Logics of agency

Chapter 2: Propositional Dynamic Logic and Theory of Action

Nicolas Troquard

ESSLLI 2016 – Bolzano

- Most of this course will consider logics of agency “proper”.
- They generally abstract away from action names.
- Different from Davidson’s treatment.
- Different from Dynamic Logic (PDL and variants).
- This chapter thus briefly covers PDL and its applications to theories of action.

Propositional Dynamic Logic (PDL) [Hoare 1969], [Pratt 1976], [Harel et al. 2001]:

Language:

- names for atomic events
- complex events are built recursively by means of imperative programming constructs
  - ";" (sequential composition),
  - "∪" (nondeterministic composition),
  - "*" (iteration),
  - "?" (test), ...

Example: the event of "felling a tree by performing the atomic 'chop' action until the tree is down":

$$\pi_{\mathit{fellTree}} = (\neg \mathit{treeDown}?; \mathbf{chop})^*; \mathit{treeDown}?$$

The event “felling a tree”:

$$\pi_{\mathit{fellTree}} = (\neg \mathit{treeDown}?; \mathbf{chop})^*; \mathit{treeDown}?$$

Equivalent:

```
1: if (not treeDown)
2:     chop
3:     goto 1
4: else
5:     done
```

Equivalent:

```
while (not treeDown)
  chop
```

- if φ then α else $\beta =_{\textit{def}} ((\varphi?;\alpha) \cup (\neg\varphi?;\beta))$
- while φ do $\alpha =_{\textit{def}} ((\varphi?;\alpha)^*;\neg\varphi?)$
- repeat α until $\varphi =_{\textit{def}} (\alpha;((\varphi?;\alpha)^*;\neg\varphi?))$
- abort $=_{\textit{def}} \bot?$
- skip $=_{\textit{def}} \top?$

Application to the theory of action:

- [Cohen & Levesque 1990]
- [van Linder et al. 96-99] (KARO)

Mental attitudes, time, ..., intention.

Outline

1. Propositional Dynamic Logic

2. Theory of intentional action

Language of PDL

Terms:

- atomic propositions Prop $= \{p, q, r, \ldots, p_1, p_2, \ldots\}$
- atomic events Evt $= \{\alpha, \beta, \ldots, \alpha_1, \alpha_2, \ldots\}$

In general, an event has the form:

$$\pi ::= \alpha \mid \varphi? \mid \pi;\pi \mid \pi \cup \pi \mid \pi^*$$

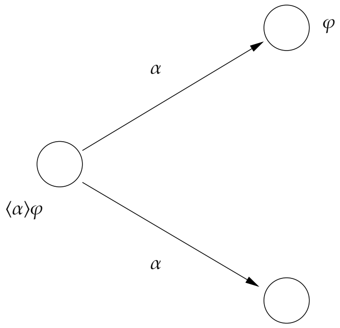
where φ is a proposition (see next), and $\alpha \in$ Evt.

The language (the set of all propositions / well-formed sentences) is given by the grammar:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle\pi\rangle\varphi$$

where $p \in$ Prop and π is an event.

$\langle\alpha\rangle\varphi$ means that "there is a possible course of event α that yields the proposition φ".



A PDL model is a tuple $\langle W, R, V \rangle$:

- W: a set of possible worlds
- R: where $R_\alpha \subseteq W \times W$ for all $\alpha \in$ Evt
- V: is a valuation function $V(p) \subseteq W$ for all $p \in$ Prop

$(w, v) \in R_\alpha$: "there is a course of α from w that ends in v"

- We can extend R naturally to general events π: $R_\pi \subseteq W \times W$.
- We can extend V naturally to general propositions φ:
  - $w \models p$ iff $w \in V(p)$
  - $w \models \langle\pi\rangle\varphi$ iff there is $v \in W$ such that $(w, v) \in R_\pi$ and $v \models \varphi$

Necessity:

$$[\pi]\varphi = \neg\langle\pi\rangle\neg\varphi$$

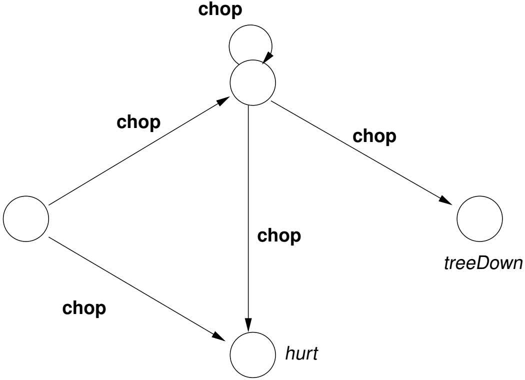
Possible execution:

$$\langle\pi\rangle\top$$

Impossible execution:

$$[\pi]\bot$$

$\pi_{fellTree} = (\neg treeDown?; \textbf{chop})^*; treeDown?$



In the left-most world:

1. $\langle \textbf{chop}; \textbf{chop} \rangle treeDown$
2. $\neg \langle \textbf{chop} \rangle treeDown$
3. $\langle \textbf{chop}; \textbf{chop}; \textbf{chop} \rangle hurt$
4. $\langle \pi_{fellTree} \rangle \top$
5. $\langle \pi_{fellTree} \rangle treeDown \wedge [\pi_{fellTree}] treeDown$
6. $\langle \pi_{fellTree} \rangle \neg treeDown$ never holds

- *hurt*? is a state event
- **chop*** is an activity event
- π_{fellTree} is an accomplishment event that necessarily culminates in world satisfying *treeDown*

Outline

[Bratman 1986]

- Intentions pose problems for the agent; the agent needs to determine a way to achieve them.
- Intentions provide a “screen of admissibility” for adopting other intentions.
- Agents “track” the success of their attempts to achieve intentions.

[Cohen & Levesque 1987, 1990] have adapted PDL models to capture actual action and intention.

Intention is choice with commitment: intention is a composite specifying what an agent choose and is commited to.

(We use the slightly simplified presentation of [Herzig & Longin 2004] [Meyer, Broersen, Herzig 2012].)

PDL SPECIALIZATION

Linear PDL:

- a world has at most one immediate successor
- more than one atomic action can yield the transition
- formally: if $(u, v_1) \in R_\alpha$ and $(u, v_2) \in R_\beta$ then $v_1 = v_2$

In this context, we use:

- $Happ_\pi\varphi =_{def} \langle\pi\rangle\varphi$
- $(IfHapp_\pi\varphi =_{def} \neg Happ_\pi\neg\varphi = [\pi]\varphi)$

We let Agt to be a set of agents.

An atomic event is now an object referring to an agent and a action he does. E.g., *i* doing α:

$$i : \alpha$$

Quantification over actions:

$$\exists \alpha Happ_{i:\alpha} \varphi$$

Future tense modalities (textbook: [Baier, Katoen 2008]):

$$\varphi U \psi \quad / \quad F\varphi \quad / \quad G\varphi$$

Beliefs:

$$Bel_i \varphi$$

Choices/Realistic Preferences (originally *Goal*):

$$Choose_i \varphi$$

i can make φ happen:

$$\exists\alpha Happ_{i:\alpha}\varphi$$

φ is true until ψ true / φ is true eventually / φ is always true:

$$\varphi U\psi \quad / \quad F\varphi \quad / \quad G\varphi$$

i believes φ:

$$Bel_i\varphi$$

i chooses φ to be true (realistic preference):

$$Choose_i\varphi$$

For every atomic event (action) $i : \alpha$ and $j : \beta$:

$$(Happ_{i:\alpha}\varphi \land Happ_{j:\beta}\top) \rightarrow Happ_{i:\beta}\varphi$$

- $Bel_i\varphi \rightarrow \neg Bel_i\neg\varphi$ (consistency, axiom D)
- $Bel_i\varphi \rightarrow Bel_iBel_i\varphi$ (positive introspection, axiom 4)
- $\neg Bel_i\varphi \rightarrow Bel_i\neg Bel_i\varphi$ (negative introspection, axiom 5)

- $Bel_i\varphi \rightarrow Choose_i\varphi$ (realism)

(We did not give the semantic constraints of Bel_i and $Choose_i$; they are standard in modal logics: R_{Bel_i} is serial, transitive, Euclidean. $R_{Choose_i} \subseteq R_{Bel_i}$.)

Achievement goal:

$$AGoal_i\varphi = Choose_i F\varphi \wedge Bel_i\neg\varphi$$

Persistent goal:

$$PGoal_i\varphi = AGoal_i\varphi \wedge (AGoal_i\varphi) U (Bel_i\varphi \vee Bel_i G\neg\varphi \vee \psi)$$

(ψ is a “superior” reason for abandoning the goal. “Mom told me I shouldn’t do it.” Not present in [Herzig & Longin 2004].)

Intention:

$$Intend_i\varphi = PGoal_i\varphi \wedge Bel_i F\exists\alpha Happ_{i:\alpha}\varphi$$

Valid principles

if $\models \varphi \leftrightarrow \psi$ then $\models \textit{Intend}_i\varphi \leftrightarrow \textit{Intend}_i\psi$

$\models \textit{Intend}_i\varphi \rightarrow \textit{Bel}_i\neg\varphi$

...

Nice:

$$\not\models \textit{Intend}_i(\varphi \wedge \psi) \rightarrow \textit{Intend}_i\varphi \wedge \textit{Intend}_i\psi$$

$$\not\models \textit{Intend}_i\varphi \wedge \textit{Intend}_i\psi \rightarrow \textit{Intend}_i(\varphi \wedge \psi)$$

Nice:

$$\not\models (\textit{Intend}_i\varphi \wedge \textit{Bel}_i(\varphi \rightarrow \psi)) \rightarrow \textit{Intend}_i\psi$$

(e.g., if I intend to go to the dentist and believe that going to the dentist will cause pain then I do not necessarily intend to have pain)

Not very nice:

$$\not\models \textit{Intend}_i\varphi \rightarrow \textit{Bel}_i\textit{Intend}_i\varphi$$

$$\not\models \neg\textit{Intend}_i\varphi \rightarrow \textit{Bel}_i\neg\textit{Intend}_i\varphi$$

Variant: obtaining introspective intention

Let us program introspective choices into the logic:

- $Choose_i\varphi \rightarrow Bel_i Choose_i\varphi$
- $\neg Choose_i\varphi \rightarrow Bel_i \neg Choose_i\varphi$

These become valid:

- $Intend_i\varphi \rightarrow Bel_i Intend_i\varphi$
- $\neg Intend_i\varphi \rightarrow Bel_i \neg Intend_i\varphi$

References I

Christel Baier and Joost-Pieter Katoen.
Principles of model checking.
MIT Press, 2008.

M. Bratman.
Intentions, plans, and practical reasons.
Harvard University Press, 1986.

Philip R. Cohen and Hector J. Levesque.
Intention is choice with commitment.
Artif. Intell., 42(2-3):213–261, 1990.

D. Harel, D. Kozen, and J. Tiuryn.
Dynamic Logic.
MIT Press, 2000.

Andreas Herzig and Dominique Longin.
C&L Intention Revisited.
In Proc. Principles of Knowledge Representation and Reasoning: Proceedings of the Ninth International Conference (KR2004), pages 527–535, 2004.

C. Hoare.
An axiomatic basis for computer programming.
Communications of the Association of Computing Machinery, 12:576–580, 1969.

J.J. Ch Meyer, J. Broersen, and A. Herzig.
BDI Logics.
In Handbook of Epistemic Logic, chapter 10. College Publications, 2015.

References II

V. Pratt.
Semantical considerations on floyd-hoare logic.
In Proceedings of the 17th IEEE Symposium on Foundations of Computer Science, pages 109–121. IEEE Computer Society, 1976.

Nicolas Troquard and Philippe Balbiani.
Propositional dynamic logic.
In Edward N. Zalta, editor, The Stanford Encyclopedia of Philosophy. Spring 2015 edition, 2015.

W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer.
An integrated modal approach to rational agents.
In M. Wooldridge and A. Rao, editors, Foundations of Rational Agency, volume 14 of Applied Logic Series, pages 133–168. Kluwer.