A Preliminary Study on Deontic Answer Set Programming

Pedro Cabalar[1], Agata Ciabattoni[2], Leon van der Torre[3]

[1]University of Corunna, SPAIN
[2]Technical University of Vienna, AUSTRIA
[3]University of Luxembourg, LUXEMBOURG

November 3rd, 2022
Cercedilla, Spain

Deontic Reasoning

- Deontic reasoning: obligation, prohibition, violation, fulfilment, contrary to duty, deontic/factual detachment, ...
- There exists a vast literature dating back to 1920's
- Family of (so-called) paradoxes (Chisholm, Ross, Gentle Murder, Good Samaritan, ...)
- Standard Deontic Logic (SDL) = Modal logic **K** plus axiom $D : \Box\varphi \rightarrow \Diamond\varphi$ meaning serial Kripke frames
  Nowadays considered unsatisfactory for many of the paradoxes or examples above
- Some paradoxes involve default reasoning but were proposed before NMR even exist!
  Others deal with obligations for compound formulas $\mathbf{O}(p \vee q)$, $\mathbf{O}(p \wedge q)$, etc, but many cases just use literals $\mathbf{O}p$, $\mathbf{O} \sim q$.

Deontic Reasoning

Let's use Answer Set Programming! It comes with two negations:

- `not p` false by default $\neg p$
- `-p` explicitly false $\sim p$

```
park :- not -can_park.    % no evidence of prohibition
park :- can_park.         % evidence of permission
```

$$park \leftarrow \neg \sim can\_park$$
$$park \leftarrow can\_park$$

Introducing deontic operators

Example: read p as “parking”

- Def. a deontic atom can be:
  - p = reality = “We have evidence that p was done”
  - $\mathbf{O}p$ = obligation = “p is obligatory"
  - $\mathbf{F}p$ = prohibition = “p is forbidden"

  Why both $\mathbf{O}p$ and $\mathbf{F}p$ as primitive?

  ☞ Paraconsistency: may coexist is some cases (we’ll see later)

- Def. an explicit literal is a deontic atom or its explicit negation:
  - $\sim p$ = reality = “We have evidence that p was not done”
  - $\sim\mathbf{O}p$ = “Explicit permission for $\sim p$"
  - $\sim\mathbf{F}p$ = “Explicit permission for p"

Introducing deontic operators

Example: read p as "parking"

- Def. a default literal can be an explicit literal or its default negation:
  - $\neg p$ = reality = "No evidence that p was done"
  - $\neg \mathbf{O}p$ = "Implicit permission for $\sim p$"
  - $\neg \mathbf{F}p$ = "Implicit permission for p"
  - $\neg \sim p$ = reality = "No evidence that p was not done"
  - $\neg \sim \mathbf{O}p$ = "No evidence on explicit permission for $\sim p$"
  - $\neg \sim \mathbf{F}p$ = "No evidence on explicit permission for p"

Deontic Programs

- A deontic logic program is a set of rules of the form

$$H_1 \vee \cdots \vee H_m \leftarrow B_1 \wedge \cdots \wedge B_n$$

with $n \geq 0$, $m \geq 0$ and B_i, H_i default literals.

Ex1 *Park when no evidence on a prohibition* (implicit permission)

$$park \leftarrow \neg \mathbf{F} park$$

Ex2 *I must normally work,*
On weekends, I have an explicit permission not to work
It is not a weekend, I decided not to work

$$\begin{array}{rcll} \mathbf{O}work & \leftarrow & \neg \sim \mathbf{O}work & \quad \sim weekend \\ \sim \mathbf{O}work & \leftarrow & weekend & \quad \sim work \end{array}$$

We derive $\mathbf{O}work \wedge \sim work$ = violation

Semantics

Ex3 *You must fight in the army* $\mathbf{O}\mathit{fight}$
You must not fight in the army $\mathbf{F}\mathit{fight}$
These two things alone should be inconsistent

Ex4 Contrary to Duty (CTD)
I must not walk in the street
If I walk in the street, I must walk on the right side of the street

$$\begin{array}{rcl} \mathbf{F}\mathit{walk} & & \\ \mathbf{O}\mathit{walk\_right} & \leftarrow & \mathit{walk} \\ \mathit{walk} & \leftarrow & \mathit{walk\_right} \\ \mathbf{O}\mathit{walk} & \leftarrow & \mathbf{O}\mathit{walk\_right} \end{array}$$

If we add the fact *walk*, we derive both $\mathbf{O}\mathit{walk}$ and $\mathbf{F}\mathit{walk}$
No inconsistency: $\mathit{walk} \wedge \mathbf{F}\mathit{walk}$ is a violation that "enables" $\mathbf{O}\mathit{walk}$

Deontic Answer Sets

Definition (Deontic Interpretation)

A *(deontic) interpretation* T is a set of explicit literals satisfying both:

1. For any deontic atom A, $\{A, \sim A\} \not\subseteq T$;
2. For any atom $p \in At$, if $\{p, \sim p\} \cap T = \emptyset$ then $\{\mathbf{O}p, \mathbf{F}p\} \not\subseteq T$.

You cannot have $\mathbf{O}p$ and $\mathbf{F}p$ but "no decision" about p

$$\bot \leftarrow \mathbf{O}p, \mathbf{F}p, \neg p, \neg \sim p$$

- Def. Π^T is the reduct of program Π w.r.t interpretation T as in standard ASP = replace $\neg L$ by $\top$ if $L \in T$ or $\bot$ otherwise.
- T is an answer set if it is a minimal model of Π^T (understanding explicit literals as "classical atoms")

Fences paradox

Ex5 *(Normally) there must be no fence*
But if you do put a fence, it must be white
By the sea, you can put a fence
You decide to put a fence
Should I paint it in white?

$$\begin{array}{rcl} \mathbf{F}\mathit{fence} & \leftarrow & \neg \sim\mathbf{F}\mathit{fence} \\ \mathbf{O}\mathit{white} & \leftarrow & \mathit{fence} \wedge \mathbf{F}\mathit{fence} \\ \sim\mathbf{F}\mathit{fence} & \leftarrow & \mathit{sea} \\ \mathit{fence} & & \end{array}$$

What if I am by the *sea*?

Practical Reasoning

- $p \wedge \mathbf{O}p$ = obligation is fulfilled
- $\sim p \wedge \mathbf{F}p$ = prohibition is fulfilled
- $\sim p \wedge \mathbf{O}p$ = obligation is violated
- $p \wedge \mathbf{F}p$ = prohibition is violated
- Sometimes we may be interested in modelling “respectful behaviour" by default (a.k.a. practical reasoning)

$$\begin{array}{rcl} \sim fence & \leftarrow & \mathbf{F}fence \wedge \neg fence \\ fence & \leftarrow & \mathbf{O}fence \wedge \neg \sim fence \end{array}$$

Deontic Equilibrium Logic

- In ASP, explicit negation $\sim p$ is only applied to atoms, not to arbitrary formulas
- Can we define its semantics as a real operator? (arbitrarily nested with others)
- $\mathcal{N}_5$ Equilibrium Logic with strong negation [Pearce 1997]
  = Equilibrium Logic + Nelson's strong negation [Nelson 1949]
- $\mathcal{X}_5$ Equilibrium Logic with explicit negation [Aguado & al ECAI 2020]
  - $\mathcal{N}_5$ does not fit with (natural extension) of program reduct
  - In $\mathcal{N}_5$, $\neg\neg\neg p \not\Leftrightarrow \neg p$ (in the scope of $\sim$)
  - $\mathcal{X}_5$ = minor variation in $\rightarrow$ truth-table that fixes those problems

Equilibrium Logic with Explicit Negation

- Negation Normal Form: $\mathcal{X}_5/\mathcal{N}_5$ common equivalences

$$
\begin{array}{rcl}
\sim\top & \Leftrightarrow & \bot \\
\sim(\varphi \wedge \psi) & \Leftrightarrow & \sim\varphi \vee \sim\psi \\
\sim\sim\varphi & \Leftrightarrow & \varphi
\end{array}
\qquad
\begin{array}{rcl}
\sim\bot & \Leftrightarrow & \top \\
\sim(\varphi \vee \psi) & \Leftrightarrow & \sim\varphi \wedge \sim\psi
\end{array}
$$

For the outermost occurrence of $\sim$ we can make the replacements

	in $\mathcal{N}_5$	in $\mathcal{X}_5$
$\sim(\alpha \to \beta)$	$\alpha \wedge \sim\beta$	$\neg\neg\alpha \wedge \sim\beta$
$\sim\neg\alpha$	α	$\neg\neg\alpha$

But not for inner occurrences:

$(p \to q) \Leftrightarrow \sim\sim(p \to q) \not\Leftrightarrow \sim(\neg\neg p \wedge \sim q) \Leftrightarrow \sim\neg\neg p \vee \sim\sim q \leftrightarrow \neg p \vee q$

Deontic Equilibrium Logic

- Can we go further and add **O** and **F** to $\mathcal{X}_5$?
- Syntax:

$$\varphi ::= p \mid \bot \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \rightarrow \varphi \mid \sim\varphi \mid \mathbf{O}\varphi \mid \mathbf{F}\varphi$$

- Abbreviations

$$\begin{aligned}
\neg\varphi &\stackrel{\text{def}}{=} (\varphi \rightarrow \bot) \\
\top &\stackrel{\text{def}}{=} \neg\bot \\
\mathbf{P}\varphi &\stackrel{\text{def}}{=} \sim\mathbf{F}\varphi \\
\widehat{\mathbf{O}}\varphi &\stackrel{\text{def}}{=} (\neg\mathbf{P}\sim\varphi \rightarrow \mathbf{O}\varphi) \\
\widehat{\mathbf{F}}\varphi &\stackrel{\text{def}}{=} (\neg\mathbf{P}\varphi \rightarrow \mathbf{F}\varphi)
\end{aligned}$$

Deontic Equilibrium Logic

- More abbreviations

$$
\begin{aligned}
\varphi \leftrightarrow \psi &\stackrel{\text{def}}{=} (\varphi \to \psi) \wedge (\psi \to \varphi) \\
\varphi \Rightarrow \psi &\stackrel{\text{def}}{=} (\varphi \to \psi) \wedge ({\sim}\psi \to {\sim}\varphi) \\
\varphi \Leftrightarrow \psi &\stackrel{\text{def}}{=} (\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi) \\
\mathbf{O}^{-}\varphi &\stackrel{\text{def}}{=} \mathbf{O}\varphi \wedge {\sim}\varphi \\
\mathbf{O}^{+}\varphi &\stackrel{\text{def}}{=} \mathbf{O}\varphi \wedge \varphi \\
\mathbf{O}^{?}\varphi &\stackrel{\text{def}}{=} \mathbf{O}\varphi \wedge \neg\varphi \wedge \neg {\sim}\varphi \\
\mathbf{O}^{+?}\varphi &\stackrel{\text{def}}{=} \mathbf{O}\varphi \wedge \neg {\sim}\varphi \\
\mathbf{O}^{-?}\varphi &\stackrel{\text{def}}{=} \mathbf{O}\varphi \wedge \neg\varphi
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{O}(\alpha \mid \beta) &\stackrel{\text{def}}{=} (\mathbf{O}\alpha \leftarrow \beta \vee \mathbf{O}^{+?}\beta) \\
&\equiv (\mathbf{O}\alpha \leftarrow \beta) \wedge (\mathbf{O}\alpha \leftarrow \mathbf{O}\beta \wedge \neg {\sim}\beta)
\end{aligned}
$$

Semantics

Definition (HT-interpretation)

A pair $M = \langle H, T \rangle$ of sets of explicit literals where $H \subseteq T$ and T is a deontic interpretation

- We define 3 worlds r (real), o (obligation), f (forbidden) and let $\overline{r} := r$, $\overline{o} := f$ and $\overline{f} := o$

Deontic Here and There

$$
\begin{array}{lll|lll}
M, w \models \top & & & M, w \not\mathrel{=}\!\mid \top & & \\
M, w \not\models \bot & & & M, w \mathrel{=}\!\mid \bot & & \\
M, w \models \varphi \land \psi & \text{if } M, w \models \varphi & & M, w \mathrel{=}\!\mid \varphi \land \psi & \text{if } M, w \mathrel{=}\!\mid \varphi & \\
 & \text{and } M, w \models \psi & & & \text{or } M, w \mathrel{=}\!\mid \psi & \\
M, w \models \varphi \lor \psi & \text{if } M, w \models \varphi & & M, w \mathrel{=}\!\mid \varphi \lor \psi & \text{if } M, w \mathrel{=}\!\mid \varphi & \\
 & \text{or } M, w \models \psi & & & \text{and } M, w \mathrel{=}\!\mid \psi &
\end{array}
$$

$$
\begin{array}{ll}
\langle H, T \rangle, w \models \varphi \rightarrow \psi & \text{if } \langle X, T \rangle, w \not\models \varphi \text{ or } \langle X, T \rangle, w \models \psi \\
 & \text{for all } X \in \{H, T\}
\end{array}
$$

$$
\langle H, T \rangle, w \mathrel{=}\!\mid \varphi \rightarrow \psi \quad \text{if } \langle T, T \rangle \models \varphi \text{ and } \langle H, T \rangle, w \mathrel{=}\!\mid \psi
$$

Deontic Here and There

$$
\begin{array}{lll|lll}
M, w \models \sim\varphi & \text{if } M, \overline{w} \dashv \varphi & & M, w \dashv \sim\varphi & \text{if } M, \overline{w} \models \varphi \\
M, w \models \mathbf{O}\varphi & \text{if } M, o \models \varphi & & M, w \dashv \mathbf{O}\varphi & \text{if } M, o \dashv \varphi \\
M, w \models \mathbf{F}\varphi & \text{if } M, f \dashv \varphi & & M, w \dashv \mathbf{F}\varphi & \text{if } M, f \models \varphi \\
M, r \models p & \text{if } p \in H & & M, r \dashv p & \text{if } \sim p \in H \\
M, o \models p & \text{if } \mathbf{O}p \in H & & M, o \dashv p & \text{if } \sim\mathbf{O}p \in H \\
M, f \models p & \text{if } \sim\mathbf{F}p \in H & & M, f \dashv p & \text{if } \mathbf{F}p \in H
\end{array}
$$

Deontic Here and There

$$
\begin{aligned}
\mathbf{O}(\varphi \vee \psi) &\equiv \mathbf{O}\varphi \vee \mathbf{O}\psi \\
\mathbf{O}(\varphi \wedge \psi) &\equiv \mathbf{O}\varphi \wedge \mathbf{O}\psi \\
\mathbf{O}\bot &\equiv \bot \\
\mathbf{O}\top &\equiv \top \\
\mathbf{O}(\varphi \rightarrow \psi) &\equiv \mathbf{O}\varphi \rightarrow \mathbf{O}\psi \\
\mathbf{O}\neg\varphi &\equiv \neg\mathbf{O}\varphi \\
\mathbf{O}\mathbf{O}\varphi &\equiv \mathbf{O}\varphi \\
\mathbf{O}\mathbf{F}\varphi &\equiv \mathbf{F}\varphi
\end{aligned}
$$

Deontic Here and There

$$
\begin{aligned}
\mathbf{F}(\varphi \vee \psi) &\equiv \mathbf{F}\varphi \wedge \mathbf{F}\psi \\
\mathbf{F}(\varphi \wedge \psi) &\equiv \mathbf{F}\varphi \vee \mathbf{F}\psi \\
\mathbf{F}\bot &\equiv \top \\
\mathbf{F}\top &\equiv \bot \\
\mathbf{F}(\varphi \rightarrow \psi) &\equiv \neg\neg\mathbf{O}\varphi \wedge \mathbf{F}\psi \\
\mathbf{F}\neg\varphi &\equiv \neg\neg\mathbf{O}\varphi \\
\mathbf{F}\mathbf{F}\varphi &\equiv \sim\mathbf{F}\varphi \\
\mathbf{F}\mathbf{O}\varphi &\equiv \sim\mathbf{O}\varphi \\
\mathbf{O} \sim\varphi &\equiv \mathbf{F}\varphi \\
\mathbf{F} \sim\varphi &\equiv \mathbf{O}\varphi
\end{aligned}
$$

- Conjecture: deontic logic programs constitute a normal form

Deontic Equilibrium Logic

Definition (Deontic Equilibrium Model)

$\langle T, T \rangle$ is a deontic equilibrium model of α if $\langle T, T \rangle, r \models \alpha$ and there is no $H \subseteq T$ s.t. $\langle H, T \rangle, r \models \alpha$.

Theorem

For program Π: $\langle T, T \rangle$ is an equilibrium model iff T is an answer set.

Chisholm's paradox

Ex6 *You ought to go to the assistance of your neighbours*
If you do so, you ought to let them know you are coming
If you do not go, you ought to not tell them you are coming
You are not going to the assistance of your neighbours

$$\begin{aligned} \mathbf{O} g \\ \mathbf{O}(t \mid g) \\ \mathbf{O}(\sim t \mid \sim g) \\ \sim g \end{aligned}$$

equivalent to ...

Ex6 *You ought to go to the assistance of your neighbours*
If you do so, you ought to let them know you are coming
If you do not go, you ought to not tell them you are coming
You are not going to the assistance of your neighbours

$$
\begin{array}{rcl}
\mathbf{O}g & & \\
\mathbf{O}t & \leftarrow & g \\
\mathbf{O}t & \leftarrow & \mathbf{O}g \wedge \neg \sim g \\
\mathbf{F}t & \leftarrow & \sim g \\
\mathbf{F}t & \leftarrow & \mathbf{F}g \wedge \neg g
\end{array}
$$

$\sim g$ ~~$\sim g$~~

We get answer set $T = \{\sim g, \mathbf{O}g, \mathbf{F}t\}$ (factual detachment)
if we remove $\sim g$ we get $T = \{\mathbf{O}g, \mathbf{O}t\}$ (deontic detachment)

Considerate murder

Ex7 *You should not kill the witness*
If you kill the witness, you should offer him a cigarette
You should not offer a cigarette
You kill the witness

$$\begin{array}{rcl} \mathbf{F}k & & \\ \mathbf{O}c & \leftarrow & \mathbf{F}k \wedge k \\ \mathbf{F}c & & \\ k & & \end{array}$$

No answer set: we would derive $\mathbf{F}c$ and $\mathbf{O}c$ but no decision about c.
If we add practical reasoning (obligations are normally fulfilled)

$$\begin{array}{rcl} c & \leftarrow & \mathbf{O}c \wedge \neg \sim c \\ \sim c & \leftarrow & \mathbf{F}c \wedge \neg c \end{array}$$

☞ 2 answer sets: $\{\mathbf{F}k, k, \mathbf{O}c, \mathbf{F}c, c\}$, $\{\mathbf{F}k, k, \mathbf{O}c, \mathbf{F}c, \sim c\}$.

Conclusions

- On the ASP level, simple extension = deontic atoms + a constraint
- As a full logic, nesting is allowed, but expressive power limited to literals (3-valued logic). Ex: $\mathbf{O}(p \vee q) \equiv \mathbf{O}p \vee \mathbf{O}q$
- Still, it offers many subtle ways of representing obligations (defaults, implicit vs explicit permissions, etc)

Future work:

- Implementation: theory atoms for deontic atoms
  `&ob{pay}, &fb{park}`
  Even for derived and nested (non-deontic) operators
  `&ob{-fence}, &ob{tell / go}, &viol{-fence} .`
- Many scenarios require future extension to Deontic Temporal Equilibrium Model
- Combination with explanations (`xclingo`)

A Preliminary Study on Deontic Answer Set Programming

Pedro Cabalar[1], Agata Ciabattoni[2], Leon van der Torre[3]

[1]University of Corunna, SPAIN
[2]Technical University of Vienna, AUSTRIA
[3]University of Luxembourg, LUXEMBOURG

November 3rd, 2022
Cercedilla, Spain

¡Gracias por la atención!