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”:

\[ \pi_{\text{fellTree}} = (\neg \text{treeDown} ; \text{chop})^* ; \text{treeDown} ? \]
The event “felling a tree”:

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

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

Equivalent:
while (not treeDown)
  chop
More imperative programming

- if $\varphi$ then $\alpha$ else $\beta = \text{def} \ ((\varphi?; \alpha) \cup (\neg \varphi?; \beta))$
- while $\varphi$ do $\alpha = \text{def} \ ((\varphi?; \alpha^*; \neg \varphi?)$
- repeat $\alpha$ until $\varphi = \text{def} \ (\alpha; ((\varphi?; \alpha^*; \neg \varphi?)$
- abort $= \text{def} \ \bot$
- skip $= \text{def} \ \top$
Theories of action

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

In general, an event has the form:

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

where \( \varphi \) is a proposition (see next), and \( \alpha \in \text{Evt} \).

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

\[
\varphi ::= p \mid \neg \varphi \mid \varphi \lor \varphi \mid \langle \pi \rangle \varphi
\]

where \( p \in \text{Prop} \) and \( \pi \) is an event.
\langle \alpha \rangle \varphi \text{ means that “there is a possible course of event } \alpha \text{ that yields the proposition } \varphi \text{”}.
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 $\alpha \in \text{Evt}$
- $V$: is a valuation function $V(p) \subseteq W$ for all $p \in \text{Prop}$

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

- We can extend $R$ naturally to general events $\pi$:
  $R_\pi \subseteq W \times W$.
- We can extend $V$ naturally to general propositions $\varphi$:
  - $w \models p$ 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:
\[ [\pi] \bot \]
\[ \pi_{\text{fellTree}} = (\neg \text{treeDown}?; \text{chop})^*; \text{treeDown}? \]

In the left-most world:

1. \( \langle \text{chop}; \text{chop} \rangle \text{treeDown} \)
2. \( \neg \langle \text{chop} \rangle \text{treeDown} \)
3. \( \langle \text{chop}; \text{chop}; \text{chop} \rangle \text{hurt} \)
4. \( \langle \pi_{\text{fellTree}} \rangle \top \)
5. \( \langle \pi_{\text{fellTree}} \rangle \text{treeDown} \land [\pi_{\text{fellTree}}] \text{treeDown} \)
6. \( \langle \pi_{\text{fellTree}} \rangle \neg \text{treeDown} \text{ never holds} \)
Types of events

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

1 Propositional Dynamic Logic

2 Theory of intentional action
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 *committed* to.

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

- \(\text{Happ}_\pi \varphi =_{\text{def}} \langle \pi \rangle \varphi\)
- \((\text{IfHapp}_\pi \varphi =_{\text{def}} \neg \text{Happ}_\pi \neg \varphi = [\pi] \varphi)\)
**Extra Language**

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 \(\alpha\):

\[i : \alpha\]

Quantification over actions:

\[\exists \alpha \text{Happ}_{i:\alpha} \varphi\]

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

\[\varphi U \varphi \quad / \quad F \varphi \quad / \quad G \varphi\]

Beliefs:

\[\text{Bel}_{i\varphi}\]

Choices/Realistic Preferences (originally Goal):

\[\text{Choose}_{i\varphi}\]
i can make $\varphi$ happen:

$$\exists \alpha \text{Happ}_{i: \alpha} \varphi$$

$\varphi$ is true until $\psi$ true / $\varphi$ is true eventually / $\varphi$ is always true:

$$\varphi \text{U}\psi \quad / \quad F\varphi \quad / \quad G\varphi$$

i believes $\varphi$:

$$\text{Bel}_{i}\varphi$$

i chooses $\varphi$ to be true (realistic preference):

$$\text{Choose}_{i}\varphi$$
PRINCIPLES OF INTENTIONAL ACTION

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

$$(\text{Happ}_{i:\alpha} \varphi \land \text{Happ}_{j:\beta} \top) \rightarrow \text{Happ}_{i:\beta} \varphi$$

- $\text{Bel}_{i} \varphi \rightarrow \neg \text{Bel}_{i} \neg \varphi$ (consistency, axiom D)
- $\text{Bel}_{i} \varphi \rightarrow \text{Bel}_{i} \text{Bel}_{i} \varphi$ (positive introspection, axiom 4)
- $\neg \text{Bel}_{i} \varphi \rightarrow \text{Bel}_{i} \neg \text{Bel}_{i} \varphi$ (negative introspection, axiom 5)

- $\text{Bel}_{i} \varphi \rightarrow \text{Choose}_{i} \varphi$ (realism)

(We did not give the semantic constraints of $\text{Bel}_{i}$ and $\text{Choose}_{i}$; they are standard in modal logics: $R_{\text{Bel}_{i}}$ is serial, transitive, Euclidean. $R_{\text{Choose}_{i}} \subseteq R_{\text{Bel}_{i}}$.)
Intention is choice with commitment

Achievement goal:

\[ A\text{Goal}_i\varphi = \text{Choose}_iF\varphi \land B\text{el}_i\lnot\varphi \]

Persistent goal:

\[ P\text{Goal}_i\varphi = A\text{Goal}_i\varphi \land (A\text{Goal}_i\varphi)U(B\text{el}_i\varphi \lor B\text{el}_i\lnot\varphi \lor \psi) \]

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

Intention:

\[ \text{Intend}_i\varphi = P\text{Goal}_i\varphi \land B\text{el}_iF\exists\alpha\text{Happ}_i:\alpha\varphi \]
Valid principles

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

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

...
Invalid principles

Nice:

\[ \neg \text{Intend}_i(\varphi \land \psi) \rightarrow \text{Intend}_i\varphi \land \text{Intend}_i\psi \]

\[ \neg \text{Intend}_i\varphi \land \text{Intend}_i\psi \rightarrow \text{Intend}_i(\varphi \land \psi) \]

Nice:

\[ \neg (\text{Intend}_i\varphi \land \text{Bel}_i(\varphi \rightarrow \psi)) \rightarrow \text{Intend}_i\psi \]

(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:

\[ \neg \text{Intend}_i\varphi \rightarrow \text{Bel}_i\text{Intend}_i\varphi \]

\[ \neg \neg \text{Intend}_i\varphi \rightarrow \text{Bel}_i\neg \text{Intend}_i\varphi \]
**Variant: obtaining introspective intention**

Let us program introspective choices into the logic:

- \( \text{Choose}_i \varphi \rightarrow \text{Bel}_i \text{Choose}_i \varphi \)
- \( \neg \text{Choose}_i \varphi \rightarrow \text{Bel}_i \neg \text{Choose}_i \varphi \)

These become valid:

- \( \text{Intend}_i \varphi \rightarrow \text{Bel}_i \text{Intend}_i \varphi \)
- \( \neg \text{Intend}_i \varphi \rightarrow \text{Bel}_i \neg \text{Intend}_i \varphi \)
Christel Baier and Joost-Pieter Katoen.  
*Principles of model checking.* 

M. Bratman.  
*Intentions, plans, and practical reasons.*  

Philip R. Cohen and Hector J. Levesque.  
*Intention is choice with commitment.*  

D. Harel, D. Kozen, and J. Tiuryn.  
*Dynamic Logic.*  

Andreas Herzig and Dominique Longin.  
*C&L Intention Revisited.*  

C. Hoare.  
*An axiomatic basis for computer programming.*  

J.J. Ch Meyer, J. Broersen, and A. Herzig.  
*BDI Logics.*  
V. Pratt.  
Semantical considerations on floyd-hoare logic.  

Nicolas Troquard and Philippe Balbiani.  
Propositional dynamic logic.  

W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer.  
An integrated modal approach to rational agents.  