



## Learning Actions Models: Qualitative Approach

**Bolander, Thomas; Gierasimczuk, Nina**

*Published in:*

Proceedings of the 5th International Workshop on Logic, Rationality, and Interaction (LORI 2015)

*Link to article, DOI:*

[10.1007/978-3-662-48561-3\\_4](https://doi.org/10.1007/978-3-662-48561-3_4)

*Publication date:*

2015

*Document Version*

Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*

Bolander, T., & Gierasimczuk, N. (2015). Learning Actions Models: Qualitative Approach. In W. van der Hoek, W. H. Holliday, & W. Wang (Eds.), *Proceedings of the 5th International Workshop on Logic, Rationality, and Interaction (LORI 2015)* (pp. 40-52). Springer. [https://doi.org/10.1007/978-3-662-48561-3\\_4](https://doi.org/10.1007/978-3-662-48561-3_4)

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Learning Actions Models: Qualitative Approach

Thomas Bolander and Nina Gierasimczuk

**Abstract.** In dynamic epistemic logic, actions are described using action models. In this paper we introduce a framework for studying learnability of action models from observations. We present first results concerning propositional action models. First we check two basic learnability criteria: finite identifiability (conclusively inferring the appropriate action model in finite time) and identifiability in the limit (inconclusive convergence to the right action model). We show that deterministic actions are finitely identifiable, while non-deterministic actions require more learning power—they are identifiable in the limit. We then move on to a particular learning method, which proceeds via restriction of a space of events within a learning-specific action model. This way of learning closely resembles the well-known update method from dynamic epistemic logic. We introduce several different learning methods suited for finite identifiability of particular types of deterministic actions.

Dynamic epistemic logic (DEL) allows analyzing knowledge change in a systematic way. The static component of a situation is represented by an epistemic model, while the structure of the dynamic component is encoded in an *action model*. An action model can be applied to the epistemic model via so-called *product update* operation, resulting in a new up-to-date epistemic model of the situation after the action has been executed. A language, interpreted on epistemic models, allows expressing conditions under which an action takes effect (so-called preconditions), and the effects of such actions (so-called postconditions). This setting is particularly useful for modeling the process of epistemic planning (see [7,1]): one can ask which sequence of actions should be executed in order for a given epistemic formula to hold in the epistemic model after the actions are executed.

The purpose of this paper is to investigate possible learning mechanisms involved in discovering the ‘internal structure’ of actions on the basis of their executions. In other words, we are concerned with qualitative learning of action models on the basis of observations of pairs of the form (initial state, resulting state). We analyze learnability of action models in the context of two learning conditions: finite identifiability (conclusively inferring the appropriate action model in finite time) and identifiability in the limit (inconclusive convergence to the right action model). The paper draws on the results from formal learning theory applied to DEL (see [11,13,12]).

Learning of action models is highly relevant in the context of epistemic planning. A planning agent might not initially know the effects of her actions, so she will initially not be able to plan to achieve any goals. However, if she can learn the relevant action models through observing the effect of the actions (either

by executing the actions herself, or by observing other agents), she will eventually learn how to plan. Our ultimate goal is to integrate learning of actions into (epistemic) planning agents. In this paper, we seek to lay the foundations for this goal by studying learnability of action models from streams of observations.

The structure of the paper is as follows. In Section 1 we recall the basic concepts and notation concerning action models and action types in DEL. In Section 2 we specify our learning framework and provide general learnability results. In Section 3 we study particular learning functions, which proceed via updating action models with new information. Finally, in Section 4 we indicate how to lift our results from the level of individual action learning to that of action library learning. In the end we briefly discuss related and further work.

## 1 Languages and action types

Let us first present the basic notions required for the rest of the article (see [6,8] for more details). Following the conventions of automated planning, we take the set of atomic propositions and the set of actions to be finite. Given a finite set  $P$  of atomic propositions, we define the (single-agent) *epistemic language* over  $P$ ,  $\mathcal{L}_{\text{epis}}(P)$ , by the following BNF:  $\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid K\phi$ , where  $p \in P$ . The language  $\mathcal{L}_{\text{prop}}(P)$  is the propositional sublanguage without the  $K\phi$  clause. When  $P$  is clear from the context, we write  $\mathcal{L}_{\text{epis}}$  and  $\mathcal{L}_{\text{prop}}$  instead of  $\mathcal{L}_{\text{epis}}(P)$  and  $\mathcal{L}_{\text{prop}}(P)$ , respectively. By means of the standard abbreviations we introduce the additional symbols  $\rightarrow$ ,  $\vee$ ,  $\leftrightarrow$ ,  $\perp$ , and  $\top$ .

**Definition 1 (Epistemic models and states).** *An epistemic model over a set of atomic propositions  $P$  is  $\mathcal{M} = (W, R, V)$ , where  $W$  is a finite set of worlds,  $R \subseteq W \times W$  is an equivalence relation, called the indistinguishability relation, and  $V : P \rightarrow \mathcal{P}(W)$  is a valuation function. An epistemic state is a pointed epistemic model  $(\mathcal{M}, w)$  consisting of an epistemic model  $\mathcal{M} = (W, R, V)$  and a distinguished world  $w \in W$  called the actual world.*

A *propositional state* (or simply *state*) over  $P$  is a subset of  $P$  (or, equivalently, a propositional valuation  $\nu : P \rightarrow \{0, 1\}$ ). We identify propositional states and singleton epistemic models via the following canonical isomorphism. A propositional state  $s \subseteq P$  is isomorphic to the epistemic model  $\mathcal{M} = (\{w\}, \{(w, w)\}, V)$  where  $V(p) = \{w\}$  if  $p \in s$  and  $V(p) = \emptyset$  otherwise. Truth in epistemic states  $(\mathcal{M}, w)$  with  $\mathcal{M} = (W, R, V)$  (and hence propositional states) is defined as usual and hence omitted.

Dynamic epistemic logic (DEL) introduces the concept of an action model for modelling the changes to states brought about by the execution of actions [6]. We here use a variant that includes postconditions [8,7], which means that actions can have both epistemic effects (changing the beliefs of agents) and ontic effects (changing the factual states of affairs).

**Definition 2 (Action models).** *An action model over a set of atomic propositions  $P$  is  $\mathcal{A} = (E, Q, \text{pre}, \text{post})$ , where  $E$  is a finite set of events;  $Q \subseteq E \times E$*

is an equivalence relation called the indistinguishability relation;  $pre : E \rightarrow \mathcal{L}_{epis}(P)$  assigns to each event a precondition;  $post : E \rightarrow \mathcal{L}_{prop}(P)$  assigns to each event a postcondition. Postconditions are conjunctions of literals (atomic propositions and their negations) or  $\top$ .<sup>1</sup>  $dom(\mathcal{A}) = E$  denotes the domain of  $\mathcal{A}$ . The set of all action models over  $P$  is denoted  $\mathbf{Actions}(P)$ .

Intuitively, events correspond to the ways in which an action changes the epistemic state, and the indistinguishability relation codes (an agent's) ability to recognize the difference between those different ways. In an event  $e$ ,  $pre(e)$  specifies what conditions have to be satisfied for it to take effect, and  $post(e)$  specifies its outcome.

*Example 1.* Consider the action of tossing a coin. It can be represented by the following action model ( $h$  means that the coin is facing heads up):

$$\mathcal{A} = \begin{array}{cc} \bullet & \bullet \\ e_1: \langle \top, h \rangle & e_2: \langle \top, \neg h \rangle \end{array}$$

We label each event by a pair whose first argument is the event's precondition while the second is its postcondition. Hence, formally we have  $\mathcal{A} = (E, Q, pre, post)$  with  $E = \{e_1, e_2\}$ ,  $Q$  is the identity on  $E$ ,  $pre(e_1) = pre(e_2) = \top$ ,  $post(e_1) = h$  and  $post(e_2) = \neg h$ . The action model encodes that tossing the coin will either make  $h$  true ( $e_1$ ) or  $h$  false ( $e_2$ ).

**Definition 3 (Product update).** Let  $\mathcal{M} = (W, R, V)$  and  $\mathcal{A} = (E, Q, pre, post)$  be an epistemic model and action model (over a set of atomic propositions  $P$ ), respectively. The product update of  $\mathcal{M}$  with  $\mathcal{A}$  is the epistemic model  $\mathcal{M} \otimes \mathcal{A} = (W', R', V')$ , where  $W' = \{(w, e) \in W \times E \mid (\mathcal{M}, w) \models pre(e)\}$ ;  $R' = \{((w, e), (v, f)) \in W' \times W' \mid wRv \text{ and } eQf\}$ ;  $V'(p) = \{(w, e) \in W' \mid post(e) \models p \text{ or } ((\mathcal{M}, w) \models p \text{ and } post(e) \not\models \neg p)\}$ . For  $e \in dom(\mathcal{A})$ , we define  $\mathcal{M} \otimes e = \mathcal{M} \otimes (\mathcal{A} \upharpoonright \{e\})$ .

The product update  $\mathcal{M} \otimes \mathcal{A}$  represents the result of executing the action  $\mathcal{A}$  in the state(s) represented by  $\mathcal{M}$ .

*Example 2.* Continuing Example 1, consider a situation of an agent seeing a coin lying heads-up, i.e., the singleton epistemic state  $\mathcal{M} = (\{w\}, \{w, w\}, V)$  with  $V(h) = \{w\}$ . Let us now calculate the result of executing the coin toss in this model.

$$\mathcal{M} \otimes \mathcal{A} = \begin{array}{cc} \bullet & \bullet \\ (w_1, e_1): h & (w_1, e_2): \neg h \end{array}$$

Here each world is labelled by the propositions being true at the world.

We say that two action models  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are *equivalent*, written  $\mathcal{A}_1 \equiv \mathcal{A}_2$ , if for any epistemic model  $\mathcal{M}$ ,  $\mathcal{M} \otimes \mathcal{A}_1 \dot{\sim} \mathcal{M} \otimes \mathcal{A}_2$ , where  $\dot{\sim}$  denotes standard bisimulation on epistemic models [17].

<sup>1</sup> We are here using the postcondition conventions from [7], which are slightly non-standard. Any action model with standard postconditions can be turned into one of our type, but it might become exponentially larger in the process [8, 7].

### 1.1 Action types

We can identify a number of different action types.

**Definition 4 (Action types).** *An action model  $\mathcal{A} = (E, Q, pre, post)$  is:*

- atomic if  $|E| = 1$ .
- deterministic if all preconditions are mutually inconsistent, that is,  $\models pre(e) \wedge pre(f) \rightarrow \perp$  for all distinct  $e, f \in E$ .
- fully observable if  $Q$  is the identity relation on  $E$ . Otherwise it is partially observable.
- precondition-free if  $pre(e) = \top$  for all  $e \in E$ .
- propositional if  $pre(e) \in \mathcal{L}_{prop}$  for all  $e \in E$ .
- universally applicable if  $\models \bigvee_{e \in E} pre(e)$ .
- normal if for all propositional literals  $l$  and all  $e \in E$ ,  $pre(e) \models l$  implies  $post(e) \not\models l$ .
- with basic preconditions if all  $pre(e)$  are conjunctions of literals (propositional atoms and their negations).
- with maximal preconditions if all  $pre(e)$  are maximally consistent conjunctions of literals (i.e., preconditions are conjunctions of literals in which each atomic proposition  $p$  occurs exactly once, either as  $p$  or as  $\neg p$ ).

Some of the notions defined above are known from existing literature [7,8,16]. The newly introduced notions are precondition-free, universally applicable, and normal actions, as well as actions with basic preconditions. Note that action types interact with each other, atomic actions are automatically both deterministic and fully observable, and precondition-free actions can only be deterministic if atomic.<sup>2</sup>

In the remainder of this section we set a uniform representation of action models that we will later on use in learning methods. We also specify and justify the restrictions we impose on action models.

**Propositionality** In this paper we are concerned with product updates of *propositional* states with *propositional* action models. Let  $s$  denote a propositional state over  $P$ , and let  $\mathcal{A} = (E, Q, pre, post)$  be any propositional action model. Using the definition above and the canonical isomorphism between propositional states and singleton epistemic states, we get that  $s \otimes \mathcal{A}$  is isomorphic to the epistemic model  $(W', R', V')$ , where  $W' = \{e \in E \mid s \models pre(e)\}$ ,  $R' = \{(e, f) \in W' \times W' \mid eQf\}$ ,  $V'(p) = \{e \in W' \mid post(e) \models p \text{ or } (s \models p \text{ and } post(e) \not\models \neg p)\}$ . If  $\mathcal{A}$  is fully observable, then the indistinguishability of  $s \otimes \mathcal{A}$  is the identity relation. This means that we can think of  $s \otimes \mathcal{A}$  as a set of propositional states (via the canonical isomorphism between singleton epistemic models and propositional states). In this case we write  $s' \in s \otimes \mathcal{A}$  to mean that  $s'$  is one of the propositional states in  $s \otimes \mathcal{A}$ . When  $\mathcal{A}$  is atomic we have  $s \otimes a = s'$  for some propositional state  $s'$  (using again the canonical isomorphism).

<sup>2</sup> The actions considered in propositional STRIPS planning (called *set-theoretic planning* in [9]) correspond to epistemic actions that are atomic and have basic post-conditions.

*Example 3.* Consider the action model  $\mathcal{A}$  of Example 1 (the coin toss). It is a precondition-free, fully observable, non-deterministic action. Consider an initial propositional state  $s = \{h\}$ . Then  $s \otimes \mathcal{A}$  is the epistemic model of Example 2. It has two worlds, one in which  $h$  is true, and another in which  $h$  is false. So we have  $\emptyset, \{h\} \in s \otimes \mathcal{A}$ , i.e., the outcome of tossing the coin is either the propositional state where  $h$  is false ( $\emptyset$ ) or the one where  $h$  is true ( $\{h\}$ ).

**Basic preconditions and normality** When preconditions are basic, pre- and postconditions are of the same simple normal form, they are conjunctions of literals. Below we show that any propositional action model can be turned into an action model having this normal form. We also show that we can ensure all action models to be normal.

**Proposition 1.** *Any propositional action model is equivalent to a normal action model with basic preconditions.*

*Proof sketch.* Take a propositional action model  $(E, Q, pre, post)$ . We first make the preconditions basic in the following way. Take any event  $e \in E$  with precondition  $\phi$ . Turn  $\phi$  into disjunctive normal form  $\bigvee_{i \in I} \bigwedge_{j \in J} p_{ij}$ . Then replace  $e$  by a set of events  $e_i, i \in I$ , where  $post(e_i) = post(e)$  and  $pre(e_i) = \bigwedge_{j \in J} p_{ij}$ . Each  $e_i$  is connected by a  $Q$ -edge to every event  $e$  was originally connected to. This is done for each event  $e \in E$ . It is easy to see that the resulting action model  $(E', Q', pre', post')$  has basic preconditions and is equivalent to the original one.

We now “normalise”  $post'$  into a new mapping  $post''$  in order to obtain an equivalent normal action model  $(E', Q', pre', post'')$ . Note that since the action model  $(E', Q', pre', post')$  has basic preconditions, the normality condition can be expressed in a particularly simple way: for all literals  $l$  and all  $e \in E$ , if  $l$  is a conjunct of  $pre(e)$  then it is not a conjunct of  $post(e)$ . For each event  $e \in E$ , we now define  $post''(e)$  from  $post'(e)$  by deleting each conjunct of  $post'(e)$  which is also a conjunct in  $pre'(e)$ . It is easy to see that this gives an equivalent action model: consider an event  $e$  and a literal  $l$  which is both a conjunct of  $pre'(e)$  and  $post'(e)$ . Since  $l$  is a conjunct of  $pre'(e)$ ,  $l$  has to be true for  $e$  to occur. Since  $l$  is a conjunct of  $post'(e)$ ,  $l$  will also be true after the event  $e$  has occurred. Hence, the event  $e$  does not affect the truth value of  $l$ , and we get an equivalent event by removing  $l$  from the postcondition.

In this paper we are only going to be concerned with propositional actions, and so, due to Proposition 1, we can restrict attention to normal actions having basic preconditions.

**Universal applicability** The condition for being universally applicable intuitively means that the action specifies an outcome no matter what state it is applied to. In this paper we will only be concerned with universally applicable action models. To understand the reason for this restriction consider the example of an action *open\_door* with singleton action model  $\langle \neg open \wedge \neg locked, open \rangle$ , i.e., if the door is currently closed and unlocked, performing *open\_door* will open it.

This action model does not specify what happens if an agent attempts *open\_door* when the door is either already open or is locked. We can easily fix this by adding another event to the action model,  $\langle open \vee locked, \top \rangle$ , expressing that if one tries to open it when locked or already open, nothing happens. More generally, any action model  $\mathcal{A} = (E, Q, pre, post)$  which is not universally applicable can be turned into a universally applicable action model by adding the following event:  $\langle \neg \vee_{e \in E} pre(e), \top \rangle$ . If an agent is learning results of an action, she should in any possible state be able to attempt executing the action, and hence the action model should specify an outcome of this attempt. For this reason, we require universal applicability.

## 2 Learning action models

First we will focus on learning an individual action, i.e., inferring semantics of a single action name. The semantics of an action name is an action model. In the following we will use the expressions *action* and *action model* interchangeably. Below we will first present general results on learnability of various types of action models, and then, in Section 3, we study particular learning methods and exemplify them.

We are concerned with learning fully observable actions (action models). Partially observable actions are generally not learnable in the strict sense to be defined below. Consider for instance an agent trying to learn an action that controls the truth value of a proposition  $p$ , but where the agent cannot observe  $p$  (events making  $p$  true and events making  $p$  false are indistinguishable). Then clearly there is no way for that agent to learn exactly how the action works. The case of fully observable actions is much simpler. If initially the agent has no uncertainty, her “belief state” can be represented by a propositional state. Executing any sequence of fully observable actions will then again lead to a propositional state. So in the case of fully observable actions, we can assume actions to make transitions between propositional states.

For the rest of this section, except in examples, we fix a set  $P$  of atomic propositions.

**Definition 5.** A stream  $\mathcal{E}$  is an infinite sequence of pairs  $(s, s')$  of propositional states over  $P$ , i.e.,  $\mathcal{E} \in (\mathcal{P}(P) \times \mathcal{P}(P))^\omega$ . The elements  $(s, s')$  of  $\mathcal{E}$  are called observations. Let  $\mathbb{N} := \mathbb{N}^+ \cup \{0\}$ , let  $\mathcal{E}$  be a stream over  $P$ , and let  $s, t \in \mathcal{P}(P)$ .  $\mathcal{E}_n$  stands for the  $n$ -th observation in  $\mathcal{E}$ .  $\mathcal{E}[n]$  stands for the initial segment of  $\mathcal{E}$  of length  $n$ , i.e.,  $\mathcal{E}_0, \dots, \mathcal{E}_{n-1}$ .  $\text{set}(\mathcal{E}) := \{(x, y) \mid (x, y) \text{ is an element of } \mathcal{E}\}$  stands for the set of all observations in  $\mathcal{E}$ ; we similarly define  $\text{set}(\mathcal{E}[n])$  for initial segments of streams.

**Definition 6.** Let  $\mathcal{E}$  be a stream over  $P$  and  $\mathcal{A}$  a fully observable action model over  $P$ . The stream  $\mathcal{E}$  is *sound* with respect to  $\mathcal{A}$  if for all  $(s, s') \in \text{set}(\mathcal{E})$ ,  $s' \in s \otimes \mathcal{A}$ . The stream  $\mathcal{E}$  is *complete* with respect to  $\mathcal{A}$  if for all  $s \subseteq P$  and all  $s' \in s \otimes \mathcal{A}$ ,  $(s, s') \in \text{set}(\mathcal{E})$ . In this paper we always assume the streams to be sound and complete. For brevity, if  $\mathcal{E}$  is sound and complete wrt  $\mathcal{A}$ , we will write: ‘ $\mathcal{E}$  is for  $\mathcal{A}$ ’.

**Definition 7 (Learning function).** A learning function is a computable  $L : (\mathcal{P}(P) \times \mathcal{P}(P))^* \rightarrow \text{Actions}(P) \cup \{\uparrow\}$ .

In other words, a learning function takes a finite sequence of observations (pairs of propositional states) and outputs an action model or a symbol corresponding to ‘undecided’.

We will study two types of learning: finite identifiability and identifiability in the limit. First let us focus on *finite identifiability*. Intuitively, finite identifiability corresponds to conclusive learning: upon observing some finite amount of action executions the learning function outputs, with certainty, a correct model for the action in question (up to equivalence). This certainty can be expressed in terms of the function being once-defined: it is allowed to output an action model only once, there is no chance of correction later on. Formally, we say that a learning function  $L$  is *(at most) once defined* if for any stream  $\mathcal{E}$  for an action over  $P$  and  $n, k \in \mathbb{N}$  such that  $n \neq k$ , we have that  $L(\mathcal{E}[n]) = \uparrow$  or  $L(\mathcal{E}[k]) = \uparrow$ .

**Definition 8.** Let  $\mathcal{X}$  be a class of action models and  $\mathcal{A} \in \mathcal{X}$ ,  $L$  be a learning function, and  $\mathcal{E}$  be a stream. We say that:

1.  $L$  finitely identifies  $\mathcal{A}$  on  $\mathcal{E}$  if  $L$  is once-defined and there is an  $n \in \mathbb{N}$  s.t.  $L(\mathcal{E}[n]) \equiv \mathcal{A}$ .
2.  $L$  finitely identifies  $\mathcal{A}$  if  $L$  finitely identifies  $\mathcal{A}$  on every stream for  $\mathcal{A}$ .
3.  $L$  finitely identifies  $\mathcal{X}$  if  $L$  finitely identifies every  $\mathcal{A} \in \mathcal{X}$ .
4.  $\mathcal{X}$  is finitely identifiable if there is a function  $L$  which finitely identifies  $\mathcal{X}$ .

The following definition and theorem are adapted from [15,14,13].

**Definition 9.** Let  $\mathcal{X} \subseteq \text{Actions}(P)$ . A set  $D_{\mathcal{A}} \subseteq \mathcal{P}(P) \times \mathcal{P}(P)$  is a definite finite tell-tale set (DFTT) for  $\mathcal{A}$  in  $\mathcal{X}$  if

1.  $D_{\mathcal{A}}$  is sound for  $\mathcal{A}$  (i.e., for all  $(s, s') \in D_{\mathcal{A}}$ ,  $s' \in s \otimes \mathcal{A}$ ),
2.  $D_{\mathcal{A}}$  is finite, and
3. for any  $\mathcal{A}' \in \mathcal{X}$ , if  $D_{\mathcal{A}}$  is sound for  $\mathcal{A}'$ , then  $\mathcal{A} \equiv \mathcal{A}'$ .

**Lemma 1.**  $\mathcal{X}$  is finitely identifiable iff there is an effective procedure  $D : \mathcal{X} \rightarrow \mathcal{P}(\mathcal{P}(P) \times \mathcal{P}(P))$ , given by  $\mathcal{A} \mapsto D_{\mathcal{A}}$ , that on input  $\mathcal{A}$  produces a definite finite tell-tale of  $\mathcal{A}$ .

*Proof.*  $[\Rightarrow]$  Assume that  $\mathcal{X}$  is finitely identifiable. Then there is a computable function  $L$  that finitely identifies  $\mathcal{X}$ . We use that function to define  $D$ . Once the learning function  $L$  identifies an action  $\mathcal{A}$  it has to give it as a definite output, and this will happen for some  $\mathcal{E}[n]$ . We then set  $D(\mathcal{A}) = \text{set}(\mathcal{E}[n])$ . It is easy to check that such  $D(\mathcal{A})$  is a definite tell-tale set.  $[\Leftarrow]$  Assume that there is an effective procedure  $D : \mathcal{X} \rightarrow \mathcal{P}(\mathcal{P}(P) \times \mathcal{P}(P))$ , that on input  $\mathcal{A}$  produces a definite finite tell-tale of  $\mathcal{A}$ . Take an enumeration of  $\mathcal{X}$  and take any  $\mathcal{A} \in \mathcal{X}$  and any  $\mathcal{E}$  for  $\mathcal{A}$ . We use  $D$  to define the learning function. At each step  $n \in \mathbb{N}$ ,  $L$  compares  $\mathcal{E}[n]$  with  $D(\mathcal{A}_1), \dots, D(\mathcal{A}_n)$ . Once, at some step  $\ell \in \mathbb{N}$ , it finds  $\mathcal{A}_k$  such that  $D(\mathcal{A}_k) \subseteq \text{set}(\mathcal{E}[\ell])$ , it outputs  $\mathcal{A}_k$ . It is easy to verify that then  $\mathcal{A}_k \equiv \mathcal{A}$ .



In other words, the finite set of observations  $D_{\mathcal{A}}$  is consistent with only one action  $\mathcal{A}$  in the class (up to equivalence of actions).  $D$  is a computable function that gives a  $D_{\mathcal{A}}$  for any action  $\mathcal{A}$ .

**Theorem 1.** *For any finite set of propositions  $P$  the set of (fully observable) deterministic propositional actions over  $P$  is finitely identifiable.*

*Proof.* We use Lemma 1, and hence define  $D$ :  $D(\mathcal{A}) = \{(s, s') \mid s \otimes \mathcal{A} = s', \text{ where } s, s' \in \mathcal{P}(P)\}$ . Let us check that indeed  $D(\mathcal{A})$  is a DFTT for  $\mathcal{A}$ . We need to show conditions 1, 2 and 3 of Definition 9 for  $D(\mathcal{A})$ . 1:  $D(\mathcal{A})$  is sound for  $\mathcal{A}$ , trivially. 2:  $D(\mathcal{A})$  is finite, because  $P$  is finite. 3: Let us take any propositional action  $\mathcal{A}'$  such that  $D(\mathcal{A})$  is sound for  $\mathcal{A}'$ . This means, by the definition of  $D$  above and the fact that  $\mathcal{A}$  and  $\mathcal{A}'$  are deterministic, that for all propositional states  $s, s'$  over  $P$ , if  $s \otimes \mathcal{A} = s'$  then  $s \otimes \mathcal{A}' = s'$ . It follows that  $s \otimes \mathcal{A} = s' \otimes \mathcal{A}'$  for all propositional states  $s$ , and hence  $\mathcal{A} \equiv \mathcal{A}'$  (since  $\mathcal{A}$  and  $\mathcal{A}'$  are propositional). Finally,  $D$  is computable because  $P$  is finite.

*Example 4.* Theorem 1 shows that deterministic actions are finitely identifiable. We will now show that this does not carry over to non-deterministic actions, that is, non-deterministic actions are in general not finitely identifiable. Consider the action of tossing a coin, given by the action model  $\mathcal{A}$  in Example 1. If in fact the coin is fake and it will always land tails (so it only consists of the event  $e_2$ ), in no finite amount of tosses the agent can exclude that the coin is fair, and that heads will start appearing in the long run (that  $e_1$  will eventually occur). So the agent will never be able to say “stop” and declare the action model to only consist of  $e_2$ . This argument can be generalised, leading to the theorem below.

**Theorem 2.** *For any finite set of propositions  $P$  the set of arbitrary (including non-deterministic) fully observable propositional actions over  $P$  is not finitely identifiable.*

*Proof.* Assume that the set of arbitrary propositional actions over  $A$  is finitely identifiable. Then there is a learning function  $L$  that finitely identifies it. Among such actions we will have two,  $\mathcal{A}$  and  $\mathcal{A}'$ , such that  $\mathcal{A}' = \mathcal{A} \upharpoonright D(\mathcal{A}')$ .<sup>3</sup> Let us now construct a stream  $\mathcal{E}$  on which  $L$  fails to finitely identify one of them. Let the  $\mathcal{E}$  start with enumerating all pairs of propositional states that are sound for the smaller action,  $\mathcal{A}'$ , and keep repeating this pattern. Since this is a stream for  $\mathcal{A}'$  indeed the learning function has to at some point output an equivalent of  $\mathcal{A}'$  (otherwise it fails to finitely identify  $\mathcal{A}'$ , which leads to contradiction). Assume that this happens at some stage  $n \in \mathbb{N}$ . Now, observe that  $\mathcal{E}[n]$  is sound with respect to  $\mathcal{A}$  too, so starting at the stage  $n + 1$  let us make  $\mathcal{E}$  enumerate the rest of remaining pairs of propositional states consistent with  $\mathcal{A}$ . That means that there is a stream  $\mathcal{E}$  for  $\mathcal{A}$  on which  $L$  does not finitely identify  $\mathcal{A}$ . Contradiction.

<sup>3</sup> For any action model  $\mathcal{A} = (E, Q, pre, post)$  and any subset  $E' \subseteq E$  we define  $a \upharpoonright E'$  as the restriction of  $\mathcal{A}$  to the domain  $E'$ , that is,  $a \upharpoonright E' = (E', Q', pre', post')$  where  $Q' = Q \cap (E')^2$ ,  $pre' = pre \upharpoonright E'$  and  $post' = post \upharpoonright E'$ .

A weaker condition of learnability, *identifiability in the limit*, allows widening the scope of learnable actions, to cover also the case of arbitrary actions. Identifiability in the limit requires that the learning function after observing some finite amount of action executions outputs a correct model (up to equivalence) for the action in question and then forever keeps to this answer (up to equivalence) in all the outputs to follow. This type of learning can be called ‘inconclusive’, because certainty cannot be achieved in finite time.

**Definition 10.** *Let  $\mathcal{X}$  be a class of action models and  $\mathcal{A} \in \mathcal{X}$ ,  $L$  be a learning function, and  $\mathcal{E}$  be a stream. We say that:*

1.  *$L$  identifies  $\mathcal{A}$  on  $\mathcal{E}$  in the limit if there is  $k \in \mathbb{N}$  such that for all  $n \geq k$ ,  $L(\mathcal{E}[n]) \equiv \mathcal{A}$ .*
2.  *$L$  identifies  $\mathcal{A}$  in the limit if  $L$  identifies  $\mathcal{A}$  in the limit on every  $\mathcal{E}$  for  $\mathcal{A}$ .*
3.  *$L$  identifies  $\mathcal{X}$  in the limit if  $L$  identifies in the limit every  $\mathcal{A} \in \mathcal{X}$ .*
4.  *$\mathcal{X}$  is identifiable in the limit if there is an  $L$  which identifies  $\mathcal{X}$  in the limit.*

The following theorem is adapted from [2].

**Theorem 3.** *For any finite set of propositions  $P$  the set of (fully observable) propositional actions over  $P$  is identifiable in the limit.*

*Proof.* The argument is similar to the proof of Theorem 1. Analogously to the concept of definite finite tell-tale set, we define a weaker notion of finite tell-tale set (FTT). Let  $P$  be a set of propositions and let  $\mathcal{X} \subseteq \text{Actions}(P)$ . A set  $D_{\mathcal{A}} \subseteq \mathcal{P}(P) \times \mathcal{P}(P)$  is a finite tell-tale set (FTT) for  $\mathcal{A}$  in  $\mathcal{X}$  if: 1.  $D_{\mathcal{A}}$  is sound for  $\mathcal{A}$  (i.e., for all  $(s, s') \in D_{\mathcal{A}}$ ,  $s' \in s \otimes \mathcal{A}$ ); 2.  $D_{\mathcal{A}}$  is finite, and 3. for any  $\mathcal{A}' \in \mathcal{X}$ , if  $D_{\mathcal{A}}$  is sound for  $\mathcal{A}'$ , then  $\mathcal{A} \equiv \mathcal{A}' \upharpoonright X$ , where  $X \subseteq \text{dom}(\mathcal{A}')$ .

Similarly to the argument for Lemma 1 one can show that  $\mathcal{X}$  is identifiable in the limit iff there is an effective procedure  $D : \mathcal{X} \rightarrow \mathcal{P}(\mathcal{P}(P) \times \mathcal{P}(P))$ , given by  $\mathcal{A} \mapsto D_{\mathcal{A}}$ , that on input  $\mathcal{A}$  enumerates a finite tell-tale of  $\mathcal{A}$ . We will omit the proof for the sake of brevity.

Now it is enough to show that indeed such a function  $D$  can be given for the set of arbitrary (fully observable, propositional) actions over  $P$ . Define  $D(\mathcal{A}) = \{(s, s') \mid s' \in s \otimes \mathcal{A}, \text{ where } s, s' \in \mathcal{P}(P)\}$ . Let us check that indeed  $D(\mathcal{A})$  is a FTT for  $\mathcal{A}$ . 1:  $D(\mathcal{A})$  is sound for  $\mathcal{A}$ , trivially. 2:  $D(\mathcal{A})$  is finite, because  $P$  is finite. 3: Let us take any propositional action  $\mathcal{A}'$  such that  $D(\mathcal{A})$  is sound for  $\mathcal{A}'$ . This means, by the definition of  $D$  above that for all propositional states  $s, s'$  over  $P$ , if  $s' \in s \otimes \mathcal{A}$  then  $s' \in s \otimes \mathcal{A}'$ . This implies that  $s \otimes \mathcal{A}$  is a submodel of  $s \otimes \mathcal{A}'$  for all propositional states  $s$ , and hence that  $\mathcal{A}$  is equivalent to a submodel of  $\mathcal{A}'$  (since actions are propositional).

Finally, again  $D$  is computable because  $P$  is finite.

Having established the general facts about finite identifiability and identifiability in the limit of propositional fully-observable actions, we will now turn to studying particular learning methods suited for such learning conditions.

### 3 Learning actions via update

Standard DEL, and in particular public announcement logic, deals with learning within epistemic models. If an agent is in a state described by an epistemic model  $\mathcal{M}$  and learns from a reliable source, that  $\phi$  is true, her state will be updated by eliminating all the worlds where  $\phi$  is false. That is, the model  $\mathcal{M}$  will be restricted to the worlds where  $\phi$  is true. This can also be expressed in terms of action models, where the learning of  $\phi$  corresponds to taking the product update of  $\mathcal{M}$  with the event model  $\langle \phi, \top \rangle$  (public announcement of  $\phi$ ).

Now we turn to learning *actions* rather than learning *facts*. Actions are represented by action models, so to learn an action means to infer the action model that describes it. Consider again the action model  $\mathcal{A}$  of Example 1. The coin toss is non-deterministic and fully observable: either  $h$  or  $\neg h$  will non-deterministically be made true and the agent is able to distinguish these two outcomes (no edge between  $e_1$  and  $e_2$ ). However, we can also think of  $\mathcal{A}$  as the hypothesis space of a *deterministic* action, that is, the action  $\mathcal{A}$  is in fact deterministically making  $h$  true or false, but the agent is currently uncertain about which one it is. Given the prior knowledge that the action in question must be deterministic, learning the action could proceed in a way analogous to that of update in the usual DEL setting.

It could for instance be that the agent knows that the coin is fake and always lands on the same side, but the agent initially does not know which. After the agent has executed the action once, she will know. She will observe either  $h$  becoming false or  $h$  becoming true, and can hence discard either  $e_1$  or  $e_2$  from her hypothesis space. She has now *learned* the correct deterministic action model for tossing the fake coin. Note the nice symmetry to learning of facts: here, learning of facts means eliminating worlds in epistemic models, learning of actions means eliminating events in action models.

*In the rest of this section, all action models are silently assumed to be: fully observable, propositional, and universally applicable. Furthermore, we can assume them to be normal and have basic preconditions, due to Proposition 1.*

#### 3.1 Learning precondition-free atomic actions

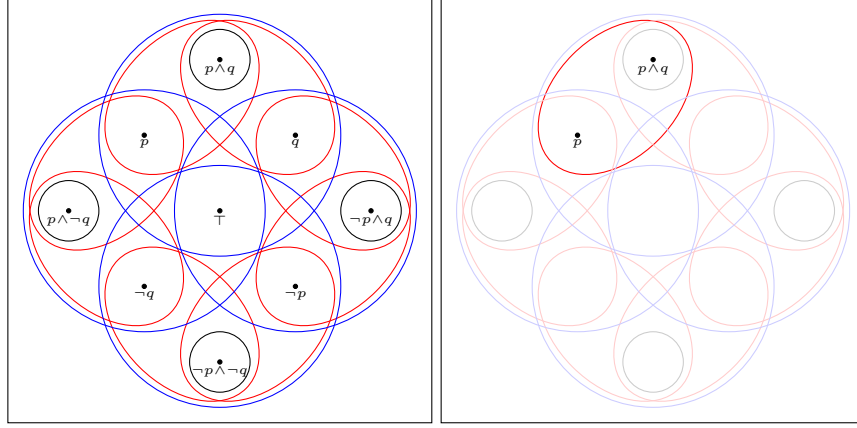
We will first propose and study an update learning method especially geared towards learning the simplest possible type of ontic actions: precondition-free atomic actions.

**Definition 11.** *For any deterministic action model  $\mathcal{A}$  and any pair of propositional states  $(s, s')$ , the update of  $\mathcal{A}$  with  $(s, s')$  is defined by  $\mathcal{A} \mid (s, s') := \mathcal{A} \upharpoonright \{e \in E \mid \text{if } \text{pre}(e) \models s \text{ then } s \otimes e = s'\}$ . For a set  $S$  of pairs of propositional states, we define:  $\mathcal{A} \mid S := \mathcal{A} \upharpoonright \{e \in E \mid \text{for all } (s, s') \in S, \text{ if } \text{pre}(e) \models s \text{ then } s \otimes e = s'\}$ .*

The update  $\mathcal{A} \mid (s, s')$  restricts the action model  $\mathcal{A}$  to the events that are consistent with observing  $s'$  as the result of executing the action in question in the state  $s$ .

**Definition 12.** The update learning function for precondition-free atomic actions over  $P$  is the learning function  $L_1$  defined by  $L_1(\mathcal{E}[n]) = \mathcal{A}_{init}^1 \upharpoonright \text{set}(\mathcal{E}[n])$  where  $\mathcal{A}_{init}^1 = (E, Q, pre, post)$  with  $E = \{\psi \mid \psi \text{ is a consistent conjunction of literals over } P\}$ ;  $Q$  is the identity relation on  $E \times E$ ;  $pre(e) = \top$  for all  $e \in E$ ;  $post(\psi) = \psi$ .

In Figure 1 we show a generic example of such update learning for  $P = \{p, q\}$ .



**Fig. 1.** On the left hand side  $\mathcal{A}_{init}^1$  for  $P = \{p, q\}$ , together with sets corresponding to possible observations. We have labelled each event  $e$  by  $post(e)$ . On the right hand side the state of learning after observing  $\mathcal{E}_0 = (\{q\}, \{p, q\})$ .

**Theorem 4.** The class of precondition-free atomic actions is finitely identifiable by the update learning function  $L_1^{update}$ , defined in the following way:

$$L_1^{update}(\mathcal{E}[n]) = \begin{cases} L_1(\mathcal{E}[n]) & \text{if } \text{card}(\text{dom}(L_1(\mathcal{E}[n]))) = 1 \\ & \text{and for all } k < n, L_1^{update}(\mathcal{E}[k]) = \uparrow; \\ \uparrow & \text{otherwise.} \end{cases}$$

*Proof.* Let  $\mathcal{A}$  denote a precondition-free atomic action over  $P$ , and let  $\mathcal{E}$  be a stream for  $\mathcal{A}$ . We show that  $L_1^{update}$  finitely identifies  $\mathcal{A}$  on  $\mathcal{E}$ . First note that  $L_1^{update}$  is obviously (at most) once-defined. Further, we need to show that for some  $n \in \mathbb{N}$ ,  $L_1^{update}(\mathcal{E}[n]) \equiv \mathcal{A}$ . By definition of  $L_1^{update}$ , it is the case only if  $\text{card}(\text{dom}(\mathcal{A}_{init}^1 \upharpoonright \text{set}(\mathcal{E}[n]))) = 1$  and  $(\mathcal{A}_{init}^1 \upharpoonright \text{set}(\mathcal{E}[n])) \equiv \mathcal{A}$ .

Since  $\mathcal{A}$  is atomic and precondition-free, it must consist of a single event of the form  $\langle \top, \psi \rangle$ . By definition of  $\mathcal{A}_{init}^1$  (Definition 12), this implies that there is an event  $e \in \text{dom}(\mathcal{A}_{init}^1)$  such that  $\mathcal{A} = \mathcal{A}_{init}^1 \upharpoonright \{e\}$ . Since  $\mathcal{E}$  is a stream for  $\mathcal{A}$ ,  $e$  is in  $\mathcal{A}_{init}^1 \upharpoonright \text{set}(\mathcal{E}[n])$ .

By Theorem 1 we know that  $\mathcal{A}$  is finitely identifiable, so, by Lemma 1, there is a DFTT  $D_{\mathcal{A}}$  for  $\mathcal{A}$ . Since  $\mathcal{E}$  is for  $\mathcal{A}$  and  $D_{\mathcal{A}}$  is by definition finite and sound for  $\mathcal{A}$ , there is  $n \in \mathbb{N}$  such that  $D_{\mathcal{A}} \subseteq \text{set}(\mathcal{E}[n])$ . By definition of DFTT and of  $\mathcal{A}_{init}^1$  we get that for all  $e' \in \mathcal{A}_{init}^1$  such that  $e' \neq e$  there is  $(s, s')$  in  $\mathcal{E}[n]$  such that  $s \otimes e' \neq s'$ , and hence  $e'$  is not in  $\mathcal{A}_{init}^1 \mid \text{set}(\mathcal{E}[n])$ . To see why that is assume the contrary, i.e., that there is an  $e' \in \mathcal{A}_{init}^1$ ,  $e' \neq e$  such that for all  $(s, s') \in \mathcal{E}[n]$  we have  $s \otimes e' = s'$ . Then  $D_{\mathcal{A}}$  is also sound for the singleton action model containing only  $e'$ . But this contradicts that  $D_{\mathcal{A}}$  is a DFTT for  $\mathcal{A}$  (since all pairs of distinct events in  $\mathcal{A}_{init}^1$  are inequivalent).

Combining the above we get that  $\mathcal{A}_{init}^1 \mid \text{set}(\mathcal{E}[n])$  contains exactly one event,  $e$ , and hence  $\mathcal{A}_{init}^1 \mid \text{set}(\mathcal{E}[n]) = \mathcal{A}_{init}^1 \upharpoonright \{e\} = \mathcal{A}$ , showing the required.

### 3.2 Learning deterministic actions with preconditions

We now turn to learning of action models with preconditions. First we only treat the case of maximal preconditions, then afterwards we generalise to arbitrary (not necessarily maximal) preconditions.

**Definition 13.** *The update learning function for deterministic action models with maximal preconditions over  $P$  is the learning function  $L_2$  defined by  $L_2(\mathcal{E}[n]) = \mathcal{A}_{init}^2 \mid \text{set}(\mathcal{E}[n])$  where  $\mathcal{A}_{init}^2 = (E, Q, \text{pre}, \text{post})$  with  $E = \{(\phi, \psi) \mid \phi \text{ is a maximally consistent conjunction of literals over } P \text{ and } \psi \text{ is a conjunction of literals over } P \text{ not containing any of the conjuncts of } \phi\}$ ;  $Q$  is the identity on  $E \times E$ ;  $\text{pre}((\phi, \psi)) = \phi$ ;  $\text{post}((\phi, \psi)) = \psi$ .*

**Theorem 5.** *The class of deterministic action models with maximal preconditions is finitely identifiable by the following update learning function  $L_2^{\text{update}}$ .*

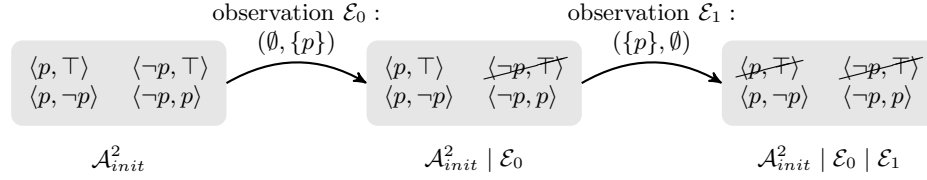
$$L_2^{\text{update}}(\mathcal{E}[n]) = \begin{cases} L_2(\mathcal{E}[n]) & \text{if for all } e, e' \in \text{dom}(L_2(\mathcal{E}[n])) \\ & \text{if } e \neq e', \text{ then } \text{pre}(e) \neq \text{pre}(e') \\ & \text{and for all } k < n, L_2^{\text{update}}(\mathcal{E}[k]) = \uparrow; \\ \uparrow & \text{otherwise.} \end{cases}$$

*Proof.* Consider any event  $e = \langle \phi, \psi \rangle$  in  $\mathcal{A}$ . Its precondition  $\phi$  is a maximally consistent conjunction of literals over  $P$ . Due to normality, its postcondition  $\psi$  can not contain any of the conjunctions of  $\phi$ . Hence  $e$  must be identical to one of the events of  $\mathcal{A}_{init}^2$ . In other words,  $\mathcal{A}$  must be isomorphic to a restriction/submodel of  $\mathcal{A}_{init}^2$ .

Let  $\mathcal{E}$  be a stream for  $\mathcal{A}$ . We show that  $L_2^{\text{update}}$  finitely identifies  $\mathcal{A}$  on  $\mathcal{E}$ .  $L_2^{\text{update}}$  is obviously (at most) once-defined. Further, we need that for some  $n \in \mathbb{N}$ ,  $L_2^{\text{update}}(\mathcal{E}[n]) \equiv \mathcal{A}$ . By Theorem 1 we know that  $\mathcal{A}$  is finitely identifiable, so, by Lemma 1, there is a DFTT  $D_{\mathcal{A}}$  for  $\mathcal{A}$ , and hence there is  $n \in \mathbb{N}$  such that  $D_{\mathcal{A}} \subseteq \text{set}(\mathcal{E}[n])$ . Firstly, note that all distinct  $e, e' \in \text{dom}(L_2(\mathcal{E}[n]))$  have distinct preconditions  $\text{pre}(e) \neq \text{pre}(e')$ . Here a similar argument applies as in the proof of Theorem 4. This gives us that  $L_2^{\text{update}}$  is guaranteed to give an answer. Secondly,

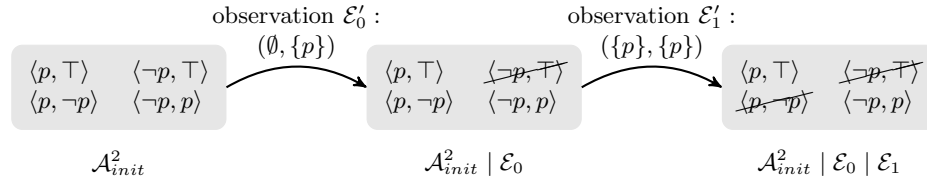
we have to show that for any  $k < n$ , if for all distinct  $e, e' \in \text{dom}(L_2(\mathcal{E}[k]))$  it is the case that  $\text{pre}(e) \neq \text{pre}(e')$ , then  $L_2(\mathcal{E}[k]) \equiv \mathcal{A}$ . Take any  $k < n$ , there are two cases. If  $\mathcal{E}[k]$  does not include a DFTT for  $\mathcal{A}$ , then there are two distinct  $e, e' \in L_2(\mathcal{E}[k])$  such that  $\text{pre}(e) = \text{pre}(e')$ . If  $\mathcal{E}[k]$  includes a DFTT for  $\mathcal{A}$ , then actions  $\mathcal{A}' \neq \mathcal{A}$  have been eliminated from  $\mathcal{A}_{init}^2$  by step  $k$ , and hence  $L_2(\mathcal{E}[k]) \equiv \mathcal{A}$ .

*Example 5.* Consider a simple scenario with a pushbutton and a light bulb. Assume there is only one proposition  $p$ : ‘the light is on’, and only one action: pushing the button. We assume an agent wants to learn the functioning of the pushbutton. There are 4 distinct possibilities: 1) the button does not affect the light (i.e., the truth value of  $p$ ); 2) it is an *on button*: it turns on the light unconditionally (makes  $p$  true); 3) it is an *off button*: it turns off the light unconditionally (makes  $p$  false); 4) it is an *on/off button* (flips the truth value of  $p$ ). If the agent is learning by update, it starts with the action model  $\mathcal{A}_{init}^2$  containing the following events:  $\langle p, \top \rangle$ ,  $\langle \neg p, \top \rangle$ ,  $\langle p, \neg p \rangle$ , and  $\langle \neg p, p \rangle$ . Note that by definition  $\mathcal{A}_{init}^2$  does not contain the events  $\langle p, p \rangle$  and  $\langle \neg p, \neg p \rangle$ , since they both have a postcondition conjunct which is also a precondition conjunct. Assume the first two observations the learner receives (the first elements of a stream  $\mathcal{E}$ ) are  $(\emptyset, \{p\})$  and  $(\{p\}, \emptyset)$ . Since the agent uses learning by update, she revises her model as follows (cf. Definition 13):



Now the agent has reached a deterministic action model  $\mathcal{A}_{init}^2 \mid \text{set}(\mathcal{E}[2])$ , and can report this to be the correct model of the action, cf. Theorem 5. Note that the two observations correspond to first pushing the button when the light is off ( $\mathcal{E}_0$ ), and afterwards pushing the button again after the light has come on ( $\mathcal{E}_1$ ). These two observations are sufficient to learn that the pushbutton is of the on/off type (it has one event that makes  $p$  true if  $p$  is currently false, and another event making  $p$  true if currently false).

Consider now another stream  $\mathcal{E}'$  where the first two elements are  $(\emptyset, \{p\})$  and  $(\{p\}, \{p\})$ . Update learning will now work as follows:



This time the learner identifies the button to be an on button, again after only two observations. It is not hard to show that in a setting with only one proposi-

tional symbol  $p$ , any deterministic action will be identified after having received the first two distinct observations.

*Example 6.* Consider learning the functioning of an  $n$ -bit binary counter, where the action to be learned is the increment operation. For  $i = 1, \dots, n$ , we use the proposition  $c_i$  to denote that the  $i$ th least significant bit is 1. Consider first the case  $n = 2$ . A possible stream for the increment operation is the following:

$$(\emptyset, \{c_1\}), \quad (\{c_1\}, \{c_2\}), \quad (\{c_2\}, \{c_2, c_1\}), \quad (\{c_2, c_1\}, \{\emptyset\}), \quad \dots$$

$$\begin{array}{ccccccc} \boxed{0} \boxed{0} \rightarrow \boxed{0} \boxed{1} & \boxed{0} \boxed{1} \rightarrow \boxed{1} \boxed{0} & \boxed{1} \boxed{0} \rightarrow \boxed{1} \boxed{1} & \boxed{1} \boxed{1} \rightarrow \boxed{0} \boxed{0} & \dots \\ c_2 \ c_1 & c_2 \ c_1 & c_2 \ c_1 & c_2 \ c_1 & \end{array}$$

Using the update learning method on this stream, it is easy to show that the learner will after the first 4 observations be able to report the correct action model containing the following events:  $\langle \neg c_2 \wedge \neg c_1, c_1 \rangle, \langle \neg c_2 \wedge c_1, c_2 \wedge \neg c_1 \rangle, \langle c_2 \wedge \neg c_1, c_1 \rangle, \langle c_2 \wedge c_1, \neg c_2 \wedge \neg c_1 \rangle$ . Note that since  $\mathcal{A}_{init}^2$  has maximal preconditions, the action model learned for an  $n$ -bit counter will necessarily contain  $2^n$  events: one for each possible configuration of the  $n$  bits. If we did not insist on maximal preconditions, we would only need  $n + 1$  events to describe the  $n$ -bit counter:  $\langle \neg c_i \wedge c_{i-1} \wedge c_{i-2} \wedge \dots \wedge c_1, c_i \wedge \neg c_{i-1} \wedge \neg c_{i-2} \wedge \dots \wedge \neg c_1 \rangle$  for all  $i = 2, \dots, n$ ,  $\langle \neg c_1, c_1 \rangle$  and  $\langle c_n \wedge \dots \wedge c_1, \neg c_n \wedge \dots \wedge \neg c_1 \rangle$ . This means that there is room for improvement in our learning method.

To allow learning of deterministic action models where preconditions are not required to be maximal we need a different learning condition. Consider learning an action on  $P = \{p\}$  that sets  $p$  true unconditionally. With non-maximal preconditions, all of the following events would be consistent with any stream for the action:  $\langle \top, p \rangle, \langle \neg p, p \rangle, \langle p, \top \rangle$ . To get to a deterministic action model, the learning function would have to delete either the first or the two latter events. We can make it work as described in the following.

For any action model  $\mathcal{A}$  we define

$$\min(\mathcal{A}) = \mathcal{A} \upharpoonright \{e \mid \text{there is no event } e' \neq e \text{ with } pre(e) \models pre(e')\}.$$

Furthermore, we define  $L_3$  to be exactly like  $L_2$  of Definition 13 except in the definition of  $E$ ,  $\phi$  can be any conjunction of literals, not only maximally consistent ones.

**Theorem 6.** *The class of deterministic action models is finitely identifiable by the following update learning function  $L_3^{update}$ .*

$$L_3^{update}(\mathcal{E}[n]) = \begin{cases} \min(L_3(\mathcal{E}[n])) & \text{if for all } s \in \mathcal{P}(P) \text{ there exists an } s' \text{ s.t.} \\ & (s, s') \in \text{set}(\mathcal{E}[n]) \text{ and for all } k < n, \\ & L_3^{update}(\mathcal{E}[k]) = \uparrow; \\ \uparrow & \text{otherwise.} \end{cases}$$

The proof of this theorem is left out. The theorem can be seen as a generalisation of Theorem 5 in that it allows the learner to learn more compact action models in which maximal consistency of preconditions is not enforced (on the contrary, by the way the min operator is defined above, the learner will learn an action model with *minimal* preconditions). For instance, in the case of the  $n$ -bit counter considered in Example 5, it can be shown that the learner will learn the action model with  $n + 1$  events instead of the one with  $2^n$  events.

## 4 Action library learning

In this section we introduce action library learning. This is the type of learning most relevant in planning. A finite set of action names is available to the agent. In order to plan a sequence of actions towards a goal it is essential to know what the corresponding actions do. As most of the results in this section are straightforward generalizations of our previous results, for the sake of space we will omit all proofs.

An action library corresponds to what is called a *planning domain* (and sometimes also an *action library*) in classical planning: a specification of the available actions and their action schemas. *Action library learning* is the learning problem where the agent is initially only given a set  $A$  of action names and has to learn the action library  $l : A \rightarrow \mathbf{Actions}(P)$ . That is, the agent initially only knows the names of the available actions, and it then learns the action models that correspond to those names.

**Definition 14 (Action library).** *Let  $P$  denote a set of atomic propositions, and let  $A$  denote a finite set, the set of action names. An action library over  $P, A$  is a mapping  $l : A \rightarrow \mathbf{Actions}(P)$  (a mapping from action names into action models). If all actions in the codomain of  $l$  enjoy property  $X$ , then  $l$  is called an  $X$  action library (e.g., a deterministic action library  $l$  is one where all action models in the codomain of  $l$  are deterministic).*

Streams and learning functions for action libraries are defined similarly to the case of individual actions. Let  $P$  be a set of atomic propositions, and  $A$  a set of action names. A *stream* over  $P, A$  is an infinite sequence of triples  $(s, a, s')$  where  $s, s'$  are propositional states over  $P$  and  $a \in A$ . Notations  $\mathcal{E}_n, \mathcal{E}[n], \text{set}(\mathcal{E})$  and  $\text{set}(\mathcal{E}[n])$  are defined similarly to Definition 5. Given  $a \in A$ , the  $a$ -*substream* of  $\mathcal{E}$  is given by  $\mathcal{E}^a = \{(s, s') \mid (s, a, s') \in \text{set}(\mathcal{E})\}$ . Let  $l$  be an action library over  $P, A$ . A *stream* for  $l$  is an infinite sequence of triples  $(s, a, s')$  where  $s, s'$  are propositional states over  $P$ ,  $a \in A$  and  $s' \in s \otimes l(a)$ . A *library learning function* over  $P, A$  is a mapping  $L : (\mathcal{P}(P) \times A \times \mathcal{P}(P))^* \rightarrow ((A \rightarrow \mathbf{Actions}(P)) \cup \{\uparrow\})$ . Given a learning function  $L$  for individual actions over  $P$  (Definition 7), we define the *induced library learning function*  $\mathbf{L}$  over  $P, A$  by

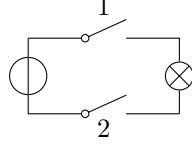
1.  $\mathbf{L}(\mathcal{E}) := \uparrow$  if  $L(\mathcal{E}^a) = \uparrow$  for some  $a \in A$ .
2. Otherwise, for all  $a \in A$ ,  $\mathbf{L}(\mathcal{E})(a) := L(\mathcal{E}^a)$ .

From Theorem 6 we then immediately get the following.



**Theorem 7.** *The class of deterministic action libraries is finitely identifiable by the library learning function  $\mathcal{L}_3^{update}$  (induced from  $L_3^{update}$  of Theorem 6).*

*Example 7.* Consider the electrical circuit below consisting of two switches (1 and 2), a voltage source (left) and a light bulb (right).



When both switches are closed, the light will be on, otherwise it will be off. Let proposition  $s_i$  denote that switch  $i$  is closed and let  $l$  denote that the light is on. Assume the available actions are  $flip_1$  and  $flip_2$  that flip switch 1 and 2, respectively. Consider an agent trying to learn how the switches and the circuit work. This agent then tries to learn an action library over  $\{s_1, s_2, l\}$ ,  $\{flip_1, flip_2\}$ . Given a stream  $\mathcal{E}$  for the action library, it can be shown that the learning function  $\mathcal{L}_3^{update}$  will eventually return the following action library that describes it (note that it can be described in many equivalent ways).

$$\begin{aligned} l(flip_1) &= \langle \neg s_1 \wedge \neg s_2, s_1 \wedge \neg l \rangle, \langle \neg s_1 \wedge s_2, s_1 \wedge l \rangle, \langle s_1 \wedge \neg s_2, \neg s_1 \wedge \neg l \rangle, \langle s_1 \wedge s_2, \neg s_1 \wedge \neg l \rangle \\ l(flip_2) &= \langle \neg s_1 \wedge \neg s_2, s_2 \wedge \neg l \rangle, \langle \neg s_1 \wedge s_2, \neg s_2 \wedge \neg l \rangle, \langle s_1 \wedge \neg s_2, s_2 \wedge l \rangle, \langle s_1 \wedge s_2, \neg s_2 \wedge \neg l \rangle \end{aligned}$$

## 5 Conclusions and related work

This paper is the first to study the problem of learnability of action models in dynamic epistemic logic (DEL). We provided an original learnability framework and several early results concerning fully observable propositional action models with respect to conclusive (finite identifiability) and inconclusive (identifiability in the limit) learnability. Apart from those general results, we proposed various learning functions which code particular learning algorithms. Here, by implementing the update method (commonly used in DEL), we demonstrated how the learning of action models can be seen as transitioning from nondeterministic to deterministic actions.

**Related work** A similar qualitative approach to learning actions has been addressed by [18] within the STRIPS planning formalism. The STRIPS setting is more general than ours in that it uses atoms of first-order predicate logic for pre- and postconditions. It is however less general in neglecting various aspects of actions which we have successfully treated in this paper: negative preconditions (i.e., negative literals as precondition conjuncts), negative postconditions, conditional effects (which we achieve through non-atomic action models). We believe that the ideas introduced here can be applied to generalize the results of [18] to richer planning frameworks allowing such action types. It is also worth mentioning here that there has been quite substantial amount of work in relating

DEL and learning theory (see [11,12] for overviews), which concerns a different setting: treating update and upgrade revision policies as long term learning methods, where learning can be seen as convergence to certain types of knowledge (see [3,4,5]). A study of abstract properties of finite identifiability in a setting similar to ours, including various efficiency considerations, can be found in [13].

**Further directions** In this short paper we only considered fully observable actions applied in fully observable states, and hence did not use the full expressive power of the DEL formalism. The latter still remains adequate, since action models provide a very well-structured and principled way of describing actions in a logical setting, and since its use opens ways to various extensions. The next steps are to cover more DEL action models: those with arbitrary pre- and postconditions, and those with partial observability and multiple agents. As described earlier, partially observable actions are not learnable in the strict sense considered above, but we can still investigate agents learning “as much as possible” given their limitations in observability. The multi-agent case is particularly interesting due to the possibility of agents with varied limitations on observability, and the possibility of communication within the learning process.

We plan to study the computational complexity of learning proposed in this paper, but also to investigate other more space-efficient learning algorithms. We are also interested algorithms that produce minimal action models. For instance, if we allow action models that have event postconditions specified as mappings from propositions to formulas (as is standard in DEL), then the action library for the circuit of Example 7 can be described using only 2 events. However, learning such minimal action descriptions might turn out to be computationally much harder. Furthermore, we here considered only what we call *reactive learning*: the learner has no influence over observations. We would also like to study the case of *proactive learning*, where the learner gets to choose which actions to execute, and hence observe their effects. This is probably the most relevant type of learning for a general learning-and-planning agent. In this context, we also plan to focus on *consecutive streams*: streams corresponding to executing sequences of actions rather than observing arbitrary state transitions. Our ultimate aim is to relate learning and planning within the framework of DEL. Those two cognitive capabilities are now investigated mostly in separation—our goal is to bridge them.

## 6 Acknowledgments

Nina Gierasimczuk is funded by an Innovational Research Incentives Scheme Veni grant 275-20-043, Netherlands Organisation for Scientific Research (NWO). We thank Martin Holm Jensen for his contributions through discussions we had in the early stages of this paper. We are also grateful to Mikko Berggren Ettienne who made us aware of [18]. We are also grateful to the anonymous reviewers for their valuable feedback.

## References

1. Andersen, M., Bolander, T., Jensen, M.: Conditional epistemic planning. In: del Cerro, L., Herzig, A., Mengin, J. (eds.) *Logics in Artificial Intelligence, Lecture Notes in Computer Science*, vol. 7519, pp. 94–106. Springer (2012)
2. Angluin, D.: Inductive inference of formal languages from positive data. *Information and Control* 45(2), 117–135 (1980)
3. Baltag, A., Gierasimczuk, N., Smets, S.: Belief revision as a truth-tracking process. In: Apt, K. (ed.) *TARK-11: Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 187–190. ACM (2011)
4. Baltag, A., Gierasimczuk, N., Smets, S.: Truth tracking by belief revision. *ILLC Prepublication Series PP-2014-20* (to appear in *Studia Logica* 2015) (2014)
5. Baltag, A., Gierasimczuk, N., Smets, S.: On the solvability of inductive problems: A study in epistemic topology. *ILLC Prepublication Series PP-2015-13* (to appear in *Proceedings of TARK-15*) (2015)
6. Baltag, A., Moss, L.S., Solecki, S.: The logic of public announcements and common knowledge and private suspicions. In: Gilboa, I. (ed.) *TARK-98: Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge*. pp. 43–56. Morgan Kaufmann (1998)
7. Bolander, T., Andersen, M.B.: Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics* 21, 9–34 (2011)
8. van Ditmarsch, H., Kooi, B.: Semantic results for ontic and epistemic change. In: Bonanno, G., van der Hoek, W., Wooldridge, M. (eds.) *LOFT 7: Logic and the Foundation of Game and Decision Theory*. pp. 87–117. Texts in Logic and Games 3, Amsterdam University Press (2008)
9. Ghallab, M., Nau, D.S., Traverso, P.: *Automated Planning: Theory and Practice*. Morgan Kaufmann (2004)
10. Gierasimczuk, N.: Learning by erasing in dynamic epistemic logic. In: Dediu, A.H., Ionescu, A.M., Martin-Vide, C. (eds.) *LATA-09: Proceedings of 3rd International Conference on Language and Automata Theory and Applications, Lecture Notes in Computer Science*, vol. 5457, pp. 362–373. Springer (2009)
11. Gierasimczuk, N.: *Knowing One’s Limits. Logical Analysis of Inductive Inference*. Ph.D. thesis, Universiteit van Amsterdam, The Netherlands (2010)
12. Gierasimczuk, N., de Jongh, D., Hendricks, V.F.: Logic and learning. In: Baltag, A., Smets, S. (eds.) *Johan van Benthem on Logical and Informational Dynamics*. Springer (2014)
13. Gierasimczuk, N., de Jongh, D.: On the complexity of conclusive update. *The Computer Journal* 56(3), 365–377 (2013)
14. Lange, S., Zeugmann, T.: Types of monotonic language learning and their characterization. In: *COLT’92: Proceedings of the 5th Annual ACM Conference on Computational Learning Theory, 1992*, pp. 377–390. ACM (1992)
15. Mukouchi, Y.: Characterization of finite identification. In: Jantke, K. (ed.) *AII’92: Proceedings of the International Workshop on Analogical and Inductive Inference, 1992, Lecture Notes in Computer Science*, vol. 642, pp. 260–267. Springer (1992)
16. Sadzik, T.: Exploring the Iterated Update Universe, *ILLC Prepublications PP-2006-26* (2006)
17. Sietsma, F., van Eijck, J.: Action emulation between canonical models. *Journal of Philosophical Logic* 42(6), 905–925 (2013)
18. Walsh, T.J., Littman, M.L.: Efficient learning of action schemas and web-service descriptions. In: *AAAI’08: Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 2*. pp. 714–719. AAAI Press (2008)