To do something else

Ju Fengkui

School of Philosophy, Beijing Normal University

May 21, 2016

Joint work with Jan van Eijck from Centrum Wiskunde & Informatica

Outline

An old idea

There was an old idea in deontic logic: *an action is prohibited if doing it would cause a bad thing; an action is permitted if it is possible of doing it without causing a bad thing; an action is obligated if refraining to do it would cause a bad thing*.

An old idea

There was an old idea in deontic logic: *an action is prohibited if doing it would cause a bad thing; an action is permitted if it is possible of doing it without causing a bad thing; an action is obligated if refraining to do it would cause a bad thing.*

The three fundamental normative notions, *prohibition*, *permission* and *obligation*, can be defined in terms of the consequences of doing actions.

An old idea

There was an old idea in deontic logic: *an action is prohibited if doing it would cause a bad thing; an action is permitted if it is possible of doing it without causing a bad thing; an action is obligated if refraining to do it would cause a bad thing.*

The three fundamental normative notions, *prohibition*, *permission* and *obligation*, can be defined in terms of the consequences of doing actions.

Anderson [1967] and Kanger [1971] followed this idea and independently developed a deontic logic. This is a classical modal logic and it leads to quite many problems.

A dynamic deontic logic

Starting from the same idea, Meyer [1988] proposed a deontic logic based on a fragment of PDL.

Starting from the same idea, Meyer [1988] proposed a deontic logic based on a fragment of PDL.

This work applies deontic operators to actions and many problems with previous deontic logics are avoided.

Two problems with Meyer’s work

Meyer’s work does not handle the normative notions very well, as the following statements are satisfiable in it:

- Killing the president is not allowed, but killing him and then surrendering to the police is.
- Rescuing the injured and then calling an ambulance is obligated, but rescuing the injured is not.

Two problems with Meyer's work

Meyer's work does not handle the normative notions very well, as the following statements are satisfiable in it:

- Killing the president is not allowed, but killing him and then surrendering to the police is.
- Rescuing the injured and then calling an ambulance is obligated, but rescuing the injured is not.

Its formalization of refraining to do something is not reasonable:

- The intersection of $\overline{\alpha}$ and α is not always empty. This means that there may be ways to refrain from α while at the same time doing α.
- The intersection of $\overline{\alpha}$ and $\alpha;\beta$ is not always empty. This means that performing $\alpha;\beta$ may be a way to refrain from doing α.

Outline

To do something else

The principle of symmetry: if doing α is doing something else than β, then doing β is also doing something else than α.

To do something else

The principle of symmetry: if doing α is doing something else than β, then doing β is also doing something else than α.

The principle of irrevocable history: if you have done something, then you will always have done it.

To do something else

The principle of symmetry: if doing α is doing something else than β, then doing β is also doing something else than α.

The principle of irrevocable history: if you have done something, then you will always have done it.

Let a and b be two different atomic actions. Fix a start point. When would we say that the agent has done something else than $a; b$?

To do something else

The principle of symmetry: if doing α is doing something else than β, then doing β is also doing something else than α.

The principle of irrevocable history: if you have done something, then you will always have done it.

Let a and b be two different atomic actions. Fix a start point. When would we say that the agent has done something else than $a; b$?

Clearly if the agent has done a, he has done something else than b. Then if he has done $a; b$, he has done something else than b. Then if he has done b, he has done something else than $a; b$.

To do something else

The principle of symmetry: if doing α is doing something else than β, then doing β is also doing something else than α.

The principle of irrevocable history: if you have done something, then you will always have done it.

Let a and b be two different atomic actions. Fix a start point. When would we say that the agent has done something else than $a; b$?

Clearly if the agent has done a, he has done something else than b. Then if he has done $a; b$, he has done something else than b. Then if he has done b, he has done something else than $a; b$.

We can not say that if the agent has done a, he has done something else than $a; b$, otherwise we have to say that if he has done $a; b$, he has done something else than $a; b$.

A sharpened idea of reducing normative notions

There are a class of states, a group of people and an agent. Some states are *bad* and some are *fine* for this group. The agent doing an action at a state changes this state to another one.

A sharpened idea of reducing normative notions

There are a class of states, a group of people and an agent. Some states are *bad* and some are *fine* for this group. The agent doing an action at a state changes this state to another one.

An action is *prohibited* at a state if the state will be bad at some point during any performance of this action.

A sharpened idea of reducing normative notions

There are a class of states, a group of people and an agent. Some states are *bad* and some are *fine* for this group. The agent doing an action at a state changes this state to another one.

An action is *prohibited* at a state if the state will be bad at some point during any performance of this action.

An action is *permitted* at a state if the state will always be fine during some performance of this action.

A sharpened idea of reducing normative notions

There are a class of states, a group of people and an agent. Some states are *bad* and some are *fine* for this group. The agent doing an action at a state changes this state to another one.

An action is *prohibited* at a state if the state will be bad at some point during any performance of this action.

An action is *permitted* at a state if the state will always be fine during some performance of this action.

An action is *obligated* at a state if the state will be bad at some point during any performance of *anything else*.

Outline

Language

Let Π_0 be a *finite* set of atomic actions and Φ_0 a *countable* set of atomic propositions.

$$\alpha ::= a \mid 0 \mid (\alpha ; \alpha) \mid (\alpha \cup \alpha) \mid \alpha^*$$
$$\phi ::= p \mid \top \mid \mathfrak{b} \mid \neg\phi \mid (\phi \wedge \phi) \mid \|\alpha\|\phi$$

Language

Let Π_0 be a *finite* set of atomic actions and Φ_0 a *countable* set of atomic propositions.

$$\begin{aligned} \alpha &::= a \mid 0 \mid (\alpha;\alpha) \mid (\alpha \cup \alpha) \mid \alpha^* \\ \phi &::= p \mid \top \mid \mathfrak{b} \mid \neg\phi \mid (\phi \wedge \phi) \mid \|\alpha\|\phi \end{aligned}$$

$\mathfrak{b}$ means that this is a *bad* world.

Language

Let Π_0 be a *finite* set of atomic actions and Φ_0 a *countable* set of atomic propositions.

$$\alpha ::= a \mid 0 \mid (\alpha;\alpha) \mid (\alpha \cup \alpha) \mid \alpha^*$$
$$\phi ::= p \mid \top \mid \mathfrak{b} \mid \neg\phi \mid (\phi \wedge \phi) \mid \|\alpha\|\phi$$

$\mathfrak{b}$ means that this is a *bad* world.

$\|\alpha\|\phi$ means that for any way to perform α, ϕ will be the case at some point in the process.

Language

Let Π_0 be a *finite* set of atomic actions and Φ_0 a *countable* set of atomic propositions.

$$\alpha ::= a \mid 0 \mid (\alpha;\alpha) \mid (\alpha \cup \alpha) \mid \alpha^*$$
$$\phi ::= p \mid \top \mid \mathfrak{b} \mid \neg\phi \mid (\phi \wedge \phi) \mid \|\alpha\|\phi$$

$\mathfrak{b}$ means that this is a *bad* world.

$\|\alpha\|\phi$ means that for any way to perform α, ϕ will be the case at some point in the process.

The dual $\langle\!\langle\alpha\rangle\!\rangle\phi$ of $\|\alpha\|\phi$ says that there is a way to perform α s.t. ϕ will be the case at all the points in the process.

Language

Let Π_0 be a *finite* set of atomic actions and Φ_0 a *countable* set of atomic propositions.

$$\alpha ::= a \mid 0 \mid (\alpha;\alpha) \mid (\alpha \cup \alpha) \mid \alpha^*$$
$$\phi ::= p \mid \top \mid \mathfrak{b} \mid \neg\phi \mid (\phi \wedge \phi) \mid \|\alpha\|\phi$$

$\mathfrak{b}$ means that this is a *bad* world.

$\|\alpha\|\phi$ means that for any way to perform α, ϕ will be the case at some point in the process.

The dual $\langle\!\langle\alpha\rangle\!\rangle\phi$ of $\|\alpha\|\phi$ says that there is a way to perform α s.t. ϕ will be the case at all the points in the process.

$F\alpha$, α is *forbidden*, is defined as $\|\alpha\|\mathfrak{b}$. $P\alpha$, α is *permitted*, is defined as $\langle\!\langle\alpha\rangle\!\rangle\neg\mathfrak{b}$.

Models

$\mathfrak{M} = (W, \{R_a \mid a \in \Pi_0\}, B, V)$ is a model where

- W is a nonempty set of states
- for any $a \in \Pi_0$, $R_a \subseteq W \times W$, and for any $a, b \in \Pi_0$, $R_a \cap R_b = \varnothing$
- $B \subseteq W$
- V is a function from Φ_0 to 2^W

Models

$\mathfrak{M} = (W, \{R_a \mid a \in \Pi_0\}, B, V)$ is a model where

- W is a nonempty set of states
- for any $a \in \Pi_0$, $R_a \subseteq W \times W$, and for any $a, b \in \Pi_0$, $R_a \cap R_b = \varnothing$
- $B \subseteq W$
- V is a function from Φ_0 to 2^W

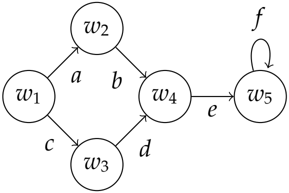
Roughly a model is a labeled transion system plus a set of bad states.

Interpretation of actions

Each finite sequence of states is called a trace. Each action α corresponds to a set S_α of traces.

Interpretation of actions

Each finite sequence of states is called a trace. Each action α corresponds to a set S_α of traces.



$$
\begin{aligned}
S_{a;b} &= \{w_1 w_2 w_4\}\\
S_{c;d} &= \{w_1 w_3 w_4\}\\
S_{(a;b)\cup(c;d)} &= \{w_1 w_2 w_4, w_1 w_3 w_4\}\\
S_{f*} &= \{w_5, w_5 w_5, w_5 w_5 w_5, \ldots\}\\
&\vdots
\end{aligned}
$$

Semantics

$\mathfrak{M}, w \Vdash \phi$, ϕ being true at w in $\mathfrak{M}$, is defined as follows:

- $\mathfrak{M}, w \Vdash \mathfrak{b} \Leftrightarrow w \in B$
- $\mathfrak{M}, w \Vdash F\alpha \Leftrightarrow$ for any trace $w_0 \dots w_n$, if $w_0 = w$ and $w_0 \dots w_n \in S_\alpha$, then $\mathfrak{M}, w_i \Vdash \mathfrak{b}$ for some $i \leq n$ s.t. $1 \leq i \leq n$
- $\mathfrak{M}, w \Vdash P\alpha \Leftrightarrow$ there is a trace $w_0 \dots w_n$ s.t. $w_0 = w$, $w_0 \dots w_n \in S_\alpha$ and $\mathfrak{M}, w_i \Vdash \neg\mathfrak{b}$ for any $i \leq n$ s.t. $1 \leq i \leq n$

Outline

Computation sequences

A computation sequence, called a *seq*, is a sequence of atomic actions. An example: $abab$ is a seq.

Computation sequences

A computation sequence, called a *seq*, is a sequence of atomic actions. An example: *abab* is a seq.

Each action α corresponds to a set $CS(\alpha)$ of seqs.

- $CS(a;b) = \{ab\}$
- $CS((a;b) \cup (c;d)) = \{ab, cd\}$
- $CS(a^*) = \{\epsilon, a, aa, \ldots\}$

The relation of x-different

Let $\sqsubseteq$ denote the relation of *initial segment*. For any seqs σ and τ, σ is *x-different* from τ, $\sigma \not\approx \tau$, if $\sigma \not\sqsubseteq \tau$ and $\tau \not\sqsubseteq \sigma$.

The relation of x-different

Let $\sqsubseteq$ denote the relation of *initial segment*. For any seqs σ and τ, σ is *x-different* from τ, $\sigma \nsim \tau$, if $\sigma \not\sqsubseteq \tau$ and $\tau \not\sqsubseteq \sigma$.

An example: b is x-different from ab, but a is not x-different from ab.

The relation of x-different

Let $\sqsubseteq$ denote the relation of *initial segment*. For any seqs σ and τ, σ is *x-different* from τ, $\sigma \not\approx \tau$, if $\sigma \not\sqsubseteq \tau$ and $\tau \not\sqsubseteq \sigma$.

An example: b is x-different from ab, but a is not x-different from ab.

For any actions α and β, α is *x-different* from β, $\alpha \not\approx \beta$, if for any seqs $\sigma \in CS(\alpha)$ and $\tau \in CS(\beta)$, $\sigma \not\approx \tau$.

The relation of x-different

Let $\sqsubseteq$ denote the relation of *initial segment*. For any seqs σ and τ, σ is *x-different* from τ, $\sigma \not\approx \tau$, if $\sigma \not\sqsubseteq \tau$ and $\tau \not\sqsubseteq \sigma$.

An example: b is x-different from ab, but a is not x-different from ab.

For any actions α and β, α is *x-different* from β, $\alpha \not\approx \beta$, if for any seqs $\sigma \in CS(\alpha)$ and $\tau \in CS(\beta)$, $\sigma \not\approx \tau$.

The relation of x-different formalizes the word "else" in the imperatives such as "don't watch cartoon and do something else". If α is x-different from β, then doing α is doing something else from β.

To do something else

For an action α, there might be many actions each of which is something else. The relation of x-different itself is not enough to handle the notion of to do something else, as the latter also involves a quantifier over actions.

To do something else

For an action α, there might be many actions each of which is something else. The relation of x-different itself is not enough to handle the notion of to do something else, as the latter also involves a quantifier over actions.

Luckily, for any α, among the actions which are something else, there is a *greatest* one in the sense that it is the union of all of them. This lets us deal with the notion of to do something else without introducing any quantifier over actions.

To do something else

Let Δ be a set of seqs. Define $\widetilde{\Delta}$, called the *opposite* of Δ, as the set $\{\tau \mid \tau \not\asymp \sigma \textit{ for any } \sigma \in \Delta\}$.

To do something else

Let Δ be a set of seqs. Define $\widetilde{\Delta}$, called the *opposite* of Δ, as the set $\{\tau \mid \tau \not\approx \sigma \textit{ for any } \sigma \in \Delta\}$.

Proposition

For any α, there is a β s.t. $CS(\beta) = \widetilde{CS(\alpha)}$.

To do something else

Let Δ be a set of seqs. Define $\widetilde{\Delta}$, called the *opposite* of Δ, as the set $\{\tau \mid \tau \not\lesssim \sigma \textit{ for any } \sigma \in \Delta\}$.

Proposition

For any α, there is a β s.t. $CS(\beta) = \widetilde{CS(\alpha)}$.

β is called the *opposite* of α, denoted as $\widetilde{\alpha}$. An example: let $\Pi_0 = \{a, b\}$; then $\widetilde{a} = b; (a \cup b)^*$.

To do something else

Let Δ be a set of seqs. Define $\widetilde{\Delta}$, called the *opposite* of Δ, as the set $\{\tau \mid \tau \not\approx \sigma \textit{ for any } \sigma \in \Delta\}$.

Proposition

For any α, there is a β s.t. $CS(\beta) = \widetilde{CS(\alpha)}$.

β is called the *opposite* of α, denoted as $\widetilde{\alpha}$. An example: let $\Pi_0 = \{a, b\}$; then $\widetilde{a} = b; (a \cup b)^*$.

Proposition

$CS(\widetilde{\alpha}) = \bigcup\{CS(\beta) \mid \beta \not\approx \alpha\}$.

To do something else

Let Δ be a set of seqs. Define $\widetilde{\Delta}$, called the *opposite* of Δ, as the set $\{\tau \mid \tau \not\approx \sigma \textit{ for any } \sigma \in \Delta\}$.

Proposition

For any α, there is a β s.t. $CS(\beta) = \widetilde{CS(\alpha)}$.

β is called the *opposite* of α, denoted as $\widetilde{\alpha}$. An example: let $\Pi_0 = \{a, b\}$; then $\widetilde{a} = b; (a \cup b)^*$.

Proposition

$CS(\widetilde{\alpha}) = \bigcup\{CS(\beta) \mid \beta \not\approx \alpha\}$.

$\widetilde{\alpha}$ is the union of all the actions x-different from α. To refrain to do α is to do something else; to do *anything else* is to do $\widetilde{\alpha}$.

Obligations

$O\alpha$, α is *obligatory*, is defined as $\|\widetilde{\alpha}\|\mathfrak{b}$; it means that no matter what else except α to do and how to do it, the state will be bad at some point in the process.

Obligations

$O\alpha$, α is *obligatory*, is defined as $\|\widetilde{\alpha}\|\mathfrak{b}$; it means that no matter what else except α to do and how to do it, the state will be bad at some point in the process.

$\mathfrak{M}, w \Vdash O\alpha \Leftrightarrow$ for any trace $w_0 \ldots w_n$, if $w_0 = w$ and $w_0 \ldots w_n \in S_{\widetilde{\alpha}}$, then $w_i \Vdash \mathfrak{b}$ for some $i \leq n$

Obligations

$O\alpha$, α is *obligatory*, is defined as $\|\widetilde{\alpha}\|\mathfrak{b}$; it means that no matter what else except α to do and how to do it, the state will be bad at some point in the process.

$\mathfrak{M}, w \Vdash O\alpha \Leftrightarrow$ for any trace $w_0 \ldots w_n$, if $w_0 = w$ and $w_0 \ldots w_n \in S_{\widetilde{\alpha}}$, then $w_i \Vdash \mathfrak{b}$ for some $i \leq n$

It can be verified that our work does not suffer from the two problem with Meyer's work.

Obligations

$O\alpha$, α is *obligatory*, is defined as $\|\widetilde{\alpha}\|\mathfrak{b}$; it means that no matter what else except α to do and how to do it, the state will be bad at some point in the process.

$\mathfrak{M}, w \Vdash O\alpha \Leftrightarrow$ for any trace $w_0 \ldots w_n$, if $w_0 = w$ and $w_0 \ldots w_n \in S_{\widetilde{\alpha}}$, then $w_i \Vdash \mathfrak{b}$ for some $i \leq n$

It can be verified that our work does not suffer from the two problem with Meyer's work.

A test $\phi?$ in trace semantics is a set of states in which ϕ is true. Trivially $F(\phi?)$ is never true and $P(\phi?)$ is always true. This means that there is no restriction on testing and testing is always free.

Is there a related algebra?

Let $\mathcal{C}$ be the set of concise sets of seqs. Is there a known algebra related with $(\mathcal{C}, \otimes, \cup, ^*, \widetilde{\ })$?

Outline

Future work

Does this logic have complete axiomatization?

Future work

Does this logic have complete axiomatization?

What is the computational complexity of it?

Future work

Does this logic have complete axiomatization?

What is the computational complexity of it?

How do we dynamize it?

Future work

Does this logic have complete axiomatization?

What is the computational complexity of it?

How do we dynamize it?

Instead of the division of good and evil states, we want to introduce the more fine-grained notion of *betterness* in further work.

Thanks!

References

A. R. Anderson. Some nasty problems in the formal logic of ethics. *Noûs*, 1(4):345–360, 1967.

J. Broersen. Action negation and alternative reductions for dynamic deontic logics. *Journal of Applied Logic*, 2(1):153–168, 2004.

F. Ju and J. Cai. Two process modalities in trace semantics. Manuscript, 2015.

F. Ju, N. Cui, and S. Li. Trace semantics for IPDL. In W. van der Hoek, Wesley H. Holliday, and W. Wang, editors, *Logic, Rationality, and Interaction*, volume 9394 of *Lecture Notes in Computer Science*, pages 169–181. Springer Berlin Heidelberg, 2015.

S. Kanger. New foundations for ethical theory. In R. Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 36–58. Springer Netherlands, 1971.

J.-J. CH. Meyer. A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29(1):109–136, 1988.

X. Sun and H. Dong. Deontic logic based on a decidable PDL with action negation. Manuscript, 2015.

R. van der Meyden. The dynamic logic of permission. *Journal of Logic and Computation*, 6(3):465–479, 1996.