# Action Theory Contraction and Minimal Change

**Ivan José Varzinczak**

IRIT – Université de Toulouse
Toulouse, France
ivan.varzinczak@irit.fr

Meraka Institute
CSIR, Pretoria, South Africa
ivan.varzinczak@meraka.org.za

**Abstract**

This work is about changing action domain descriptions in dynamic logic. We here revisit the semantics of action theory contraction, giving more robust operators that express minimal change based on a notion of distance between models. We then define syntactical contraction operators and establish their correctness w.r.t. our semantics. Finally we show that our operators satisfy the PDL-counterpart of the standard postulates for theory change adopted in the literature.

## Introduction and Motivation

Let an intelligent agent be designed to perform rationally in a dynamic world, and suppose she should reason about the dynamics of an automatic coffee machine. Suppose that the agent believes that a coffee is a hot beverage. Now suppose that some day she gets a coffee and observes it is cold. In such a case, the agent must change her beliefs about the relationship between the propositions "I have a coffee" and "I have a hot beverage". This example is an instance of the problem of changing propositional belief bases and is largely addressed in the literature about belief change (Gärdenfors 1988) and belief update (Katsuno and Mendelzon 1992).

Next, let our agent believe that whenever buying a coffee on the machine, she gets a hot beverage. This means that in every state of the world that follows the execution of buying, the agent possesses a hot beverage. Some day, it may happen that the machine is running out of cups, and then after buying, the coffee runs through the shelf and so the agent holds no hot beverage.

Imagine now the agent believes that if she has a token, then it is always possible to buy coffee. However, during a blackout, even with a token the agent does not manage to order her coffee on the machine.

The last two examples illustrate situations where changing the beliefs about the behavior of the action of buying coffee is mandatory. In the first one, buying coffee, once believed to be deterministic, has now to be seen as nondeterministic, or alternatively to have a different outcome in a more specific context (e.g. if there is no cup in the machine). In the second example, the executability of the action under concern is questioned in the light of new information showing a context that was not known to preclude its execution.

Such cases of theory change are very important in logical descriptions of dynamic domains: it may always happen that one discovers that an action actually has a behavior that is different from that one has always believed it had.

Up to now, theory change has been studied mainly for knowledge bases in classical logics, both in terms of revision and update. Since (Fuhrmann 1989), only in a few recent works it has been considered in the realm of modal logics, viz. in epistemic logic (Hansson 1999) and in dynamic logics (Herzig, Perrussel, and Varzinczak 2006), and in action languages (Eiter et al. 2005). Some other works (Shapiro et al. 2000; Jin and Thielscher 2005) have investigated revision of beliefs about *facts* of the world. In our scenario, this would concern for example the truth of *token* in a given state: the agent believes she has a token, but is wrong about that and might subsequently be forced to revise her beliefs about the current state of affairs. Such belief revision operations do not modify the agent's beliefs about the *action laws*. In opposition to that, here we are interested exactly in such modifications.

## Logical Preliminaries

Our base formalism is Propositional Dynamic Logic (PDL) without the $*$ operator (Harel, Tiuryn, and Kozen 2000).

### Action Theories in Dynamic Logic

Let $\mathfrak{Act} = \{a_1, a_2, \ldots\}$ be the set of *atomic actions* of a given domain. An example of atomic action is *buy*. To each action $a$ there is associated a modal operator $[a]$. $\mathfrak{Prop} = \{p_1, p_2, \ldots\}$ denotes the set of *propositional constants*, also called *fluents* or *atoms*. Examples of those are *token* ("the agent has a token") and *coffee* ("the agent holds a coffee"). The set of all literals is $\mathfrak{Lit} = \{\ell_1, \ell_2, \ldots\}$, where each $\ell_i$ is either $p$ or $\neg p$, for some $p \in \mathfrak{Prop}$. If $\ell = \neg p$, then we identify $\neg\ell$ with $p$. By $|\ell|$ we denote the atom in $\ell$.

We use $\varphi, \psi, \ldots$ to denote *Boolean formulas*, an example of which is *coffee* $\rightarrow$ *hot*. $\mathfrak{Fml}$ is the set of all Boolean formulas. A propositional valuation $v$ is a *maximally consistent* set of literals. We denote by $v \Vdash \varphi$ the fact that $v$ satisfies $\varphi$. By $val(\varphi)$ we denote the set of all valuations satisfying $\varphi$.

$\models_{\mathsf{CPL}}$ is the classical consequence relation. $Cn(\varphi)$ denotes all logical consequences of $\varphi$ in classical propositional logic.

With $IP(\varphi)$ we denote the set of *prime implicants* (Quine 1952) of $\varphi$. By $\pi$ we denote a prime implicant, and $atm(\pi)$ is the set of atoms occurring in $\pi$. For given $\ell$ and $\pi$, $\ell \in \pi$ abbreviates '$\ell$ is a literal of $\pi$'.

We will use $\Phi, \Psi, \dots$ to denote complex formulas (formulas with modal operators). An example of a complex formula is $\neg coffee \rightarrow [buy]coffee$. $\langle a \rangle$ is the dual operator of $[a]$ ($\langle a \rangle\Phi =_{\text{def}} \neg[a]\neg\Phi$).

A $\mathsf{PDL}$-*model* is a tuple $\mathscr{M} = \langle W, R \rangle$ where $W$ is a set of valuations, and $R$ maps action constants $a$ to accessibility relations $R_a \subseteq W \times W$. Given $\mathscr{M}$, $\models^{\mathscr{M}}_{w} p$ ($p$ is true at world $w$ of model $\mathscr{M}$) if $w \Vdash p$; $\models^{\mathscr{M}}_{w} [a]\Phi$ if $\models^{\mathscr{M}}_{w'} \Phi$ for every $w'$ s.t. $(w, w') \in R_a$; truth conditions for the other connectives are as usual. By $\mathcal{M}$ we will denote a set of $\mathsf{PDL}$-models.

$\mathscr{M}$ is a model of $\Phi$ (noted $\models^{\mathscr{M}} \Phi$) if and only if $\models^{\mathscr{M}}_{w} \Phi$ for all $w \in W$. $\mathscr{M}$ is a model of a set of formulas $\Sigma$ (noted $\models^{\mathscr{M}} \Sigma$) if and only if $\models^{\mathscr{M}} \Phi$ for every $\Phi \in \Sigma$. $\Phi$ is a *consequence of the global axioms* $\Sigma$ in all $\mathsf{PDL}$-models (noted $\Sigma \models_{\mathsf{PDL}} \Phi$) if and only if for every $\mathscr{M}$, if $\models^{\mathscr{M}} \Sigma$, then $\models^{\mathscr{M}} \Phi$.

With $\mathsf{PDL}$ we can state laws describing the behavior of actions. Following the tradition in the reasoning about actions community, we here distinguish three types of them.

**Static Laws** A *static law* is a formula $\varphi \in \mathfrak{Fml}$. It is a formula that characterizes the possible states of the world. An example of static law is *coffee* $\rightarrow$ *hot*: if the agent holds a coffee, then she holds a hot beverage. The set of all static laws of a domain is denoted by $\mathcal{S}$.

**Effect Laws** An *effect law for a* is of the form $\varphi \rightarrow [a]\psi$, where $\varphi, \psi \in \mathfrak{Fml}$. Effect laws are formulas relating an action to its effects, which can be conditional. The consequent $\psi$ is the effect that always obtains when $a$ is executed in a state where the antecedent $\varphi$ holds. If $a$ is a nondeterministic action, then $\psi$ is typically a disjunction. An example of such a law is $token \rightarrow [buy]hot$: whenever the agent has a token, after buying, she has a hot beverage. If $\psi$ is inconsistent we have a special kind of effect law that we call an *inexecutability law*. For example, $\neg token \rightarrow [buy]\bot$ says that *buy* cannot be executed if the agent has no token. The set of effect laws of a domain is denoted by $\mathcal{E}$.

**Executability Laws** An *executability law for a* has the form $\varphi \rightarrow \langle a \rangle\top$, with $\varphi \in \mathfrak{Fml}$. It stipulates the context in which $a$ is guaranteed to be executable. (In $\mathsf{PDL}$, the operator $\langle a \rangle$ is used to express executability, $\langle a \rangle\top$ thus reads "$a$'s execution is possible".) For instance, $token \rightarrow \langle buy \rangle\top$ says that buying can be executed whenever the agent has a token. The set of all executability laws of a domain is denoted by $\mathcal{X}$.

Given $a$, $\mathcal{E}_a$ (resp. $\mathcal{X}_a$) will denote the set of only those effect (resp. executability) laws about $a$.

**Action Theories** $\mathcal{T} = \mathcal{S} \cup \mathcal{E} \cup \mathcal{X}$ is an *action theory*.

For the sake of clarity, we will here abstract from the frame problem (McCarthy and Hayes 1969) and the ramification problem (Finger 1987), and assume that the agent's theory contains all frame axioms (cf. (Herzig, Perrussel, and Varzinczak 2006) for a contraction approach within a solution to the frame problem). The action theory of our example will be:

$$\mathcal{T} = \left\{ \begin{array}{c} coffee \rightarrow hot, token \rightarrow \langle buy \rangle\top, \\ \neg coffee \rightarrow [buy]coffee, token \rightarrow [buy]\neg token, \\ \neg token \rightarrow [buy]\bot, \neg token \rightarrow [buy]\neg token, \\ coffee \rightarrow [buy]coffee, hot \rightarrow [buy]hot \end{array} \right\}$$

Figure 1 below shows a $\mathsf{PDL}$-model for the theory $\mathcal{T}$.

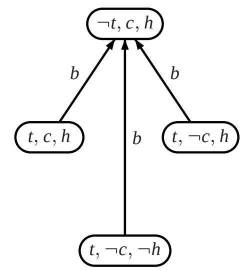


Figure 1: A model for the coffee machine scenario. *b*, *t*, *c*, and *h* stand for, respectively, *buy*, *token*, *coffee*, and *hot*.

Sometimes it will be useful to consider models whose possible worlds are *all* the possible worlds allowed by $\mathcal{S}$:

**Definition 1** *Let $\mathcal{T} = \mathcal{S} \cup \mathcal{E} \cup \mathcal{X}$ be an action theory. Then $\mathscr{M} = \langle W, R \rangle$ is the* big model *of $\mathcal{T}$ if and only if:*

- $W = val(\mathcal{S})$*; and*
- $R_a = \{(w, w') : \forall.\varphi \rightarrow [a]\psi \in \mathcal{E}_a$, *if* $\models^{\mathscr{M}}_{w} \varphi$ *then* $\models^{\mathscr{M}}_{w'} \psi\}$.

Figure 2 below depicts the big model of $\mathcal{T}$.

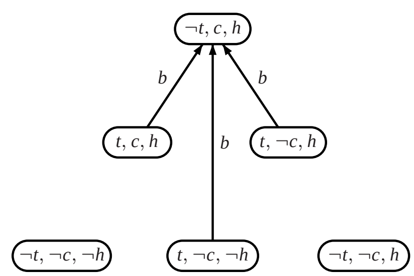


Figure 2: The big model for the coffee machine scenario.

## Elementary Atoms

Given $\varphi \in \mathfrak{Fml}$, $E(\varphi)$ denotes the elementary atoms *actually* occurring in $\varphi$. For example, $E(\neg p_1 \wedge (\neg p_1 \vee p_2)) = \{p_1, p_2\}$. An atom $p$ is *essential* to $\varphi$ if and only if $p \in E(\varphi')$ for all $\varphi'$ such that $\models_{\mathsf{CPL}} \varphi \leftrightarrow \varphi'$. For instance, $p_1$ is essential to $\neg p_1 \wedge (\neg p_1 \vee p_2)$. $E!(\varphi)$ will denote the essential atoms of $\varphi$. (If $\varphi$ is a tautology or a contradiction, then $E!(\varphi) = \emptyset$.)

For $\varphi \in \mathfrak{Fml}$, $\varphi*$ is the set of all $\varphi' \in \mathfrak{Fml}$ such that $\varphi \models_{\mathsf{CPL}} \varphi'$ and $E(\varphi') \subseteq E!(\varphi)$. For instance, $p_1 \vee p_2 \notin$

$p_1*$, as $p_1 \models_{\mathsf{CPL}} p_1 \vee p_2$ but $E(p_1 \vee p_2) \not\subseteq E!(p_1)$. Clearly, $E(\bigwedge \varphi*) = E!(\bigwedge \varphi*)$. Moreover, whenever $\models_{\mathsf{CPL}} \varphi \leftrightarrow \varphi'$, then $E!(\varphi) = E!(\varphi')$ and also $\varphi* = \varphi'*$.

**Theorem 1 (Least atom-set theorem (Parikh 1999))** *$\models_{\mathsf{CPL}} \varphi \leftrightarrow \bigwedge \varphi*$, and $E(\varphi*) \subseteq E(\varphi')$ for every $\varphi'$ s.t. $\models_{\mathsf{CPL}} \varphi \leftrightarrow \varphi'$.*

Thus for every $\varphi \in \mathfrak{Fml}$ there is a unique least set of elementary atoms such that $\varphi$ may equivalently be expressed using only atoms from that set.[1] Hence, $Cn(\varphi) = Cn(\varphi*)$.

## Prime Valuations

Given a valuation $v$, $v' \subseteq v$ is a *subvaluation*. For $W$ a set of valuations, a subvaluation $v'$ *satisfies* $\varphi \in \mathfrak{Fml}$ modulo $W$ (noted $v' \Vdash_W \varphi$) if and only if $v \Vdash \varphi$ for all $v \in W$ such that $v' \subseteq v$. A subvaluation $v$ *essentially satisfies* $\varphi$ modulo $W$ ($v \Vdash^!_W \varphi$) if and only if $v \Vdash_W \varphi$ and $\{|\ell| : \ell \in v\} \subseteq E!(\varphi)$.

**Definition 2** *Let $\varphi \in \mathfrak{Fml}$ and $W$ be a set of valuations. A subvaluation $v$ is a* prime subvaluation *of $\varphi$ (modulo $W$) if and only if $v \Vdash^!_W \varphi$ and there is no $v' \subseteq v$ s.t. $v' \Vdash^!_W \varphi$.*

A prime subvaluation of a formula $\varphi$ is one of the weakest states of truth in which $\varphi$ is true. (Notice the similarity with the syntactical notion of prime implicant (Quine 1952).)

By $base(\varphi, W)$ we denote the set of all prime subvaluations of $\varphi$ modulo $W$.

**Theorem 2** *Let $\varphi \in \mathfrak{Fml}$ and $W$ be a set of valuations. Then for all $w \in W$, $w \Vdash \varphi$ if and only if $w \Vdash \bigvee_{v \in base(\varphi,W)} \bigwedge_{\ell \in v} \ell$.*

## Closeness Between Models

When contracting a formula from a model, we will perform a change in its structure. Because there can be several different ways of modifying a model (not all of them minimal), we need a notion of distance between models to identify those that are closest to the original one.

As we are going to see in more depth in the sequel, changing a model amounts to modifying its possible worlds or its accessibility relation. Hence, the distance between two PDL-models will depend upon the distance between their sets of worlds and accessibility relations. These here will be based on the *symmetric difference* between sets, defined as $X \dot{-} Y = (X \setminus Y) \cup (Y \setminus X)$.

**Definition 3** *Let $\mathscr{M} = \langle W, R\rangle$ be a model. $\mathscr{M}' = \langle W', R'\rangle$ is* as close to *$\mathscr{M}$ as $\mathscr{M}'' = \langle W'', R''\rangle$, noted $\mathscr{M}' \preceq_{\mathscr{M}} \mathscr{M}''$, if and only if*

- *either $W \dot{-} W' \subseteq W \dot{-} W''$*
- *or $W \dot{-} W' = W \dot{-} W''$ and $R \dot{-} R' \subseteq R \dot{-} R''$*

(Notice that other distance notions are also possible, like e.g. considering the *cardinality* of symmetric differences.)

[1] The dual notion (redundant atoms) is also addressed in the literature, e.g. in (Herzig and Rifi 1999), with similar purposes.

## Semantics of Contraction

When contracting a law $\Phi$ we must ensure that $\Phi$ becomes invalid in at least one (possibly new) model of the dynamic domain. Because there can be lots of models to consider, we start with a *set* $\mathcal{M}$ of models in which $\Phi$ is (potentially) valid. Thus contracting $\Phi$ amounts to make it no longer valid in this set of models. What are the operations that must be carried out to achieve that? Throwing models out of $\mathcal{M}$ does not work, since $\Phi$ will keep on being valid in all models of the remaining set. Thus we should *add* new models to $\mathcal{M}$. Which models? Well, models in which $\Phi$ is not true. But not any of such models: taking models falsifying $\Phi$ that are too different from our original models will certainly violate minimal change.

Hence, we shall take some model $\mathscr{M} \in \mathcal{M}$ as basis and manipulate it to get a new model $\mathscr{M}'$ in which $\Phi$ is not true. In PDL, the removal of a law $\Phi$ from a model $\mathscr{M}$ amounts to modifying the possible worlds or the accessibility relation in $\mathscr{M}$ so that $\Phi$ becomes false. Such an operation gives as result a *set* $\mathscr{M}^-_\Phi$ of models, each of which is no longer a model of $\Phi$. But if there are several candidates, which ones should we choose? We shall take those that are *minimal* modifications of the original $\mathscr{M}$. Note that there can be more than one $\mathscr{M}'$ that is minimal. Because adding just one of these new models is enough to invalidate $\Phi$, we take all possible combinations $\mathcal{M} \cup \{\mathscr{M}'\}$ of expanding our set of models by one minimal model. The result will be a *set of sets of models*. In each set of models there will be one $\mathscr{M}'$ falsifying $\Phi$.

## Contraction of Executability Laws

Intuitively, to contract an executability law $\varphi \to \langle a\rangle\top$ in *one* model, we *remove arrows* leaving $\varphi$-worlds. To success the operation, we have to guarantee that in the resulting model there is at least one $\varphi$-world with no departing $a$-arrow.

**Definition 4** *Let $\mathscr{M} = \langle W, R\rangle$ be a* PDL*-model. Then $\mathscr{M}' = \langle W', R'\rangle \in \mathscr{M}^-_{\varphi \to \langle a\rangle\top}$ if and only if*

- *$W' = W$*
- *$R' \subseteq R$*
- *If $(w, w') \in R \setminus R'$, then $\models^{\mathscr{M}}_w \varphi$*
- *There is $w \in W'$ s.t. $\not\models^{\mathscr{M}'}_w \varphi \to \langle a\rangle\top$*

To get minimal change, we want such an operation to be minimal w.r.t. the original model: we should remove a minimum set of arrows sufficient to get the desired result.

**Definition 5** *Let $\mathscr{M}$ be a* PDL*-model and $\varphi \to \langle a\rangle\top$ an executability law. Then*

$$contraction(\mathscr{M}, \varphi \to \langle a\rangle\top) = \bigcup \min\{\mathscr{M}^-_{\varphi \to \langle a\rangle\top}, \preceq_{\mathscr{M}}\}$$

And now we define the sets of possible models resulting from the contraction of an executability in a set of models:

**Definition 6** *Let $\mathcal{M}$ be a set of models, and $\varphi \to \langle a\rangle\top$ an executability law. Then $\mathcal{M}^-_{\varphi \to \langle a\rangle\top} = \{\mathcal{M}' : \mathcal{M}' = \mathcal{M} \cup \{\mathscr{M}'\}, \mathscr{M}' \in contraction(\mathscr{M}, \varphi \to \langle a\rangle\top), \mathscr{M} \in \mathcal{M}\}$.*

In our example, consider $\mathcal{M} = \{\mathscr{M}\}$, where $\mathscr{M}$ is the model in Figure 2. When the agent discovers that even with a token she does not manage to buy a coffee anymore, she has to change her models in order to admit models with states where *token* is the case but from which there is no *buy*-transition at all. Because having just one of such worlds in each new model is enough, taking those resulting models whose accessibility relations are maximal guarantees minimal change. Hence we get $\mathcal{M}^-_{token \to \langle buy \rangle \top} = \{\mathcal{M} \cup \{\mathscr{M}'_i\} : 1 \leq i \leq 3\}$, where each $\mathscr{M}'_i$ is depicted in Figure 3.

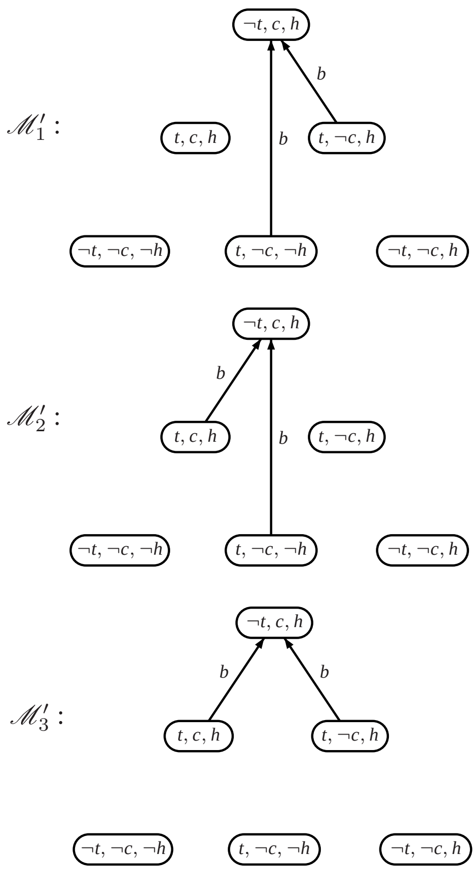


Figure 3: Models resulting from contracting the executability law *token* $\to \langle$*buy*$\rangle\top$ in the model $\mathscr{M}$ of Figure 2.

## Contraction of Effect Laws

When the agent discovers that there may be cases when after buying she gets no hot beverage, she must give up the law *token* $\to$ [*buy*]*hot* in her models. This means that *token* $\wedge$ $\langle$*buy*$\rangle\neg$*hot* shall now be admitted in at least one world of some of her new models of beliefs. Hence, to contract an effect law $\varphi \to [a]\psi$ from a given model, we have to *put new arrows* leaving $\varphi$-worlds to worlds satisfying $\neg\psi$.

In our example, when contracting *token* $\to$ [*buy*]*hot* in the model of Figure 2, we add arrows from *token*-worlds to $\neg$*hot*-worlds. The challenge in such an operation is in guaranteeing minimal change: because *coffee* $\to$ *hot*, and then $\neg$*hot* $\to$ $\neg$*coffee*, this should also give $\langle$*buy*$\rangle\neg$*coffee* ($\neg$*coffee* is *relevant* to $\neg$*hot*). Hence, we can add arrows from *token*-worlds to $\neg$*hot* $\wedge$ $\neg$*coffee* $\wedge$ *token*-worlds, as well as to $\neg$*hot* $\wedge$ $\neg$*coffee* $\wedge$ $\neg$*token* (Figure 4). Pointing the arrow to $\neg$*hot* $\wedge$ $\neg$*coffee* $\wedge$ *token* would make us lose the effect $\neg$*token*, true after every execution of *buy* in the original model. How to preserve this law while allowing for the new transition to a $\neg$*hot*-world?

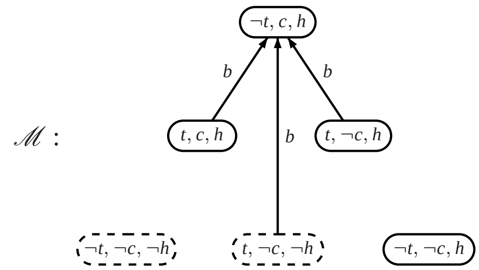


Figure 4: Candidate $\neg$*hot*-worlds to receive arrows from *token*-worlds.

When pointing a new arrow leaving a world $w$ it is enough to preserve old effects only in $w$ (because the remaining structure of the model keeps unchanged after adding this new arrow). The operation we must carry out is observing what is true in $w$ and in the candidate target world $w'$: what changes from $w$ to $w'$ ($w' \setminus w$) must be what is obliged to do so; what does not change from $w$ to $w'$ ($w \cap w'$) must be what is either obliged or allowed to do so.

This means that the only things allowed to change w.r.t. $w$ in the candidate target world are those that are forced to change: they are relevant to $\neg\psi$ or to another effect that applies in $w$. Every change outside that is not an intended one. Similarly, we want the literals preserved in the target world to be those that are relevant to $\neg\psi$ or to some other effect that applies in $w$ or that are usually preserved in $w$. Every preservation outside those may make us lose some law.

Here is where prime subvaluations play their role: the worlds one should aim the new arrow at are those whose difference w.r.t. $w$ are literals that are relevant, and whose similarity w.r.t. $w$ are literals we know may not change.

**Definition 7** *Let $\mathscr{M} = \langle W, R \rangle$, $w, w' \in W$, $\mathcal{M}$ be such that $\mathscr{M} \in \mathcal{M}$, and $\varphi \to [a]\psi$ an effect law. Then $w'$ is a* relevant target world *of $w$ w.r.t. $\varphi \to [a]\psi$ for $\mathscr{M}$ in $\mathcal{M}$ if and only if*

- $\models^{\mathscr{M}}_{w} \varphi$, $\not\models^{\mathscr{M}}_{w'} \psi$
- *for all $\ell \in w' \setminus w$*
  - *either there is $v \in base(\neg\psi, W)$ s.t. $v \subseteq w'$ and $\ell \in v$*
  - *or there are $\psi' \in \mathfrak{Fml}$, $v' \in base(\psi', W)$ s.t. $v' \subseteq w'$, $\ell \in v'$, and $\models^{\mathscr{M}_i}_{w} [a]\psi'$ for every $\mathscr{M}_i \in \mathcal{M}$*
- *for all $\ell \in w \cap w'$*
  - *either there is $v \in base(\neg\psi, W)$ s.t. $v \subseteq w'$ and $\ell \in v$*
  - *or there are $\psi' \in \mathfrak{Fml}$, $v' \in base(\psi', W)$ s.t. $v' \subseteq w'$, $\ell \in v'$, and $\models^{\mathscr{M}_i}_{w} [a]\psi'$ for every $\mathscr{M}_i \in \mathcal{M}$*
  - *or there is $\mathscr{M}_i \in \mathcal{M}$ such that $\not\models^{\mathscr{M}_i}_{w} [a]\neg\ell$*

*By RelTgt$(w, \varphi \to [a]\psi, \mathscr{M}, \mathcal{M})$ we denote the set of all relevant target worlds of $w$ w.r.t. $\varphi \to [a]\psi$ for $\mathscr{M}$ in $\mathcal{M}$.*

We need the set of models $\mathcal{M}$ (and here we can suppose it contains all models of the theory we want to change) because preserving effects depends on what other effects hold in the other models that interest us. One needs to take them into account in the local operation of changing one model:[2]

**Definition 8** *Let $\mathscr{M} = \langle W, R\rangle$ be a* PDL*-model and $\mathcal{M}$ be such that $\mathscr{M} \in \mathcal{M}$. Then $\mathscr{M}' = \langle W', R'\rangle \in \mathscr{M}^{-}_{\varphi\to[a]\psi}$ if and only if*

- $W' = W$
- $R \subseteq R'$
- $(w, w') \in R' \setminus R$ *implies* $w' \in RelTgt(w, \varphi \to [a]\psi, \mathscr{M}, \mathcal{M})$
- *There is $w \in W'$ s.t.* $\not\models^{\mathscr{M}'}_{w} \varphi \to [a]\psi$

As having just one world where the law is no longer true in each model is enough, taking those resulting models whose difference w.r.t. the original accessibility relation is minimal guarantees minimal change:

**Definition 9** *Let $\mathscr{M}$ be a* PDL*-model and $\varphi \to [a]\psi$ an effect law. Then*

$$contraction(\mathscr{M}, \varphi \to [a]\psi) = \bigcup \min\{\mathscr{M}^{-}_{\varphi\to[a]\psi}, \preceq_{\mathscr{M}}\}$$

Now we can define the possible sets of models resulting from contracting an effect law from a set of models:

**Definition 10** *Let $\mathcal{M}$ be a set of models, and $\varphi \to [a]\psi$ an effect law. Then $\mathcal{M}^{-}_{\varphi\to[a]\psi} = \{\mathcal{M}' : \mathcal{M}' = \mathcal{M} \cup \{\mathscr{M}'\}, \mathscr{M}' \in contraction(\mathscr{M}, \varphi \to [a]\psi), \mathscr{M} \in \mathcal{M}\}$.*

Taking again $\mathcal{M} = \{\mathscr{M}\}$, for $\mathscr{M}$ as in Figure 2, after contracting the effect law *token* $\to$ [*buy*]*hot* from $\mathcal{M}$, we get $\mathcal{M}^{-}_{token\to[buy]hot} = \{\mathcal{M} \cup \{\mathscr{M}'_i\} : 1 \leq i \leq 3\}$, where all $\mathscr{M}'_i$s are as depicted in Figure 5.

If $\varphi$ is not satisfied by $\mathscr{M}$ or $\psi$ is true in $\mathscr{M}$, of course we do not succeed in falsifying $\varphi \to [a]\psi$. In these cases, prior to do that we must change our set of possible states.

## Contraction of Static Laws

When contracting a static law in a model, we want to admit at least one possible state falsifying it. Intuitively this means that we should add new worlds to the original model. This is quite easy. A delicate issue however is what to do with the accessibility relation: should new arrows leave/arrive at the new world? If no arrow leaves the new added world, we may lose an executability law. If some arrow leaves it, we may lose an effect law, the same holding if we add an arrow pointing to the new world. If no arrow arrives at the new world, what about the intuition? Do we want to have an unreachable state?

All this discussion shows how drastic a change in the static laws may be: it is a change in the underlying structure (possible states) of the world! Changing it may have as consequence the loss of an effect law or an executability law.

[2]We do not need $\mathcal{M}$ in the local contraction of executabilities $\mathscr{M}^{-}_{\varphi\to\langle a\rangle\top}$ as all effects are preserved along the removal of arrows.

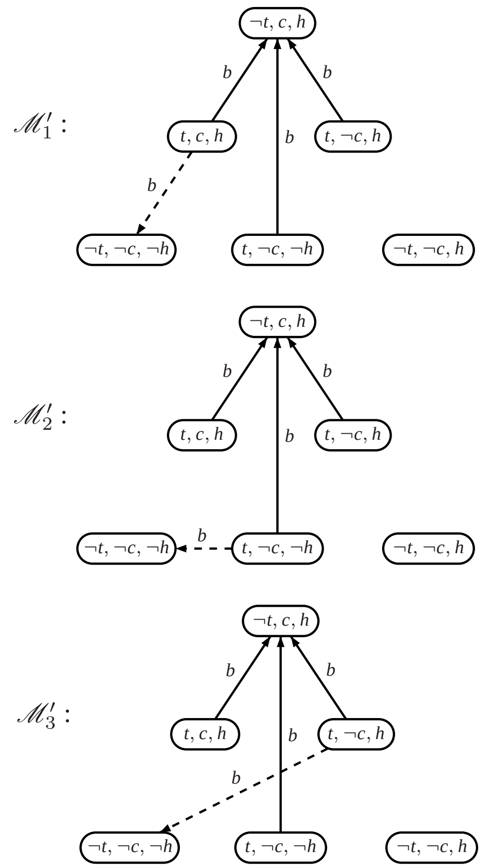


Figure 5: Models resulting from contracting the effect law *token* $\to$ [*buy*]*hot* in the model $\mathscr{M}$ of Figure 2. The new arrows are the dashed ones.

What we can do is choose which laws we accept to lose and postpone their change (by the other operators).

The tradition in the reasoning about actions community says that executability laws are, in general, more difficult to state than effect laws, and hence are more likely to be incorrect. Relying on this, in (Herzig, Perrussel, and Varzinczak 2006) no change in the accessibility relation is made, what means preserving effect laws and postponing correction of executability laws. We here embrace this solution. It is controversial whether this approach is in line with the intuition or not (see (Varzinczak 2008a) for an alternative). Anyway, with the information we have at hand, this is the safest way of contracting static laws.

**Definition 11** *Let $\mathscr{M} = \langle W, R\rangle$ be a* PDL*-model. Then $\mathscr{M}' = \langle W', R'\rangle \in \mathscr{M}^{-}_{\varphi}$ if and only if*

- $W \subseteq W'$
- $R = R'$
- *There is $w \in W'$ s.t.* $\not\models^{\mathscr{M}'}_{w} \varphi$

The minimal modifications of one model are as expected:

**Definition 12** *Let $\mathscr{M}$ be a model and $\varphi$ a static law. Then*

$$contraction(\mathscr{M}, \varphi) = \bigcup \min\{\mathscr{M}^{-}_{\varphi}, \preceq_{\mathscr{M}}\}$$

And we define the sets of models resulting from contracting a static law from one set of models:

**Definition 13** *Let $\mathcal{M}$ be a set of models, and $\varphi$ a static law. Then $\mathcal{M}^{-}_{\varphi} = \{\mathcal{M}' : \mathcal{M}' = \mathcal{M} \cup \{\mathscr{M}'\}, \mathscr{M}' \in contraction(\mathscr{M}, \varphi), \mathscr{M} \in \mathcal{M}\}$.*

In our example, contracting the static law *coffee* $\rightarrow$ *hot* from $\mathcal{M} = \{\mathscr{M}\}$, with $\mathscr{M}$ as in Figure 2, will give us $\mathcal{M}^{-}_{coffee \rightarrow hot} = \{\mathcal{M} \cup \{\mathscr{M}'_1\}, \mathcal{M} \cup \{\mathscr{M}'_2\}\}$, where each $\mathscr{M}'_i$ is as depicted in Figure 6.

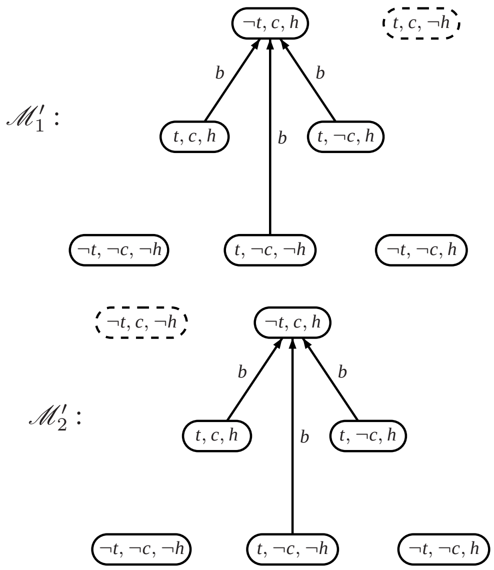


Figure 6: Models resulting from contracting the static law *coffee* $\rightarrow$ *hot* in the model $\mathscr{M}$ of Figure 2. The new added *coffee* $\wedge \neg$*hot*-worlds are dashed.

Notice that by not modifying the accessibility relation all the effect laws are preserved with minimal change. Moreover, our approach is also intuitive: when learning that a new state is now possible, we do not necessarily know all the behavior of the action in the new added state.

## Syntactic Operators for Contraction

We now turn our attention to the definition of a syntactical counterpart of our semantic operators. As (Nebel 1989) says, "[...] finite bases usually represent [...] laws, and when we are forced to change the theory we would like to stay as close as possible to the original [...] base." Hence, besides the definition of syntactical operators, we should also guarantee that they perform minimal change.

By $\mathcal{T}^{-}_{\Phi}$ we denote in the sequel the result of contracting a law $\Phi$ from the set of laws $\mathcal{T}$.

### Contracting Executability Laws

For the case of contracting an executability law $\varphi \rightarrow \langle a \rangle \top$ from an action theory, the first thing we do is to ensure that the action *a* is still executable (if that was so) in all those contexts where $\neg\varphi$ is the case. Second, in order to get minimality, we must make *a* executable in *some* contexts where $\varphi$ is true, viz. all $\varphi$-worlds but one. This means that we can have several action theories as outcome.

Algorithm 1 gives a syntactical operator to achieve this.

**Algorithm 1** Contraction of an executability law

**input:** $\mathcal{T}, \varphi \rightarrow \langle a \rangle \top$
**output:** $\mathcal{T}^{-}_{\varphi \rightarrow \langle a \rangle \top}$ /* a set of theories */
  **if** $\mathcal{T} \models_{\mathsf{PDL}} \varphi \rightarrow \langle a \rangle \top$ **then**
    **for all** $\pi \in IP(\mathcal{S} \wedge \varphi)$ **do**
      **for all** $A \subseteq \overline{atm(\pi)}$ **do**
        $\varphi_A := \bigwedge_{\substack{p_i \in \overline{atm(\pi)} \\ p_i \in A}} p_i \wedge \bigwedge_{\substack{p_i \in \overline{atm(\pi)} \\ p_i \notin A}} \neg p_i$
        **if** $\mathcal{S} \not\models_{\mathsf{CPL}} (\pi \wedge \varphi_A) \rightarrow \bot$ **then**
          $\mathcal{T}' := (\mathcal{T} \setminus \mathcal{X}_a) \cup \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\}$
          $\mathcal{T}^{-}_{\varphi \rightarrow \langle a \rangle \top} := \mathcal{T}^{-}_{\varphi \rightarrow \langle a \rangle \top} \cup \{\mathcal{T}'\}$
  **else**
    $\mathcal{T}^{-}_{\varphi \rightarrow \langle a \rangle \top} := \{\mathcal{T}\}$

As an example, contracting *token* $\rightarrow \langle buy \rangle \top$ from our theory $\mathcal{T}$ would give us three theories. One of them is:

$$\mathcal{T}'_1 = \left\{ \begin{array}{c} coffee \rightarrow hot,\ \neg coffee \rightarrow [buy]coffee, \\ token \rightarrow [buy]\neg token,\ \neg token \rightarrow [buy]\bot, \\ \neg token \rightarrow [buy]\neg token,\ coffee \rightarrow [buy]coffee, \\ hot \rightarrow [buy]hot, \\ (token \wedge \neg coffee \wedge hot) \rightarrow \langle buy \rangle \top, \\ (token \wedge \neg coffee \wedge \neg hot) \rightarrow \langle buy \rangle \top \end{array} \right\}$$

### Contracting Effect Laws

When contracting an effect law $\varphi \rightarrow [a]\psi$ from a theory $\mathcal{T}$, intuitively we should change some effect laws that preclude $\neg\psi$ in target worlds. In order to cope with minimality, we must change only those laws that are relevant to $\varphi \rightarrow [a]\psi$.

Let $\mathcal{E}^{\varphi,\psi}_a$ denote the minimum subset of the effect laws in $\mathcal{E}_a$ such that $\mathcal{S}, \mathcal{E}^{\varphi,\psi}_a \models_{\mathsf{PDL}} \varphi \rightarrow [a]\psi$. In the case where the theory is modular (Herzig and Varzinczak 2005) (see further), interpolation guarantees that such a set always exists. Moreover, note that there can be more than one such a set, in which case we denote them $(\mathcal{E}^{\varphi,\psi}_a)_1, \ldots, (\mathcal{E}^{\varphi,\psi}_a)_n$. Let

$$\mathcal{E}^{-}_a = \bigcup_{1 \leq i \leq n} (\mathcal{E}^{\varphi,\psi}_a)_i$$

The laws in $\mathcal{E}^{-}_a$ will serve as guideline to get rid of $\varphi \rightarrow [a]\psi$ in the theory.

The first thing that we must do is to ensure that action *a* still has effect $\psi$ (if that was so) in all those contexts in which $\varphi$ does not hold. This means that we shall weaken the laws in $\mathcal{E}^{\varphi,\psi}_a$ specializing them to $\neg\varphi$.

Second, we need to preserve all old effects in all $\varphi$-worlds but one. To achieve that, we specialize the above laws to each possible valuation satisfying $\varphi$ but one. In the left $\varphi$-valuation, we must ensure that action *a* has either its old effects or $\neg\psi$ as outcome. We achieve that by weakening the *consequent* of the laws in $\mathcal{E}^{-}_a$.

Finally, in order to get minimal change, we must ensure that all literals in this $\varphi$-valuation that are not forced to change in $\neg\psi$-worlds should be preserved. We do this by stating an effect law of the form $(\varphi_k \wedge \ell) \to [a](\psi \vee \ell)$, where $\varphi_k$ is the above $\varphi$-valuation. The reason why this is needed is clear: there can be several $\neg\psi$-valuations, and as far as we want at most one to be reachable from $\varphi_k$, we should force it to be the one whose difference to $\varphi_k$ is minimal.

Again, the result will be a set of theories. Algorithm 2 below gives the operator.

**Algorithm 2** Contraction of an effect law

**input:** $\mathcal{T}, \varphi \to [a]\psi$
**output:** $\mathcal{T}^-_{\varphi\to[a]\psi}$ /* a set of theories */
  **if** $\mathcal{T} \models_{\overline{\mathsf{PDL}}} \varphi \to [a]\psi$ **then**
    **for all** $\pi \in IP(\mathcal{S} \wedge \varphi)$ **do**
      **for all** $A \subseteq \overline{atm(\pi)}$ **do**
        $\varphi_A := \bigwedge_{\substack{p_i \in \overline{atm(\pi)} \\ p_i \in A}} p_i \wedge \bigwedge_{\substack{p_i \in \overline{atm(\pi)} \\ p_i \notin A}} \neg p_i$
        **if** $\mathcal{S} \not\models_{\mathsf{CPL}} (\pi \wedge \varphi_A) \to \bot$ **then**
          **for all** $\pi' \in IP(\mathcal{S} \wedge \neg\psi)$ **do**

$$\begin{array}{l} \mathcal{T}' := (\mathcal{T} \setminus \mathcal{E}^-_a) \cup \\ \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \to [a]\psi_i : \varphi_i \to [a]\psi_i \in \mathcal{E}^-_a\} \cup \\ \{(\varphi_i \wedge \pi \wedge \varphi_A) \to [a](\psi_i \vee \pi') : \varphi_i \to [a]\psi_i \in \mathcal{E}^-_a\} \end{array}$$

            **for all** $L \subseteq \mathfrak{Lit}$ **do**
              **if** $\mathcal{S} \models_{\mathsf{CPL}} (\pi \wedge \varphi_A) \to \bigwedge_{\ell \in L} \ell$ **and** $\mathcal{S} \not\models_{\mathsf{CPL}} (\pi' \wedge \bigwedge_{\ell \in L} \ell) \to \bot$ **then**
                **for all** $\ell \in L$ **do**
                  **if** $\mathcal{T} \not\models_{\overline{\mathsf{PDL}}} (\pi \wedge \varphi_A \wedge \ell) \to [a]\neg\ell$ **or** $\ell \in \pi'$ **then**
                    $\mathcal{T}' := \mathcal{T}' \cup \{(\pi \wedge \varphi_A \wedge \ell) \to [a](\psi \vee \ell)\}$
            $\mathcal{T}^-_{\varphi\to[a]\psi} := \mathcal{T}^-_{\varphi\to[a]\psi} \cup \{\mathcal{T}'\}$
  **else**
    $\mathcal{T}^-_{\varphi\to[a]\psi} := \{\mathcal{T}\}$

For instance, contracting the effect law *token* $\to [buy]hot$ from $\mathcal{T}$ will give us three resulting theories, one of them is $\mathcal{T}'_1 =$

$$\left\{ \begin{array}{c} coffee \to hot,\ token \to \langle buy \rangle \top, \\ token \to [buy]\neg token,\ \neg token \to [buy]\bot, \\ \neg token \to [buy]\neg token, \\ (coffee \wedge \neg(token \wedge coffee \wedge hot)) \to [buy]coffee, \\ (hot \wedge \neg(token \wedge coffee \wedge hot)) \to [buy]hot, \\ (\neg coffee \wedge \neg(token \wedge coffee \wedge hot)) \to [buy]coffee, \\ (token \wedge coffee \wedge hot) \to [buy](coffee \vee \neg hot), \\ (token \wedge coffee \wedge hot) \to [buy](hot \vee \neg coffee) \end{array} \right\}$$

### Contracting Static Laws

Finally, in order to contract a static law from a theory, we can use any standard contraction/revision operator $\ominus$ for classical propositional logic to change the set of static laws $\mathcal{S}$. Because contracting static laws means *admitting* new possible states (cf. the semantics), it may be the case that just modifying $\mathcal{S}$ is not enough.

Since we in general do not necessarily know the behavior of the actions in a new discovered state of the world, a careful approach is to change the theory so that all action laws remain the same in the contexts where the contracted law is the case. In our example, if when contracting the law *coffee* $\to$ *hot* we are not sure whether *buy* is still executable or not, we should weaken our executability laws specializing them to the context *coffee* $\to$ *hot*, and then make *buy* a priori inexecutable in all $\neg$(*coffee* $\to$ *hot*) contexts.

Algorithm 3 below formalizes such an operation.

**Algorithm 3** Contraction of a static law

**input:** $\mathcal{T}, \varphi$
**output:** $\mathcal{T}^-_{\varphi}$ /* a set of theories */
  **if** $\mathcal{S} \models_{\mathsf{CPL}} \varphi$ **then**
    **for all** $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$ **do**

$$\mathcal{T}' := \begin{array}{l} ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ \{(\varphi_i \wedge \varphi) \to \langle a \rangle \top : \varphi_i \to \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ \{\neg\varphi \to [a]\bot\} \end{array}$$

      $\mathcal{T}^-_{\varphi} := \mathcal{T}^-_{\varphi} \cup \{\mathcal{T}'\}$
  **else**
    $\mathcal{T}^-_{\varphi} := \{\mathcal{T}\}$

In our running example, contracting the law *coffee* $\to$ *hot* from $\mathcal{T}$ produces two theories, one of them is

$$\mathcal{T}'_1 = \left\{ \begin{array}{c} \neg(\neg token \wedge coffee \wedge \neg hot), \\ (token \wedge coffee \to hot) \to \langle buy \rangle \top, \\ \neg coffee \to [buy]coffee, token \to [buy]\neg token, \\ \neg token \to [buy]\bot, \neg token \to [buy]\neg token, \\ coffee \to [buy]coffee, hot \to [buy]hot, \\ (coffee \wedge \neg hot) \to [buy]\bot \end{array} \right\}$$

Observe that the effect laws are not affected by the change: as far as we do not state executabilities for the new world, all the effect laws remain true in it.

If the knowledge engineer is not happy with the added inexecutability law $(coffee \wedge \neg hot) \to [buy]\bot$, she can contract it from the theory using Algorithm 2.

## Correctness of the Operators

Here we show that our algorithms are correct w.r.t. our semantics for action theory contraction. Before doing that, we need a definition.

**Definition 14 (Modularity (Herzig and Varzinczak 2005))** *An action theory* $\mathcal{T}$ *is* modular *if and only if for every* $\varphi \in \mathfrak{Fml}$, *if* $\mathcal{T} \models_{\overline{\mathsf{PDL}}} \varphi$, *then* $\mathcal{S} \models_{\mathsf{CPL}} \varphi$.

For an example of a non-modular theory, suppose in our action theory $\mathcal{T}$ we had stated the law $\langle buy \rangle \top$ instead of *token* $\to \langle buy \rangle \top$. Then $\mathcal{T} \models_{\overline{\mathsf{PDL}}} token$ and $\mathcal{S} \not\models_{\mathsf{CPL}} token$.

In (Herzig and Varzinczak 2005) algorithms are given to check whether $\mathcal{T}$ satisfies the principle of modularity and also to make $\mathcal{T}$ satisfy it, if that is not the case.

**Theorem 3** *$\mathcal{T}$ is modular if and only if its big model is a model of $\mathcal{T}$.*

Modular theories have interesting properties. For example, if $\mathcal{T}$ is modular, then its consistency can be checked by just checking consistency of the set of static laws $\mathcal{S}$ alone. Deduction of effect laws does not need the executability ones and vice versa. Prediction of an effect of a sequence of actions $a_1; \ldots; a_n$ does not need the effect laws for actions other than $a_1, \ldots, a_n$. This also applies to plan validation when deciding whether $\langle a_1; \ldots; a_n \rangle \varphi$ is the case. For more results on modularity, see (Herzig and Varzinczak 2007).

The following theorem (see Appendix A for the proof) establishes that the semantic contraction of the law $\Phi$ from the set of models of the action theory $\mathcal{T}$ produces models of some contracted theory in $\mathcal{T}^-_\Phi$.

**Theorem 4** *Let $\mathcal{T}$ be modular, and $\Phi$ be a law. For all $\mathcal{M}' \in \mathcal{M}^-_\Phi$ such that $\models^{\mathscr{M}} \mathcal{T}$ for every $\mathscr{M} \in \mathcal{M}$, there is $\mathcal{T}' \in \mathcal{T}^-_\Phi$ such that $\models^{\mathscr{M}'} \mathcal{T}'$ for every $\mathscr{M}' \in \mathcal{M}'$.*

The next theorem establishes the other way round: models of theories in $\mathcal{T}^-_\Phi$ are all models of the semantical contraction of $\Phi$ from models of $\mathcal{T}$. (The proof is in Appendix B.)

**Theorem 5** *Let $\mathcal{T}$ be modular, $\Phi$ a law, and $\mathcal{T}' \in \mathcal{T}^-_\Phi$. For all $\mathscr{M}'$ such that $\models^{\mathscr{M}'} \mathcal{T}'$, there is $\mathcal{M}' \in \mathcal{M}^-_\Phi$ such that $\mathscr{M}' \in \mathcal{M}'$ and $\models^{\mathscr{M}} \mathcal{T}$ for every $\mathscr{M} \in \mathcal{M}$.*

Hence our operators are correct w.r.t. the semantics.

## Assessment of Postulates for Change

We now analyze our operator's behavior w.r.t. Katsuno and Mendelzon's classical contraction postulates. (Due to space limitations, proofs are omitted here. They are all available at (Varzinczak 2008a).)

**Theorem 6** *$\mathcal{T} \models_{\mathsf{PDL}} \mathcal{T}'$, for all $\mathcal{T}' \in \mathcal{T}^-_\Phi$.*

This result means our operators satisfy the PDL-version of Katsuno and Mendelzon's (C1) postulate about *monotonicity*. Such a postulate is not satisfied by the operators given in (Herzig, Perrussel, and Varzinczak 2006): there, when removing e.g. an executability law $\varphi \rightarrow \langle a \rangle \top$ one may make $\varphi \rightarrow [a]\bot$ valid in all models of the resulting theory.

**Theorem 7** *If $\mathcal{T} \not\models_{\mathsf{PDL}} \Phi$, then $\models_{\mathsf{PDL}} \mathcal{T} \leftrightarrow \mathcal{T}'$, for all $\mathcal{T}' \in \mathcal{T}^-_\Phi$.*

This corresponds to Katsuno and Mendelzon's (C2) postulate about *preservation*. Whenever $\mathcal{T} \not\models_{\mathsf{PDL}} \Phi$, then the models of the resulting theory are exactly the models of $\mathcal{T}$, because these are the minimal models falsifying $\Phi$.

**Theorem 8** *Let $\mathcal{T} = \mathcal{S} \cup \mathcal{E} \cup \mathcal{X}$ be consistent, and $\Phi$ be an executability or an effect law such that $\mathcal{S} \not\models_{\mathsf{PDL}} \Phi$. If $\mathcal{T}$ is modular, then $\mathcal{T}' \not\models_{\mathsf{PDL}} \Phi$ for every $\mathcal{T}' \in \mathcal{T}^-_\Phi$*

Thus, under modularity our operators satisfy the *success* postulate (C3). Still under modularity and the assumption that the classical contraction operator satisfies Katsuno and Mendelzon's (C4) postulate, our operations also satisfy it:

**Theorem 9** *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be modular. If $\models_{\mathsf{PDL}} \mathcal{T}_1 \leftrightarrow \mathcal{T}_2$ and $\models_{\mathsf{PDL}} \Phi_1 \leftrightarrow \Phi_2$, then for each $\mathcal{T}'_1 \in (\mathcal{T}_1)^-_{\Phi_2}$ there is $\mathcal{T}'_2 \in (\mathcal{T}_2)^-_{\Phi_1}$ such that $\models_{\mathsf{PDL}} \mathcal{T}'_1 \leftrightarrow \mathcal{T}'_2$, and vice-versa.*

Thanks to modularity, our operators also satisfy Katsuno and Mendelzon's (C5) postulate, *recovery*:

**Theorem 10** *Let $\mathcal{T}$ be modular. $\mathcal{T}' \cup \{\Phi\} \models_{\mathsf{PDL}} \mathcal{T}$, for all $\mathcal{T}' \in \mathcal{T}^-_\Phi$.*

**Theorem 11** *If $\mathcal{T}$ is modular, then every $\mathcal{T}' \in \mathcal{T}^-_\Phi$ is also modular.*

Besides satisfying all postulates for contraction, our operators also preserve modularity. This is a nice property, since it means that modularity can be checked/ensured once for all during the theory's evolution.

## Related Work

To the best of our knowledge, the first work on updating action theories is that by (Li and Pereira 1996) in a narrative-based action description language (Gelfond and Lifschitz 1993). Contrary to us, however, they investigate the problem of updating the narrative with new observed *facts* and (possibly) with occurrences of actions that explain those facts.

This amounts to updating a given state/configuration of the world (in our terms, what is true in a possible world) and focusing on the models of the narrative in which some actions took place (in our terms, the models of the action theory with a particular sequence of action executions). Clearly the models of the action laws remain the same.

(Liberatore 2000) proposes an action language in which one can express a given semantics for belief update, like (Winslett 1988) and (Katsuno and Mendelzon 1992). Update operations are then expressed as action laws in a theory.

The main difference between Liberatore's work and Li and Pereira's is that Liberatore's framework allows for abductively adding to the action theory new effect propositions (effect laws, in our terms) that consistently explain the occurrence of an event.

The work by (Eiter et al. 2005) is similar to ours in that they also propose a framework for updating action laws. They mainly investigate the case where e.g. a new effect law shall be added to the description. This problem is the dual of contraction and is then closer to *revision*.

In Eiter *et al.*'s approach, action theories are also described in a variant of a narrative-based action language. Like here, the semantics is in terms of transition systems. Contrary to us, the minimality condition on the outcome of the update is in terms of inclusion of sets of laws, which means the approach is more syntax-oriented than ours.

Both their framework and ours can be qualified as constraint-based update, in that the update is carried out relative to a set of laws that one wants to hold in the result. Here for example, all changes in the action laws are relative to the static laws in $\mathcal{S}$.

One difference between our approach and Eiter *et al.*'s is that there it is also possible to update a theory relatively to e.g. executability laws: when expanding $\mathcal{T}$ with a new effect law, one may want to constrain the change so that the action under concern is guaranteed to be executable in the result. This may of course require the withdrawal of some static law. Hence, in Eiter *et al.*'s framework, static laws do not have the same status as in ours.

## Concluding Remarks

The contributions of the present work are as follows:

- What is the meaning of removing a law $\Phi$ from an action theory $\mathcal{T}$? How to get minimal change, i.e., how to keep as much knowledge about other laws as possible? We answered these questions with Definitions 6, 10 and 13.
- How to syntactically contract an action theory so that its result corresponds to the intended semantics? We answered this question with Algorithms 1–3 and Theorems 4 and 5.
- Is our method closer to update or revision? Does it comply with the standard postulates for classical theory change and what are the differences w.r.t. that? We answered these questions with Theorems 6–11.

We have shown the importance that modularity has in action theory change. Under modularity, our operators satisfy all Katsuno and Mendelzon's postulates for contraction. This shows that our modularity notion is fruitful. Moreover, considering future modifications one should perform on the theory, since modularity is preserved by our operators, it suffices to check/ensure it only once.

Here we presented the case for contraction. We are currently investigating the definition of the revision counterpart of action theory change. The first results on this issue are available in (Varzinczak 2008b).

Our ongoing research is on how to contract not only laws but *any* PDL-formula. Definitions 4, 8 and 11 show up to be important for better understanding the case of general formulas: the modifications to perform in a given model in order to falsify a general formula will also comprise removal/addition of arrows and worlds. The definition of a more general contraction method will thus benefit from our present constructions.

## Acknowledgements

The author is thankful to Andreas Herzig and Laurent Perrussel for interesting discussions on the subject of this work. He is also grateful to the anonymous referees for their useful comments on an earlier version of this paper.

This work has been partially supported by a fellowship from the government of the FEDERATIVE REPUBLIC OF BRAZIL. Grant: CAPES BEX 1389/01-7.

## References

Eiter, T.; Erdem, E.; Fink, M.; and Senko, J. 2005. Updating action domain descriptions. In *Proc. 19th Intl. Joint Conf. on Artificial Intelligence (IJCAI'05)*, 418–423. Morgan Kaufmann.

Finger, J. 1987. *Exploiting constraints in design synthesis*. Ph.D. Dissertation, Stanford University, Stanford.

Fuhrmann, A. 1989. On the modal logic of theory change. In *The Logic of Theory Change*, 259–281.

Gärdenfors, P. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge, MA: MIT Press.

Gelfond, M., and Lifschitz, V. 1993. Representing action and change by logic programs. *Journal of Logic Programming* 17(2/3&4):301–321.

Hansson, S. 1999. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers.

Harel, D.; Tiuryn, J.; and Kozen, D. 2000. *Dynamic Logic*. Cambridge, MA: MIT Press.

Herzig, A., and Rifi, O. 1999. Propositional belief base update and minimal change. *Artificial Intelligence* 115(1):107–138.

Herzig, A., and Varzinczak, I. 2005. On the modularity of theories. In *Advances in Modal Logic*, volume 5. King's College Publications. 93–109.

Herzig, A., and Varzinczak, I. 2007. Metatheory of actions: beyond consistency. *Artificial Intelligence* 171:951–984.

Herzig, A.; Perrussel, L.; and Varzinczak, I. 2006. Elaborating domain descriptions. In *Proc. 17th Eur. Conf. on Artificial Intelligence (ECAI'06)*, 397–401. IOS Press.

Jin, Y., and Thielscher, M. 2005. Iterated belief revision, revised. In *Proc. 19th Intl. Joint Conf. on Artificial Intelligence (IJCAI'05)*, 478–483. Morgan Kaufmann.

Katsuno, H., and Mendelzon, A. 1992. On the difference between updating a knowledge base and revising it. In *Belief revision*. Cambridge University Press. 183–203.

Li, R., and Pereira, L. 1996. What is believed is what is explained. In *Proc. 13th Natl. Conf. on Artificial Intelligence (AAAI'96)*, 550–555. AAAI Press/MIT Press.

Liberatore, P. 2000. A framework for belief update. In *Proc. 7th Eur. Conf. on Logics in Artificial Intelligence (JELIA'2000)*, 361–375.

McCarthy, J., and Hayes, P. 1969. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence*, volume 4. Edinburgh University Press. 463–502.

Nebel, B. 1989. A knowledge level analysis of belief revision. In *Proc. Intl. Conf. on Knowledge Representation and Reasoning (KR'89)*, 301–311. Morgan Kaufmann.

Parikh, R. 1999. Beliefs, belief revision, and splitting languages. In *Logic, Language and Computation*, volume 2 of *CSLI Lecture Notes*. CSLI Publications. 266–278.

Quine, W. V. O. 1952. The problem of simplifying truth functions. *American Mathematical Monthly* 59:521–531.

Shapiro, S.; Pagnucco, M.; Lespérance, Y.; and Levesque, H. 2000. Iterated belief change in the situation calculus. In *Proc. 7th Intl. Conf. on Knowledge Representation and Reasoning (KR'2000)*, 527–538. Morgan Kaufmann.

Varzinczak, I. 2008a. Action theory change revisited: contracting action laws and minimal change. Technical Report IRIT/RT–2008-2–FR, IRIT, Toulouse.

Varzinczak, I. 2008b. Action theory revision. Technical Report IRIT/RT–2008-1–FR, IRIT, Toulouse.

Winslett, M.-A. 1988. Reasoning about action using a possible models approach. In *Proc. 7th Natl. Conf. on Artificial Intelligence (AAAI'88)*, 89–93.

## Appendix A: Proof of Theorem 4

**Lemma 1** $\mathcal{T} \models_{\mathsf{PDL}} \mathcal{T}'$.

For the proof of this lemma, the reader is invited to check (Varzinczak 2008a).

**Proof of Theorem 4**

Let $\mathcal{M} = \{\mathscr{M} : \models^{\mathscr{M}} \mathcal{T}\}$, and $\mathcal{M}' \in \mathcal{M}^{-}_{\Phi}$. We show that there is $\mathcal{T}' \in \mathcal{T}^{-}_{\Phi}$ such that $\models^{\mathscr{M}'} \mathcal{T}'$ for every $\mathscr{M}' \in \mathcal{M}'$.

By definition, each $\mathscr{M}' \in \mathcal{M}'$ is such that either $\models^{\mathscr{M}'} \mathcal{T}$ or $\not\models^{\mathscr{M}'} \Phi$. Because $\mathcal{T}^{-}_{\Phi} \neq \emptyset$, there must be $\mathcal{T}' \in \mathcal{T}^{-}_{\Phi}$. If $\models^{\mathscr{M}'} \mathcal{T}$, by Lemma 1 $\models^{\mathscr{M}'} \mathcal{T}'$ and we are done. Let's then suppose that $\not\models^{\mathscr{M}'} \Phi$. We analyze each case.

Let $\Phi$ have the form $\varphi \rightarrow \langle a \rangle \top$ for some $\varphi \in \mathfrak{Fml}$. Then $\mathscr{M}' = \langle W', R' \rangle$, where $W' = W$, $R' = R \setminus R^{\varphi}_{a}$, with $R^{\varphi}_{a} = \{(w, w') : \models^{\mathscr{M}}_{w} \varphi$ and $(w, w') \in R_a\}$, for some $\mathscr{M} \in \mathcal{M}$.

Let $u \in W'$ be such that $\not\models^{\mathscr{M}'}_{u} \varphi \rightarrow \langle a \rangle \top$, i.e., $\models^{\mathscr{M}'}_{u} \varphi$ and $R'_a(u) = \emptyset$.

Because $u \Vdash \varphi$, there must be $v \in base(\varphi, W')$ such that $v \subseteq u$. Let $\pi = \bigwedge_{\ell \in v} \ell$. Clearly $\pi$ is a prime implicant of $\mathcal{S} \wedge \varphi$. Let also $\varphi_A = \bigwedge_{\ell \in u \setminus v} \ell$, and consider

$$\mathcal{T}' = (\mathcal{T} \setminus \mathcal{X}_a) \cup \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\}$$

(Clearly, $\mathcal{T}'$ is a theory produced by Algorithm 1.)

It is enough to show that $\mathscr{M}'$ is a model of the new added laws. Given $(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top \in \mathcal{T}'$, for every $w \in W'$, if $\models^{\mathscr{M}'}_{w} \varphi_i \wedge \neg(\pi \wedge \varphi_A)$, then $\models^{\mathscr{M}'}_{w} \varphi_i$, from what it follows $\models^{\mathscr{M}}_{w} \varphi_i$. Because $\models^{\mathscr{M}} \varphi_i \rightarrow \langle a \rangle \top$, there is $w' \in W$ such that $w' \in R_a(w)$. We need to show that $(w, w') \in R'_a$. If $\not\models^{\mathscr{M}}_{w} \varphi$, then $R^{\varphi}_{a} = \emptyset$, and $(w, w') \in R'_a$. If $\models^{\mathscr{M}}_{w} \varphi$, either $w = u$, and then from $\models^{\mathscr{M}'}_{u} \pi \wedge \varphi_A$ we conclude $\models^{\mathscr{M}'}_{u} (\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top$, or $w \neq u$ and then we must have $(w, w') \in R'_a$, otherwise there is $S^{\varphi}_{a} \subset R^{\varphi}_{a}$ such that $R \dot{-} (R \setminus S^{\varphi}_{a}) \subset R \dot{-} (R \setminus R^{\varphi}_{a})$, and then $\mathscr{M}'' = \langle W', R \setminus S^{\varphi}_{a} \rangle$ is such that $\not\models^{\mathscr{M}''} \varphi \rightarrow \langle a \rangle \top$ and $\mathscr{M}'' \preceq_{\mathscr{M}} \mathscr{M}'$, a contradiction because $\mathscr{M}'$ is minimal w.r.t. $\preceq_{\mathscr{M}}$. Thus $(w, w') \in R'_a$, and then $\models^{\mathscr{M}'}_{w} \langle a \rangle \top$. Hence $\models^{\mathscr{M}'} \mathcal{T}'$.

Now let $\Phi$ be of the form $\varphi \rightarrow [a]\psi$, for $\varphi, \psi \in \mathfrak{Fml}$. Then $\mathscr{M}' = \langle W', R' \rangle$, where $W' = W$, $R' = R \cup R^{\varphi, \neg\psi}_{a}$, with

$$R^{\varphi, \neg\psi}_{a} = \{(w, w') : w' \in RelTgt(w, \varphi \rightarrow [a]\psi, \mathscr{M}, \mathcal{M})\}$$

for some $\mathscr{M} = \langle W, R \rangle \in \mathcal{M}$.

Let $u \in W'$ be such that $\not\models^{\mathscr{M}'}_{u} \varphi \rightarrow [a]\psi$. Then there is $u' \in W'$ such that $(u, u') \in R'_a$ and $\not\models^{\mathscr{M}'}_{u'} \psi$. Because $u \Vdash \varphi$, there is $v \in base(\varphi, W')$ such that $v \subseteq u$, and as $u' \Vdash \neg\psi$, there must be $v' \in base(\neg\psi, W')$ such that $v' \subseteq u'$. Let $\pi = \bigwedge_{\ell \in v} \ell$, $\varphi_A = \bigwedge_{\ell \in u \setminus v} \ell$, and $\pi' = \bigwedge_{\ell \in v'} \ell$. Clearly $\pi$ (resp. $\pi'$) is a prime implicant of $\mathcal{S} \wedge \varphi$ (resp. $\mathcal{S} \wedge \neg\psi$).

Now let $\mathcal{E}^{-}_{a} = \bigcup_{1 \leq i \leq n} (\mathcal{E}^{\varphi, \psi}_{a})_i$ and let the theory

$$\mathcal{T}' = (\mathcal{T} \setminus \mathcal{E}^{-}_{a}) \cup$$
$$\{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}^{-}_{a}\} \cup$$
$$\{(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi') : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}^{-}_{a}\} \cup$$
$$\left\{ \begin{array}{r} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell) : \ell \in L, \text{ for } L \subseteq \mathfrak{Lit} \text{ s.t.} \\ \mathcal{S} \not\models_{\mathsf{CPL}} (\pi' \wedge \bigwedge_{\ell \in L} \ell) \rightarrow \bot, \text{ and} \\ \ell \in \pi' \text{or } \mathcal{T} \not\models_{\mathsf{PDL}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell \end{array} \right\}$$

(Clearly, $\mathcal{T}'$ is a theory produced by Algorithm 2.)

In order to show that $\mathscr{M}'$ is a model of $\mathcal{T}'$, it is enough to show that it is a model of the added laws. Given $(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i \in \mathcal{T}'$, for every $w \in W'$, if $\models^{\mathscr{M}'}_{w} \varphi_i \wedge \neg(\pi \wedge \varphi_A)$, then $\models^{\mathscr{M}'}_{w} \varphi_i$, and then $\models^{\mathscr{M}}_{w} \varphi_i$. Because $\models^{\mathscr{M}} \varphi_i \rightarrow [a]\psi_i$, $\models^{\mathscr{M}}_{w'} \psi_i$ for all $w' \in W$ such that $(w, w') \in R_a$. We need to show that $R'_a(w) = R_a(w)$. If $\not\models^{\mathscr{M}}_{w} \varphi$, then $R^{\varphi, \neg\psi}_{a} = \emptyset$, and then $R'_a(w) = R_a(w)$. If $\models^{\mathscr{M}}_{w} \varphi$, then either $w = u$, and from $\models^{\mathscr{M}'}_{u} \pi \wedge \varphi_A$ we conclude $\models^{\mathscr{M}'}_{u} (\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i$, or $w \neq u$, and then we must have $R^{\varphi, \neg\psi}_{a} = \emptyset$, otherwise there would be $S^{\varphi, \neg\psi}_{a} \subset R^{\varphi, \neg\psi}_{a}$ such that $R \dot{-} (R \cup S^{\varphi, \neg\psi}_{a}) \subset R \dot{-} (R \cup R^{\varphi, \neg\psi}_{a})$, and then $\mathscr{M}'' = \langle W', R \cup S^{\varphi, \neg\psi}_{a} \rangle$ would be such that $\not\models^{\mathscr{M}''} \varphi \rightarrow [a]\psi$ and $\mathscr{M}'' \preceq_{\mathscr{M}} \mathscr{M}'$, a contradiction since $\mathscr{M}'$ is minimal w.r.t. $\preceq_{\mathscr{M}}$. Hence $R'_a(w) = R_a(w)$, and $\models^{\mathscr{M}'}_{w'} \psi_i$ for all $w'$ such that $(w, w') \in R'_a$.

Now, given $(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi')$, for every $w \in W'$, if $\models^{\mathscr{M}'}_{w} \varphi_i \wedge \pi \wedge \varphi_A$, then $\models^{\mathscr{M}'}_{w} \varphi_i$, and then $\models^{\mathscr{M}}_{w} \varphi_i$. Because, $\models^{\mathscr{M}} \varphi_i \rightarrow [a]\psi_i$, we have $\models^{\mathscr{M}}_{w'} \psi_i$ for all $w' \in W$ such that $(w, w') \in R_a$, and then $\models^{\mathscr{M}'}_{w'} \psi_i$ for every $w' \in W'$ such that $(w, w') \in R'_a \setminus R^{\varphi, \neg\psi}_{a}$. Now, given $(w, w') \in R^{\varphi, \neg\psi}_{a}$, $\models^{\mathscr{M}'}_{w'} \pi'$, and the result follows.

Now, for each $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$, for every $w \in W'$, if $\models^{\mathscr{M}'}_{w} \pi \wedge \varphi_A \wedge \ell$, then $\models^{\mathscr{M}'}_{w} \varphi$, and then $\models^{\mathscr{M}}_{w} \varphi$. Because $\models^{\mathscr{M}} \varphi \rightarrow [a]\psi$, we have $\models^{\mathscr{M}}_{w'} \psi$ for every $w' \in W$ such that $(w, w') \in R_a$, and then $\models^{\mathscr{M}'}_{w'} \psi$ for all $w' \in W'$ such that $(w, w') \in R'_a \setminus R^{\varphi, \neg\psi}_{a}$. It remains to show that $\models^{\mathscr{M}'}_{w'} \ell$ for every $w' \in W'$ such that $(w, w') \in R^{\varphi, \neg\psi}_{a}$. Since $\mathscr{M}'$ is minimal, it is enough to show that $\models^{\mathscr{M}'}_{u'} \ell$ for every $\ell \in \mathfrak{Lit}$ such that $\models^{\mathscr{M}'}_{u} \pi \wedge \varphi_A \wedge \ell$. If $\ell \in \pi'$, the result follows. Otherwise, suppose $\not\models^{\mathscr{M}'}_{u'} \ell$. Then

- either $\neg\ell \in \pi'$, then $\pi'$ and $\ell$ are unsatisfiable, and in this case Algorithm 2 has not put the law $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$ in $\mathcal{T}'$, a contradiction;
- or $\neg\ell \in u' \setminus v'$. In this case, there is a valuation $u'' = (u' \setminus \{\neg\ell\}) \cup \{\ell\}$ such that $u'' \not\Vdash \psi$. We must have $u'' \in W'$, otherwise there will be $L' = \{\ell_i : \ell_i \in u''\}$ such that

$\mathcal{T} \models_{\overline{\mathsf{PDL}}} (\pi' \wedge \bigwedge_{\ell_i \in L'} \ell_i) \rightarrow \bot$, and, because $\mathcal{T}$ is modular, $\mathcal{S} \models_{\overline{\mathsf{CPL}}} (\pi' \wedge \bigwedge_{\ell_i \in L'} \ell_i) \rightarrow \bot$, and then Algorithm 2 has not put the law $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$ in $\mathcal{T}'$, a contradiction. Then $u'' \in W'$, and moreover $u'' \notin R_a^{\varphi,\neg\psi}(u)$, otherwise $\mathscr{M}'$ is not minimal. As $u'' \setminus u \subset u' \setminus u$, the only reason why $u'' \notin R_a^{\varphi,\neg\psi}(u)$ is that there is $\ell' \in u \cap u''$ such that $\models^{\mathscr{M}_i} \bigwedge_{\ell_j \in u} \ell_j \rightarrow [a]\neg\ell'$ for every $\mathscr{M}_i \in \mathcal{M}$ if and only if $\ell' \notin v'$ for any $v' \in base(\neg\psi, W')$ such that $v' \subseteq u''$. Clearly $\ell' = \ell$, and because $\ell \notin \pi'$, we have $\models^{\mathscr{M}_i} \bigwedge_{\ell_j \in u} \ell_j \rightarrow [a]\neg\ell$ for every $\mathscr{M}_i \in \mathcal{M}$. Then $\mathcal{T} \models_{\overline{\mathsf{PDL}}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell$, and Algorithm 2 has not put the law $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$ in $\mathcal{T}'$, contradiction.

Hence we have $\models^{\mathscr{M}'}_{w'} \psi \vee \ell$ for every $w' \in W'$ such that $(w, w') \in R'_a$.

Putting the above results together, we get $\models^{\mathscr{M}'} \mathcal{T}'$.

Let now $\Phi$ be some propositional $\varphi$. Then $\mathscr{M}' = \langle W', R' \rangle$, where $W \subseteq W'$, $R' = R$, is minimal w.r.t. $\preceq_{\mathscr{M}}$, i.e., $W'$ is a minimum superset of $W$ such that there is $u \in W'$ with $u \not\Vdash \varphi$. Because we have assumed the syntactical classical contraction operator is correct w.r.t. its semantics and is moreover minimal, then there must be $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$ such that $W' = val(\mathcal{S}^-)$. Hence $\models^{\mathscr{M}'} \mathcal{S}^-$.

As $R' = R$, every effect law of $\mathcal{T}$ remains true in $\mathscr{M}'$.

Now, let

$$\mathcal{T}' = \begin{array}{l} ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ \{(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ \{\neg\varphi \rightarrow [a]\bot\} \end{array}$$

(Clearly, $\mathcal{T}'$ is a theory produced by Algorithm 3.)

For every $(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top \in \mathcal{T}'$ and every $w \in W'$, if $\models^{\mathscr{M}'}_{w} \varphi_i \wedge \varphi$, then $R_a(w) \neq \emptyset$, because $\models^{\mathscr{M}}_{w} \varphi_i \rightarrow \langle a \rangle \top$. Given $\neg\varphi \rightarrow [a]\bot$, for every $w \in W'$, if $\models^{\mathscr{M}'}_{w} \neg\varphi$, then $w = u$, and $R_a(w) = \emptyset$.

Putting all these results together, we have $\models^{\mathscr{M}'} \mathcal{T}'$. ∎

## Appendix B: Proof of Theorem 5

**Lemma 2** *Let $\Phi$ be a law. If $\mathcal{T}$ is modular, then every $\mathcal{T}' \in \mathcal{T}^-_{\Phi}$ is modular.*

**Proof:** Let $\Phi$ be nonclassical, and suppose there is $\mathcal{T}' \in \mathcal{T}^-_{\Phi}$ such that $\mathcal{T}'$ is not modular. Then there is some $\varphi' \in \mathfrak{Fml}$ such that $\mathcal{T}' \models_{\overline{\mathsf{PDL}}} \varphi'$ and $\mathcal{S}' \not\models_{\overline{\mathsf{CPL}}} \varphi'$, where $\mathcal{S}'$ is the set of static laws in $\mathcal{T}'$. By Lemma 1, $\mathcal{T} \models_{\overline{\mathsf{PDL}}} \mathcal{T}'$, and then we have $\mathcal{T} \models_{\overline{\mathsf{PDL}}} \varphi'$. Because $\Phi$ is nonclassical, $\mathcal{S}' = \mathcal{S}$. Thus $\mathcal{S} \not\models_{\overline{\mathsf{CPL}}} \varphi'$, and hence $\mathcal{T}$ is not modular.

Let now $\Phi$ be some $\varphi \in \mathfrak{Fml}$. Then

$$\mathcal{T}' = \begin{array}{l} ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ \{(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ \{\neg\varphi \rightarrow [a]\bot\} \end{array}$$

for some $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$.

Suppose $\mathcal{T}$ is modular, and let $\varphi' \in \mathfrak{Fml}$ be such that $\mathcal{T}' \models_{\overline{\mathsf{PDL}}} \varphi'$ and $\mathcal{S}^- \not\models_{\overline{\mathsf{CPL}}} \varphi'$.

As $\mathcal{S}^- \not\models_{\overline{\mathsf{CPL}}} \varphi'$, there is $v \in val(\mathcal{S}^-)$ such that $v \not\Vdash \varphi'$. If $v \in val(\mathcal{S})$, then $\mathcal{S} \not\models_{\overline{\mathsf{CPL}}} \varphi'$, and as $\mathcal{T}$ is modular, $\mathcal{T} \not\models_{\overline{\mathsf{PDL}}} \varphi'$. By Lemma 1, $\mathcal{T} \models_{\overline{\mathsf{PDL}}} \mathcal{T}'$, and we have $\mathcal{T}' \not\models_{\overline{\mathsf{PDL}}} \varphi'$, a contradiction. Hence $v \notin val(\mathcal{S})$. Moreover, we must have $v \not\Vdash \varphi$, otherwise $\ominus$ has not worked as expected.

Let $\mathscr{M} = \langle W, R \rangle$ be such that $\models^{\mathscr{M}} \mathcal{T}'$. (We extend $\mathscr{M}$ to another model of $\mathcal{T}'$.) Let $\mathscr{M}' = \langle W', R' \rangle$ be such that $W' = W \cup \{v\}$ and $R' = R$. To show that $\mathscr{M}'$ is a model of $\mathcal{T}'$, it suffices to show that $v$ satisfies every law in $\mathcal{T}'$. As $v \in val(\mathcal{S}^-)$, $\models^{\mathscr{M}'}_{v} \mathcal{S}^-$. Given $\neg\varphi \rightarrow [a]\bot \in \mathcal{T}'$, as $v \not\Vdash \varphi$ and $R'_a(v) = \emptyset$, $\models^{\mathscr{M}'}_{v} \neg\varphi \rightarrow [a]\bot$. Now, for every $\varphi_i \rightarrow [a]\psi_i \in \mathcal{T}'$, if $\models^{\mathscr{M}'}_{v} \varphi_i$, then we trivially have $\models^{\mathscr{M}'}_{v'} \psi_i$ for every $v'$ such that $(v, v') \in R'_a$. Finally, given $(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top \in \mathcal{T}'$, as $v \not\Vdash \varphi$, the formula trivially holds in $v$. Hence $\models^{\mathscr{M}'} \mathcal{T}'$, and because there is $v \in W'$ such that $\not\models^{\mathscr{M}'}_{v} \varphi'$, we have $\mathcal{T}' \not\models_{\overline{\mathsf{PDL}}} \varphi'$, a contradiction. Hence for all $\varphi' \in \mathfrak{Fml}$ such that $\mathcal{T}' \models_{\overline{\mathsf{PDL}}} \varphi'$, $\mathcal{S}^- \models_{\overline{\mathsf{CPL}}} \varphi'$, and then $\mathcal{T}'$ is modular. ∎

For the proof of the following three lemmas, please refer to (Varzinczak 2008a).

**Lemma 3** *If $\mathscr{M}_{big} = \langle W_{big}, R_{big} \rangle$ is a model of $\mathcal{T}$, then for every $\mathscr{M} = \langle W, R \rangle$ such that $\models^{\mathscr{M}} \mathcal{T}$ there is a minimal (w.r.t. set inclusion) extension $R' \subseteq R_{big} \setminus R$ such that $\mathscr{M}' = \langle val(\mathcal{S}), R \cup R' \rangle$ is a model of $\mathcal{T}$.*

**Lemma 4** *Let $\mathcal{T}$ be modular, and $\Phi$ be a law. Then $\mathcal{T} \models_{\overline{\mathsf{PDL}}} \Phi$ if and only if every $\mathscr{M}' = \langle val(\mathcal{S}), R' \rangle$ such that $\models^{\langle W, R \rangle} \mathcal{T}$ and $R \subseteq R'$ is a model of $\Phi$.*

**Lemma 5** *Let $\mathcal{T}$ be modular, $\Phi$ a law, and $\mathcal{T}' \in \mathcal{T}^-_{\Phi}$. If $\mathscr{M}' = \langle val(\mathcal{S}'), R' \rangle$ is a model of $\mathcal{T}'$, then there is $\mathcal{M} = \{\mathscr{M} : \mathscr{M} = \langle val(\mathcal{S}), R \rangle$ and $\models^{\mathscr{M}} \mathcal{T}\}$ such that $\mathscr{M}' \in \mathcal{M}'$ for some $\mathcal{M}' \in \mathcal{M}^-_{\Phi}$.*

**Proof of Theorem 5**

From the hypothesis that $\mathcal{T}$ is modular and Lemma 2, $\mathcal{T}'$ is also a modular theory. Then $\mathscr{M}' = \langle val(\mathcal{S}'), R \rangle$ is a model of $\mathcal{T}'$, by Lemmas 3 and 4. From this and Lemma 5 the result follows. ∎

# 2008

This paper was published in the *Proceedings, Eleventh International Conference on Principles of Knowledge Representation and Reasoning*, edited by Gerhard Brewka and Jérôme Lang (Menlo Park, California: AAAI Press).

The 2008 KR conference was held in Sydney, Australia, 16–19 September, 2008.