridge, MA: The MIT Press, 1995.

15. J. Forrester. Gentle murder, or the adverbial samaritan. *Journal of Philosophy*, 81:193–197, 1984.

16. P. Girard. *Modal Logic for Belief and Preference Change*. PhD thesis, University of Amsterdam, 2008.

17. G. Governatori and A. Rotolo. Logic of violations: A gentzen system for reasoning with contrary-to-duty obligations. *Australasian Journal of Logic*, 3, 2005.

18. D. Grossi and P. Turrini. Short-sight in extensive games. In V. Conitzer and M. Winikoff, editors, *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. IFAAMAS, 2012.

19. J. Hansen, G. Pigozzi, and L. van der Torre. Ten philosophical problems in deontic logic. In Guido Boella, Leon van der Torre, and Harko Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in DROPS Proceedings. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, 2007.

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

21. R. Hilpinen. On normative change. In E. Morscher and R. Stranzinger, editors, *Ethics. Foundations, Problems, and Applications. Proceedings of The 5th International Wittgenstein Symposium*, pages 155–164. Hölder-Pichler-Tempsky, 1980.

22. J. Horty. Nonmonotonic foundations of deontic logic. In D. Nute, editor, *Defeasible Deontic Logic*, pages 17–44. Kluwer, 1997.

23. J. Horty. *Agency and Deontic Logic*. Oxford University Press, 2001.

24. A. J. I. Jones and I. Pörn. Ideality, sub-ideality and deontic logic. *Synthese*, pages 275–290, 1985.

25. S. Kanger. New fondations for ethical theory. In R. Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 36–58. Reidel Publishing Company, 1971.

26. B. Kooi and A. Tamminga. Conditional obligations in strategic situations. In M. Singh G. Boella, G Pigozzi and H. Verhagen, editors, *3rd International Workshop on Normative Multiagent Systems*, pages 188–200. Kluwer Academic Publishers, 2008.

27. D. Lewis. Semantic analysis for dyadic deontic logic. In S. Stenlund, editor, *Logical Theory and Semantical Analysis*, pages 1–14. Reidel Publishing Company, 1974.

28. F. Liu. *Changing for the Better. Preferemce dynamics and Agent Diversity*. PhD thesis, University of Amsterdam, 2008.

29. F. Liu. *Reasoning about Preference Dynamics*, volume 354 of *Synthese Library*. Springer, 2011.

30. F. Liu. A two-level perspective on preference. *Journal of Philosophical Logic*, 40(3):421–439, 2011.

31. D. Makinson. Five faces of minimality. *Studia Logica*, 52:339–379, 1993.

32. D. Makinson and L. van der Torre. Permission from an input/output perspective. *Journal of Philosophical Logic*, 32:391–416, 2003.

33. R. van der Meyden. The dynamic logic of permission. *Journal of Logic and Computation*, 6:465–479, 1996.

34. G. E. Moore. *Principia Ethica*. Cambridge University Press, 1903.

35. A. Nayak, P. Nelson, and H. Polansky. Belief change as change in epistemic entrenchment. *Synthese*, 109:143–174, 1996.

36. E. Pacuit, R. Parikh, and E. Cogan. The logic of knowledge based on obligation. *Synthese*, 149:311–341, 2006.

37. H. Prakken and M. Sergot. Contrary-to-duty obligations. *Studia Logica*, 57:91–115, 1996.

38. H. Prakken and M. Sergot. Dyadic deontic logic and contrary-to-duty obligations. In D. Nute, editor, *Defeasible Deontic Logic*, volume 263 of *Synthese Library Series*. Kluwer Academic Publishers, 1997.

39. A. Ross. *Directives and Norms*. Routledge & Kegan Paul, London, 1968.

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

41. X. Sun and F. Liu. Consequentialist deontic logic for decisions and games. under submission, 2012.

42. Y.-H. Tan and L. van der Torre. How to combine ordering and minimizing in a deontic logic based on preferences. In M. Brown and J. Carmo, editors, *Deontic Logic, Agency and Normative Systems, DEON '96: Third International Workshop on Deontic Logic in Computer Science*, pages 216–232, 1996.

43. J. van Benthem. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logic*, 14(2), 2004.

44. J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011.

45. J. van Benthem, J. Gerbrandy, T. Hoshi, and E. Pacuit. Merging frameworks for interaction. *Journal of Philosophical Logic*, 38(5):491–526, 2009.

46. J. van Benthem, P. Girard, and O. Roy. Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, 38:83–125, 2009.

47. J. van Benthem and D. Grossi. Normal forms for priority graphs. Technical Report PP-2011-02, Institute for Logic, Language and Computation, 2011.

48. J. van Benthem and E. Pacuit. Dynamic logics of evidence-based beliefs. *Studia Logica*, 99(1):61–92, 2011.

49. L. van der Torre. *Reasoning about obligations: defeasibility in preference-based deontic logic*. PhD thesis, Erasmus University Rotterdam, 1997.
50. L. van der Torre and Y.-H. Tan. The many faces of defeasibility in defeasible deontic logic: Essays in nonmonotonic normative reasoning. In D. Nute, editor, *Defeasible Deontic Logic*, volume 263 of *Synthese Library*, pages 79–121. Kluwer Academic Publishers, 1997.
51. J. van Eck. A system of temporally relative modal and deontic predicate logic and its philosophical applications. *Logique et Analyse*, 99:249—290, 1982.
52. B. van Fraassen. Values and the heart's command. *The Journal of Philosophy*, 70(1):5–19, 1973.
53. G. H. Von Wright. *Norm and Action. A Logical Inquiry*. Routledge, London, 1963.
54. T. Yamada. Logical dynamics of some speech acts that affect obligations and preferences. *Synthese*, 165(2):295–315, 2008.

A   Equivalence of P-graph formats

This appendix is devoted to establish the equivalence of two different formats of P-graphs, the one used throughout the paper, and the one we alluded to in Remark 1 and which, in a special limited form, had been proposed in [4] (see also footnote 4).

Definition 9 has introduced P-graphs where elements are ordered according to logical entailment. In this appendix, which recapitulates results from [47], we show how each P-graph (even without the logical entailment constraint) can be put in two syntactically different, but semantically equivalent formats.

Definition 9 (General P-graphs). *Let $\mathcal{L}(\mathbf{P})$ be a propositional language built on the set of atoms $\mathbf{P}$. A P-graph is a tuple $\mathcal{G} = \langle \Phi, \prec \rangle$ such that:*

- $\Phi \subset \mathcal{L}(\mathbf{P})$ *with* $|\Phi| < \omega$;
- $\prec$ *is a strict order on Φ.*

Given a P-graph $\mathcal{G}$, we denote with $\uparrow_{\mathcal{G}} \varphi = \{\psi \in \Phi \mid \varphi \prec \psi$ or $\psi = \varphi\}$ the *upset* of φ in $\mathcal{G}$. We also denote with $\uparrow_{\mathcal{G}}^+ \varphi = \{\psi \in \Phi \mid \varphi \prec \psi\}$ the *strict upset* of φ in $\mathcal{G}$. When clear from the context, we will drop the index and write simply $\uparrow \varphi$ and $\uparrow^+ \varphi$.

A.1   Exclusive normal form

We introduce the first type of normal form:

Definition 10 (Exclusive normal form for P-graphs). *Let $\mathcal{G} = \langle \Phi, \prec \rangle$ be a P-graph. The exclusive normal form of $\mathcal{G}$ is a P-graph $\mathcal{G}_{ex} = \langle \Phi_{ex}, \prec_{ex} \rangle$ such that:*

- $\Phi_{ex} = 2^{\Phi}$. *Each element $\Psi \in \Phi_{ex}$ has to be read as a finite conjunction:*

$$\bigwedge \Psi \wedge \bigwedge \neg(\Phi - \Psi)$$

that is, the conjunction of all properties in Ψ and all the negations of the properties not in Ψ.
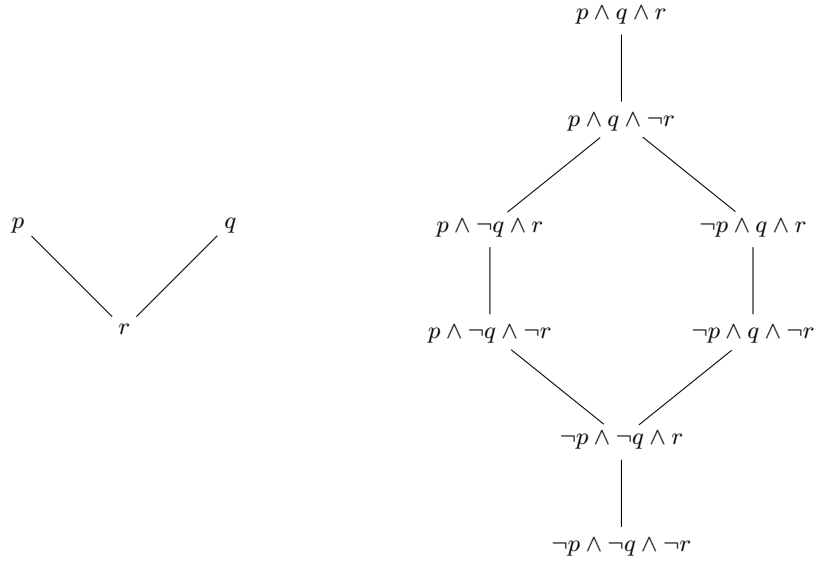
Fig. 5. Hasse diagrams of a general P-graph (left) and its exclusive normal form (right)

- *The relation $\prec_{ex}$ is defined as follows:*

$$\Psi \prec_{ex} \Psi' \Longleftrightarrow \exists \varphi \in \Phi : [\varphi \in \Psi' \text{ and } \varphi \notin \Psi$$
$$\text{and } \forall \varphi' : [\varphi' \notin \Psi' \text{ and } \varphi \in \Psi \Longrightarrow \varphi' \prec \varphi]].$$

An example of the exclusive normal form of a graph is given in Figure 5. It should be clear by the construction of the exclusive normal forms that these graphs consist of logically disjoint elements. Notice also that the number of elements in these graphs is bound by the cardinality of Φ being equal to $2^{|\Phi|}$.

Recall the lexicographic order-derivation rule `lex` introduced in Remark 1. We can now prove a simple normal form theorem guaranteeing that every P-graph has an equivalent exclusive normal form.

Lemma 1 (Adequacy of exclusive normal forms). *For each graph $\mathcal{G}$ and valuation $\mathcal{I}$:* $\preceq_{\mathcal{G}}^{\texttt{lex}} = \preceq_{\mathcal{G}_{ex}}^{\texttt{lex}}$.

Proof. We prove the claim by showing that, for any valuation $\mathcal{I}$:

(i) $|S|_{\preceq_{\mathcal{G}}^{\texttt{lex}}} = |S|_{\preceq_{\mathcal{G}_{ex}}^{\texttt{lex}}}$, i.e., the two preorders give rise to the same equivalence classes;

(ii) $\prec_{\mathcal{G}}^{\texttt{lex}} = \prec_{\mathcal{G}_{ex}}^{\texttt{lex}}$, i.e., the strict part of the two preorders is the same.

As to (i) it suffices to observe that all equivalence classes $|s|_{\preceq_{\mathcal{G}}^{\texttt{lex}}}$ yielded by $\preceq_{\mathcal{G}}^{\texttt{lex}}$ must be the truth-sets—under $\mathcal{I}$—of some conjunctions of length $|\Phi|$ of formulae of the form $\bigwedge \Psi \wedge \bigwedge \neg(\Phi \setminus \Psi)$ with $\Psi \subseteq \Phi$. But these are precisely the disjoint

Fig. 6. Hasse diagrams of a P-graph with logically disjoint elements (left) and its inclusive normal form (right)

elements of the normal form. As to (ii), it is proven by the following series of equivalences:

$$s \prec_{\mathcal{G}}^{\mathtt{lex}} s' \iff \exists \varphi \in \Phi : s \notin \llbracket \varphi \rrbracket \text{ and } s' \in \llbracket \varphi \rrbracket \text{ and } \forall \varphi' : [s \in \llbracket \varphi' \rrbracket \text{ and } s' \notin \llbracket \varphi \rrbracket \implies \varphi' \prec \varphi]$$
$$\iff \forall \Psi, \Psi' \in \Psi_{ex} : [\text{ if } \llbracket \Psi \rrbracket = |s|_{\sim_{\mathcal{G}}^{\mathtt{lex}}} \text{ and } \llbracket \Psi' \rrbracket = |s'|_{\sim_{\mathcal{G}}^{\mathtt{lex}}} \text{ then } \Psi \prec_{ex} \Psi']$$
$$\iff \exists \Xi \in \Phi_{ex} : [s \notin \llbracket \Xi \rrbracket \text{ and } s' \in \llbracket \Xi \rrbracket \text{ and } \forall \Xi' : [s \in \llbracket \Xi' \rrbracket \text{ and } s' \notin \llbracket \Xi' \rrbracket \implies \Xi' \prec_{ex} \Xi]]$$
$$\iff s \prec_{\mathcal{G}_{ex}}^{\mathtt{lex}} s'$$

where, to simplify notation, by Ψ, Ξ we denote the finite conjunction of the elements of Ψ, Ξ and of the negations of the members of its complement. The first equivalence holds by the definition of **lex**. The second one holds by the definition of exclusive normal form and the fact that equivalence classes are definable as Boolean compounds of elements of the graph. The third one holds by the fact that those compounds are logically disjoint.

A.2   Inclusive normal form

Definition 11 (Inclusive normal forms). *Let $\mathcal{G} = \langle \Phi, \prec \rangle$ be a P-graph with logically disjoint elements. The inclusive normal form of $\mathcal{G}$ is a P-graph $\mathcal{G}_{in} = \langle \Phi_{in}, \prec_{in} \rangle$ such that:*

- *$\Phi_{in} = \uparrow \mathcal{G}$, that is, the set of upsets of $\mathcal{G}$. To simplify notation, each element $\Psi \in \Phi_{in}$ will be read also as the finite disjunction: $\bigvee \Psi$.*
- *The relation $\prec_{in}$ is defined as follows:*

$$\Psi \prec_{in} \Psi' \iff \Psi' \subset \Psi.$$

An example of inclusive normal form is given in Figure 6. Notice that this form is defined only for graphs with logically disjoint elements.

Like for the case of exclusive normal forms, we obtain a lemma guaranteeing that every exclusive P-graph has an equivalent inclusive normal form.

Lemma 2 (Adequacy of inclusive normal forms). *For each graph $\mathcal{G}$ and valuation $\mathcal{I}$: $\preceq_{\mathcal{G}}^{\mathtt{lex}} = \preceq_{\mathcal{G}_{in}}^{\mathtt{sub}}$.*

Proof. We prove the claim by the following series of equivalences:

$$s \preceq_{\mathcal{G}}^{\text{lex}} s' \iff \exists \varphi \in \Phi : (s \notin [\![\varphi]\!] \text{ and } s' \in [\![\varphi]\!] \text{ and } \forall \varphi' : (s \in [\![\varphi']\!] \text{ and } s' \notin [\![\varphi']\!] \implies \varphi' \prec \varphi)$$

$$\iff \forall \Psi, \Psi' \in \Phi_{in} : [\text{ if } [\![\Psi]\!] = \uparrow |s|_{\sim_{\mathcal{G}}^{\text{lex}}} \text{ and } [\![\Psi']\!] = \uparrow |s'|_{\sim_{\mathcal{G}}^{\text{lex}}} \text{ then } \Psi' \subseteq \Psi]$$

$$\iff \forall \Xi \in \Phi_{in} : s \in [\![\Xi]\!] \implies s' \in [\![\Xi]\!]$$

$$\iff s \preceq_{\mathcal{G}_{in}}^{\text{sub}} s'$$

where $\uparrow_{\mathcal{G}_{ex}} |s|_{\sim_{\mathcal{G}}^{\text{lex}}}$ denotes the upset of the equivalence class of s in the exclusive normal form of $\mathcal{G}$. The first equivalence holds by definition. The second holds holds under the assumption that $\mathcal{G}$ is exclusive, and hence that it has yields pre-orders with just as many equivalence classes as the elements of its domain. It just says that for any two elements of the inclusive normal form, if they coincide with the upsets of the equivalence classes of s and s' respectively, then the latter is included in the former. But this is to say—third equivalence—that for any property, if s satisfies it, so does s'.

So, any graph of disjoint components has an equivalent inclusive normal form. What about general graphs? In that case an inclusive normal form can be obtained from the exclusive normal form of the graph. To make it clear, the inclusive normal form of the graph given in Figure 5 (left) is not the graph given in Figure 6 (right), but a graph obtained from Figure 5 (right) which consists of all the upsets of the latter, ordered by set inclusion.

A.3   Graph equivalences

We can now pull together Lemmata 1 and 2 into one characterization theorem showing that the class of P-graphs defines, by lexicographic derivation, the same class of pre-orders that can be derived via the exclusive normal forms of those graphs by lexicographic derivation, or by subsumption-based derivation from the inclusive normal form of the exclusive normal form of the graph.

Theorem 4 (Equivalence of classes of graphs). *For any P-graph $\mathcal{G}$ and valuation $\mathcal{I}$:* $\preceq_{\mathcal{G}}^{\text{lex}} = \preceq_{\mathcal{G}_{in}}^{\text{lex}} = \preceq_{\mathcal{G}_{in}}^{\text{sub}}$.

Put it in a different way, the result can be illustrated by the following commutative diagram:

$$
\begin{array}{ccc}
\mathcal{G}_{ex} & \xrightarrow{\;in\;} & (\mathcal{G}_{ex})_{in} \\
{\scriptstyle ex}\uparrow & {\scriptstyle\text{lex}}\searrow & \downarrow{\scriptstyle\text{sub}} \\
\mathcal{G} & \xrightarrow[\text{lex}]{} & \preceq
\end{array}
$$

The pre-order $\preceq$ obtained from a general graph $\mathcal{G}$ via `lex` can alternatively be obtained by first extracting the exclusive normal form of $\mathcal{G}$ and then applying `lex`, or by extracting the inclusive normal form of (the exlusive normal form of) $\mathcal{G}$ and then applying `sub`.