Case Study: Priority Structures in Deontic Logic

Fenrong Liu

Tsinghua University & University of Amsterdam

8-10 July 2015, IRIT, Université Paul Sabatier, Toulouse

Joint work with

Johan van Benthem (Amsterdam & Stanford)
Davide Grossi (Liverpool)

Outline

Ideality

- Deontic notions involve a normative “ideality ordering”.
- This ordering ranks possible situations according to how well they meet moral demands.

“[...] 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.”

[Moore, *Principia Ethica*, 1922]

Ideality and Deontic logic

- Going back to [Hanson 1969], deontic logic has long used models involving a binary betterness/ideality ordering $\preceq$.
- Dyadic obligation $\mathbf{O}(\varphi \mid \psi)$ is then interpreted in terms of world comparison $s \preceq t$ (absolute obligation has $\psi = \top$):

$$\mathcal{M}, s \models \mathbf{O}(\varphi \mid \psi) \iff Max_{\preceq}(||\psi||_{\mathcal{M}}) \subseteq ||\varphi||_{\mathcal{M}} \qquad (1)$$

- Similar ideas of maximizing along an ordering occur in conditional logic, doxastic logic, and defeasible reasoning.

What is the source of ideality?

- Order structures offer a simple workable model for interpreting the meaning of deontic operators.
- But where do these orders come from? How are they defined?
- Two main themes for today:
  - **Reasons** that underlie ideality relations: resulting in a two-level modeling format.
  - **Dynamical changes** in deontic ordering, using both levels of this richer perspective.

Background for this lecture

- F. Liu, *Reasoning about Preference Dynamics*, Springer-Verlag, 2011. Synthese Library, Vol. 354.
- J. van Benthem and F. Liu, Deontic Logic and Changing Preference, *IFCOLOG Journal of Logics and their Applications*, 1(2): 1–46, 2014.
- J. van Benthem, D. Grossi and F. Liu, Priority Structures in Deontic Logic. *Theoria*, 80(2): 116–152, 2014.

Reasons for ideality ordering

"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."

St. Paul, First letter to the Corinthians, Standard American Bible, 1901.

This text identifies three deontically relevant properties of states:

$$(\neg t \vee m \vee \neg m) \prec (\neg t \vee m) \prec \neg t \qquad (2)$$

For simplicity, we omit the 'burn' (b) that would give $\neg t \vee m \vee (\neg m \wedge b)$.

Priority 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 such that, for all propositions $\varphi, \psi \in \Phi$, if $\varphi \prec \psi$ then ψ logically implies φ.

$\varphi \prec \psi$ means that the logically stronger formula ψ is strictly better (or, deontically more important) than the weaker formula φ.

More on P-graphs

- A P-graph is a finite graph of formulae from a propositional language, where each formula logically implies its immediate successors in the order.
- P-graphs with $\prec$ a strict linear order are called *P-sequences*.
- Each P-graph $\langle \Phi, \prec \rangle$ defines in a straightforward way a set of P-sequences as the longest sequences that can be built with pairs of formulae from the relation $\prec$.

Example of a P-graph

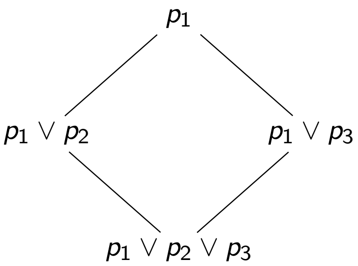


Figure: A P-graph. Convention: the higher, the better.

The graph defines two P-sequences:

$$\langle p_1, p_1 \vee p_2, p_1 \vee p_2 \vee p_3 \rangle \text{ and } \langle p_1, p_1 \vee p_3, p_1 \vee p_2 \vee p_3 \rangle.$$

From priority to ideality

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 **induced betterness relation** $\preceq_{\mathcal{G}} \subseteq S^2$ is defined as follows:

$$s \preceq_{\mathcal{G}} s' \quad := \quad \forall \varphi \in \Phi : s \in ||\varphi||_{\mathcal{I}} \Rightarrow s' \in ||\varphi||_{\mathcal{I}}. \qquad (3)$$

This definition (denote as ‘`sub`’) 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 $\prec$-lower properties in the graph.

Representation theorem

Fact

(a) Any P-graph generates a pre-order (reflexive and transitive).

(b) From any given pre-order, a P-graph can be constructed with a set of formulas which generate that pre-order.

Outline

Modal language

The basic **modal language of betterness** $\mathcal{L}(\forall, \preceq)$ is built from a set **P** of atoms according to the following inductive syntax:

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

We will relate this language to a standard deontic one in a while.

Models

A **model** for $\mathcal{L}(\forall, \preceq)$ on the set of atoms $\mathbf{P}$ is a standard modal structure $\mathcal{M} = \langle S, \preceq, \mathcal{I} \rangle$ where:

i) S is a non-empty set of states;
ii) $\preceq$ is a preorder over S ($s \preceq t$ stands intuitively for the earlier ideality order: "t is at least as good as s");
iii) $\mathcal{I} : \mathbf{P} \longrightarrow 2^S$.

It is also useful to define the usual strict suborder $\prec$ ("strictly better than") of the given $\preceq$: $s \prec t$ iff $s \preceq t$ and $t \not\preceq s$.

Truth definition

Let $\mathcal{M} = \langle S, \preceq, \mathcal{I} \rangle$ be a model. Truth for a formula $\varphi \in \mathcal{L}(\forall, \preceq)$ in a pointed model $(\mathcal{M}, s)$ is defined inductively:

$$\begin{array}{rcl} \mathcal{M}, s \models p & \iff & s \in \mathcal{I}(p) \\ \mathcal{M}, s \models [\preceq]\varphi & \iff & \forall s' \in S \text{ s.t. } s \preceq s' : \mathcal{M}, s' \models \varphi \\ \mathcal{M}, s \models [\forall]\,\varphi & \iff & \forall s' \in S : \mathcal{M}, s' \models \varphi \end{array}$$

Axiomatics

- A complete axiomatic proof calculus for our system consists of the standard modal logic **S4** for betterness, **S5** axioms for the universal modality, and one inclusion axiom $[\forall]\ \varphi \to [\preceq]\varphi$.
- This logic is known to be sound and strongly complete for preorders. Its uses go back at least to [Boutilier 1994].
- A richer system adding the strict betterness modality $[\prec]$:

  Van Benthem, Girard, and Roy, Everything else being equal: A modal logic for ceteris paribus preferences, *Journal of Philosophical Logic*, 2009, 38(1): 83–125.

Connecting betterness to deontics

The modal language can define the earlier deontic key notion.

In any model $\mathcal{M} = \langle S, \preceq, \mathcal{I} \rangle$ (modulo some technical conditions on the ordering that we omit in this lecture), it holds that:

$\mathcal{M}, s \models [\forall]\ (\psi \rightarrow \langle\preceq\rangle(\psi \wedge [\preceq](\psi \rightarrow \varphi))$ [conditional 'best']

$\Longleftrightarrow Max_{\preceq}(||\psi||_{\mathcal{M}}) \subseteq ||\varphi||_{\mathcal{M}}$ [relative maximality]

$\Longleftrightarrow \mathcal{M}, s \models \mathbf{O}\,(\varphi \mid \psi)$ [dyadic obligation]

Finer structure to be represented

- Standard deontic scenarios usually have more information than mere betterness order when you close-read their text: they also give reasons and criteria.
- To capture this additional deontic information, priority graphs come into play as a richer structure underneath (but correlated with) the betterness order.
- In the next part, we show how, in particular, deontic contrary-to-duty scenarios can be seen as priority graphs.

Outline

Contrary-to-duty scenario I: gentle murder

Here is the problem: 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:

- *It is obligatory under the law that Smith not murder Jones.*
- *It is obligatory that, if Smith murders Jones, Smith murders Jones gently.*

[Forrester 1984]

Gentle murder: from text to priority graph

The scenario mentions two relevant sets of states:

- Those in which Smith does not murder Jones, represented by the formula $\neg m$.
- Those in which either Smith does not murder Jones or he does murder Jones, but gently, i.e., $\neg m \vee (m \wedge g)$.

This generates a P-sequence $\mathcal{B} = \langle \Phi, \prec \rangle$ where

- $\Phi = \{\neg m, (\neg m \vee (m \wedge g))\}$
- $\prec$ is such that $(\neg m \vee (m \wedge g)) \prec \neg m$.

Gentle murder: priority and betterness

- For any set of states S and valuation $\mathcal{I}$ for the relevant atoms m and g, we get an induced betterness relation $\preceq_{\mathcal{B}}$.
- The relation $\preceq_{\mathcal{B}}$ orders states in three disjoint clusters:
  - The most ideal states: $\neg m$
  - Strictly worse but not worst: $m \wedge g$
  - Strictly worst: $m \wedge \neg g$, i.e., the states that do not satisfy any of the formulae in the P-sequence.

Another benchmark: Chisholm scenario

We follow the formulation proposed by [Aqvist 1967]:

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.

From scenario to priority graph

- The three items 1, 3 and 4 specify a priority sequence:

  $\neg r \vee (r \wedge p) \prec (\neg r \wedge \neg p)$

  (r: "Smith robs Jones" and p : "Smith is punished")

- Using dyadic obligations, 1, 3 and 4 say that

  $\mathbf{O}(\neg r \mid \top), \mathbf{O}(p \mid r), \mathbf{O}(\neg p \mid \neg r)$.

- The main difference with the gentle murder scenario is the **factual statement** "Smith robs Jones" (item 2).

Summary

- We think of P-graphs as the natural formalization of a CTD: some norms are given in the text, and the stated deontic obligations can be computed going ‘down the line’.
- What is the difference with standard semantic models?

  A mere betterness order on states has forgotten its origins – whereas now we have these available as reasons for the ordering, as well as a structure of relevant exception zones.

Outline

Dynamics for deontic agents

Deontic structure lives in a dynamic world of agency.

Two major ways in which obligations can change:

- Getting new information about the world,
- Changing the norms for evaluating worlds.

Update with hard information

The standard simplest pilot setting for information change:

$$\mathcal{M}, s \models [!\varphi]\psi \iff \text{if } \mathcal{M}, s \models \varphi \text{ then } \mathcal{M}|\varphi, s \models \psi \tag{4}$$

$$[!\varphi]\,[\preceq]\psi \;\leftrightarrow\; (\varphi \rightarrow [\preceq][!\varphi]\psi) \tag{5}$$

$$[!\varphi]\,\mathbf{O}\,(\psi \mid \chi) \;\leftrightarrow\; (\varphi \rightarrow \mathbf{O}\,([!\varphi]\psi \mid (\varphi \wedge [!\varphi]\chi)) \tag{6}$$

Similar dynamics can handle changing knowledge and belief entangled with deontic notions [van Benthem & Liu, 2014].

The Chisholm scenario revisited

$$\mathbf{O}(\neg r \mid \top) \wedge \mathbf{O}(p \mid r) \wedge \mathbf{O}(\neg p \mid \neg r) \rightarrow [!r]\mathbf{O}(p \mid \top) \qquad (7)$$

- The role of the factual statement suggests a dynamic-epistemic event where new information becomes available. It becomes settled/known that "Smith robs Jones".
- This triggers precise computable normative consequences, namely, in this case, that Smith ought to be punished.
- The model changes under the update, and in particular, the conditional obligation now becomes an absolute one.

One level up: 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 new graph $\mathcal{G}^\psi = \langle \Phi^\psi, \prec^\psi \rangle$ where:

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

Harmony of the two dynamics

Theorem (Harmony of announcement and restriction)

The following diagram connecting our two deontic levels commutes for all P-graphs $\mathcal{G}$, propositional formulae φ, and valuations $\mathcal{I}$:

$$
\begin{array}{ccc}
\mathcal{G} & \longrightarrow & \mathcal{G}^{\psi} \\
\downarrow \mathtt{sub} & & \downarrow \mathtt{sub} \\
\langle S, \preceq_{\mathcal{G}}, \mathcal{I} \rangle & \xrightarrow{!\psi} & \langle ||\psi||, \preceq_{\mathcal{G}^{\psi}}, \mathcal{I}|_{\psi} \rangle
\end{array}
$$

Deontic dynamics: postfixing norms

Now let us move from information change to norm change.

Consider the simple one-point P-sequence $\langle \neg m \rangle$. This induces a total preorder with all $\neg m$ states above all m states: "It is obligatory under the law that Smith not murder Jones".

Now a lawgiver comes in, and introduces a sub-ideal obligation:

"if Smith murders Jones, he must do so gently".

This can be done by postfixing with the property $\neg m \vee g$:

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

More deontic dynamics: prefixing norms

Suppose we start with the P-sequence:

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

Now we want to introduce a stronger norm, like "It is obligatory under the law that Smith not murder Jones and that Smith not be aggressive against Jones".

This can be achieved by prefixing the property $\neg m \wedge \neg a$, where a stands for "Smith is aggressive against Jones":

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

Pictorial representation

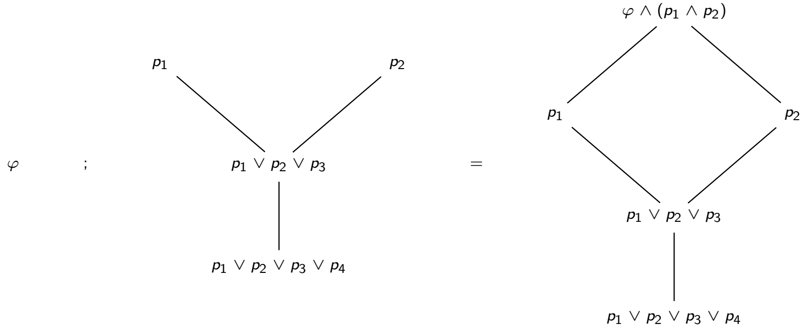


Figure: Hasse diagrams illustrating pre-fixing φ to the original P-graph.

More Hasse diagrams

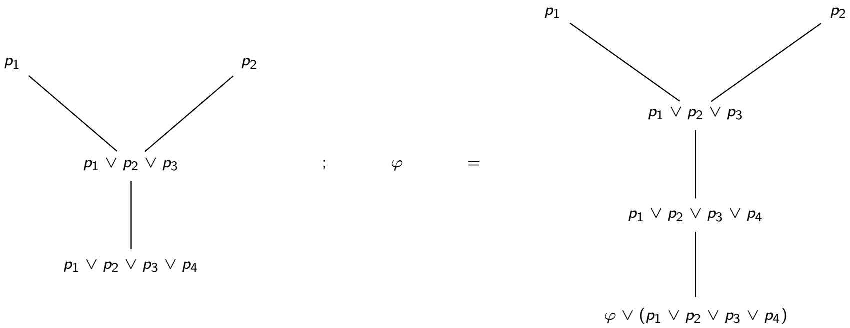


Figure: Hasse diagrams for post-fixing φ to an earlier P-graph.

Some formal definitions

Let $\mathcal{G} = \langle \Phi, \prec \rangle$ be a P-graph, and φ a propositional formula:

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

Radical upgrade

As it happens, a counterpart at the lower betterness level of deontic structure already exists in work on belief revision:

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:

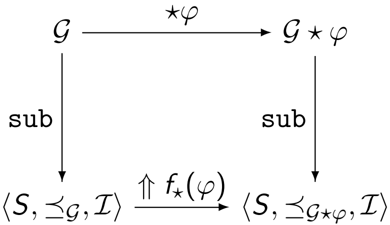
$$\preceq^{\Uparrow\varphi} = (?\varphi; \preceq; ?\varphi) \cup (?\neg\varphi; \preceq; ?\neg\varphi) \cup (?\neg\varphi; S^2; ?\varphi).$$

In other words, a radical upgrade $\Uparrow \varphi$ changes the current order $\preceq$ to one where all φ-states become better than all $\neg\varphi$-states, while, within those two zones, the old ordering remains intact.

Two-level harmony once more

Theorem (P-graph pre-postfixing matches radical upgrade)

The following diagram between our two deontic levels commutes for all P-graphs $\mathcal{G}$, propositional formulae φ, and valuations $\mathcal{I}$:



Here $\mathcal{G} \star \varphi$ denotes either pre-fixing $\varphi; \mathcal{G}$ or post-fixing $\mathcal{G}; \varphi$ of $\mathcal{G}$ by φ and accordingly, $f_\star(\varphi)$ is $\varphi \wedge \bigwedge Max(\mathcal{G})$ or $\varphi \vee \bigvee \min(\mathcal{G})$.

Summary

- Norms represented as P-graphs induce ideality orderings. Basic norms have just one property, inducing bipartitions.
- Normative systems (sets of norms) can be represented as complex P-graphs. In particular, CTD scenarios can often be represented faithfully by linear P-sequences.
- Dynamic changes at our two levels can have perfect matches.
- But the richer priority level suggests a richer dynamics than that of pure betterness ordering. What other meaningful operations can be performed on P-graphs?
- AIM: a general graph-based theory of normative systems.

Further graph operations: norm merge

First normative system:

1. It is obligatory under the law that Smith not murder Jones.
2. If Smith murders Jones, he must do so gently.

Second normative system:

1. It is obligatory under the law that Smith not murder Jones.
2. If Smith murders Jones, he must do so *quickly*.

Norm change diagrams

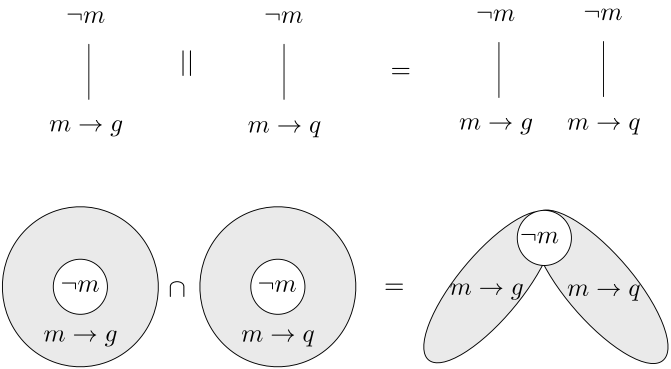


Figure: Merging P-sequences by parallel composition of P-sequences, and underneath, as intersection of induced total preorders.

Harmony for parallel composition

Theorem (Parellel composition matches intersection of orders)

With notations as before, the following diagram commutes:

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

Moving P-graphs into deontic language

- Graph dynamics leads to extended dynamic deontic logic.
- Introducing priority graphs explicitly inside the deontic language can be done, with complete logics resulting. See Patrick Girard's dissertation, 2008.
- Recent line: merge with ceteris-paribus preference logic.

Further operations and tracking

- What about other deontically relevant operations? Say, **deleting** a norm, simplifying a legal system?
- Things can then be more complex, some graph-level operations are **irreducible** and lack counterparts at the lower betterness level: no commuting diagrams.
- General issue for different representation levels (information, evidence, norms). When can natural dynamic updates at one level be **tracked** faithfully by updates at another level?

Outline

Conclusions

- We have shown how well-known deontic scenarios can be mined for more structure than just deontic inferences. A framework is proposed for representing not just standard ideality orderings, but also underlying sources or reasons.
- Our framework supports both information dynamics and norm change, and we saw how these lead to deontic obligation change. The dynamics worked at two levels, and we have explored some tracking results.

Future directions

- Add belief or knowledge into the picture, and study the many interactions between ideality, priorities and belief.
- In particular, the analogies with belief dynamics show a unity in logical modeling that can be exploited further. E.g., working at both levels, we can also study reasons for beliefs. One relevant approach: dynamic logics of "evidence models".
- Add social aspects, and explore how norms can be socially constructed or changed through multi-agent interaction.