

A Logic for Specifying Partially Observable Stochastic Domains

Gavin Rens^{1,2} and Thomas Meyer^{1,2} and Alexander Ferrein³ and Gerhard Lakemeyer³
{grens, tmeyer}@meraka.org.za {ferrein, gerhard}@cs.rwth-aachen.de

¹CSIR Meraka Institute, Pretoria, South Africa

²University of KwaZulu-Natal, School of Computer Science, South Africa

³RWTH Aachen University, Informatik, Germany

Abstract

We propose a novel modal logic for specifying agent domains where the agent’s actuators and sensors are noisy, causing uncertainty in action and perception. The logic draws both on POMDP theory and logics of action and change. The development of the logic builds on previous work in which a simple multi-modal logic was augmented with first-class *observation* objects. These *observations* can then be used to represent the set of observations in a POMDP model in a natural way. In this paper, a subset of the simple modal logic is taken for the new logic, in which modal operators may not be nested. The modal operators are then extended with notions of probability. It will be shown how stochastic domains can be specified, including new kinds of axioms dealing with perception and a *frame solution* for the proposed logic.

1 Introduction and Motivation

In the physical real world, or in extremely complex engineered systems, things are not black-and-white. We live in a world where there can be shades of truth and degrees of belief. Part of the problem is that agents’ actuators and sensors are noisy, causing uncertainty in their action and perception. In this paper, we propose a novel logic that draws on partially observable Markov decision process (POMDP) theory and on logics for reasoning about action and change, combining both in a coherent language to model change and uncertainty.

Imagine a robot that is in need of an oil refill. There is an open can of oil on the floor within reach of its gripper. If there is nothing else in the robot’s gripper, it can grab the can (or miss it, or knock it over) and it can drink the oil by lifting the can to its ‘mouth’ and pouring the contents in (or miss its mouth and spill). The robot may also want to confirm whether there is anything left in the oil-can by weighing its contents. And once holding the can, the robot may wish to place it back on the floor. In situations where the oil-can is full, the robot gets 5 units of reward for grabbing the can, and it gets 10 units for a *drink* action. Otherwise, the robot gets no rewards. Rewards motivate an agent to behave as desired.

The domain is (partially) formalized as follows. The robot has the set of actions $\mathcal{A} = \{grab, drink, weigh, replace\}$

with expected meanings. The robot can perceive observations only from the set $\Omega = \{obsNil, obsLight, obsMedium, obsHeavy\}$. When the robot performs a *weigh* action (i.e., it activates its ‘weight’ sensor) it will perceive either *obsLight*, *obsMedium* or *obsHeavy*; for other actions, it will perceive *obsNil*. The robot experiences its environs through three Boolean features: $\mathfrak{P} = \{full, drank, holding\}$ meaning respectively that the oil-can is full, that the robot has drunk the oil and that it is currently holding something in its gripper.

Given a formalization \mathcal{K} of our scenario, the robot may have the following queries:

- Is the probability of perceiving that the oil-can is light 0.7 when the can is not full, and have I drunk the oil, and am I holding the can? Does $(obsLight \mid weigh)_{0.7}(\neg full \wedge drank \wedge holding)$ follow from \mathcal{K} ?
- If the oil-can is empty and I’m not holding it, is there a 0.9 probability that I’ll be holding it after grabbing it, and a 0.1 probability that I’ll have missed it? Does $(\neg full \wedge \neg holding) \rightarrow ([grab]_{0.9}(\neg full \wedge holding) \wedge [grab]_{0.1}(\neg full \wedge \neg holding))$ follow from \mathcal{K} ?

In order for robots and intelligent agents in stochastic domains to reason about actions and observations, they must first have a model of the domain over which to reason. For example, a robot may need to represent available knowledge about its *grab* action in its current situation. It may need to represent that when ‘grabbing’ the oil-can, there is a 5% chance that it will knock over the oil-can. As another example, if the robot has access to information about the weight of an oil-can, it may want to represent the fact that the can weighs heavy with a 90% chance in ‘situation A’, but that it is heavy with a 98% chance in ‘situation B’.

Logic-based artificial intelligence for agent reasoning is well established. In particular, a domain expert choosing to represent domains with a *logic* can take advantage of the progress made in cognitive robotics [Levesque and Lakemeyer, 2008] to specify domains in a compact and transparent manner. Modal logic is considered to be well suited to reasoning about beliefs and changing situations.

POMDP theory has proven to be a good general framework for formalizing *dynamic, stochastic* systems. A drawback of traditional POMDP models is that they cannot include information about general facts and laws. Moreover, succinct axioms describing the dynamics of a domain cannot be writ-

ten in POMDP theory. In this work, we develop a logic that will further our goal of combining modal logic with POMDP theory. That is, here we design a modal logic that can represent POMDP problems specifically for reasoning tasks in cognitive robotics (with domain axioms). The logic for actual decision-making will be developed in later work. To facilitate the correspondence between POMDPs and an agent logic, we require *observation objects* in the logic to correspond to the POMDPs’ set of observations. Before the introduction of the Logic of Actions and Observations (LAO) [Rens *et al.*, 2010], no modal logic had explicit observations as first-class elements; sensing was only dealt with via special actions or by treating actions in such a way that they somehow get hold of observations. LAO is also able to accommodate models of nondeterminism in the actions and models of uncertainty in the observations. But in LAO, these notions are non-probabilistic.

In this paper we present the Specification Logic of Actions and Observations with Probability (SLAOP). SLAOP is derived from LAO and thus also considers observations as first-class objects, however, a probabilistic component is added to LAO for expressing uncertainty more finely. We have invented a new knowledge representation framework for our observation objects, based on the established approaches for specifying the behavior of actions.

We continue our motivation with a look at the related work, in Section 2. Section 3 presents the logic and Section 4 provides some of the properties that can be deduced. Section 5 illustrates domain specification with SLAOP, including a solution to the frame problem. Section 6 concludes the paper.

2 Related Work

Although SLAOP uses probability theory, it is not for reasoning *about* probability; it is for reasoning about (probabilistic) actions and observations. There have been many frameworks for reasoning *about* probability, but most of them are either not concerned with *dynamic environments* [Fagin and Halpern, 1994; Halpern, 2003; Shirazi and Amir, 2007] or they *are* concerned with change, but they are not actually *logics* [Boutilier *et al.*, 2000; Bonet and Geffner, 2001]. Some probabilistic logics for reasoning about action and change do exist [Bacchus *et al.*, 1999; Iocchi *et al.*, 2009], but they lack some desirable attributes, for example, a solution to the frame problem, nondeterministic actions, or catering for sensing. There are some logics that come closer to what we desire [Weerdt *et al.*, 1999; Van Diggelen, 2002; Gabaldon and Lakemeyer, 2007; Van Benthem *et al.*, 2009], that is, they are modal and they incorporate notions of probability, but they were not created with POMDPs in mind and they don’t take observations as first-class objects. One non-logical formalism for representing POMDPs [Boutilier and Poole, 1996] exploits structure in the problems for more compact representations. In (logic-based) cognitive robotics, such compact representation is the norm, for example, specifying only local effects of actions, and specifying a value related to a *set* of states in only one statement.

On the other hand, there are three formalisms for specifying POMDPs that employ *logic-based* representation [Wang

and Schmolze, 2005; Sanner and Kersting, 2010; Poole, 1998]. But for two of these, the frameworks are not *logics* *per se*. The first [Wang and Schmolze, 2005] is based on *Functional STRIPS*, “which is a simplified first-order language that involves constants, functions, and predicate symbols but does not involve variables and quantification”. Their representations of POMDPs are relatively succinct and they have the advantage of using first-order predicates. The STRIPS-like formalism is geared specifically towards planning, though, and their work does not mention reasoning about general facts. Moreover, in their approach, action-nondeterminism is modeled by associating sets of deterministic action-outcomes per nondeterministic action, whereas SLAOP will model nondeterminism via action *effects*—arguably, ours is a more natural and succinct method. Sanner and Kersting [2010] is similar to the first formalism, but instead of Functional STRIPS, they use the situation calculus to model POMDPs. Although reified situations make the meaning of formulae perspicuous, and reasoning with the situation calculus, in general, has been accepted by the community, when actions are nondeterministic, ‘action histories’ cause difficulties in our work: The set of possible alternative histories is unbounded and some histories may refer to the same state [Rens, 2010, Chap. 6]. When, in future work, SLAOP is extended to express belief states (i.e., sets of possible alternative states), dealing with duplicate states will be undesirable.

The Independent Choice Logic [Poole, 1998] is relatively different from SLAOP; it is an extension of Probabilistic Horn Abduction. Due to its difference, it is hard to compare to SLAOP, but it deserves mentioning because it shares its application area with SLAOP and both are inspired by decision theory. The future may tell which logic is better for certain representations and for reasoning over the representations.

Finally, SLAOP was not conceived as a new approach to represent POMDPs, but as the underlying specification language in a larger *meta-language* for reasoning robots that include notions of probabilistic uncertainty. The choice of POMDPs as a semantic framework is secondary.

3 Specification Logic of Actions and Observations with Probability

SLAOP is a non-standard modal logic for POMDP specification for robot or intelligent agent design. The specification of robot movement has a ‘single-step’ approach in SLAOP. As such, the syntax will disallow nesting of modal operators; sentences with *sequences* of actions, like $[grab][drink][replace]drank$ are not allowed. Sentences will involve at most unit actions, like $[grab]holding \vee [drink]drank$. Nevertheless, the ‘single-step’ approach is sufficient for *specifying* the probabilities of transitions due to action executions. The logic to be defined in a subsequent paper will allow an agent to query the probability of some propositional formula φ after an *arbitrary* sequence of actions and observations.

3.1 Syntax

The vocabulary of our language contains four sorts:

1. a finite set of *fluents* (alias *propositional atoms*) $\mathfrak{F} = \{p_1, \dots, p_n\}$,
2. a finite set of names of atomic *actions* $\mathfrak{A} = \{\alpha_1, \dots, \alpha_n\}$,
3. a finite set of names of atomic *observations* $\Omega = \{\varsigma_1, \dots, \varsigma_n\}$,
4. a countable set of names $\mathfrak{Q} = \{q_1, q_2, \dots\}$ of *rational numbers* in \mathbb{Q} .

From now on, denote $\mathbb{Q} \cap (0, 1]$ as \mathbb{Q}^\cap . We refer to elements of $\mathfrak{A} \cup \Omega \cup \mathfrak{Q}$ as *constants*. We are going to work in a multi-modal setting, in which we have modal operators $[\alpha]_q$, one for each $\alpha \in \mathfrak{A}$, and predicates $(\varsigma \mid \alpha)_q$ and $(\varsigma \mid \alpha)^\diamond$, for each pair in $\Omega \times \mathfrak{A}$.

Definition 3.1 Let $\alpha, \alpha' \in \mathfrak{A}$, $\varsigma, \varsigma' \in \Omega$, $q \in (\mathfrak{Q} \cap (0, 1])$, $r, c \in \mathfrak{Q}$ and $p \in \mathfrak{F}$. The language of SLAOP, denoted \mathcal{L}_{SLAOP} , is the least set of Φ defined by the grammars:

$$\begin{aligned} \varphi &::= p \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi. \\ \Phi &::= \varphi \mid [\alpha]_q\varphi \mid (\varsigma \mid \alpha)_q \mid (\varsigma \mid \alpha)^\diamond \mid \alpha = \alpha' \mid \\ &\quad \varsigma = \varsigma' \mid \text{Reward}(r) \mid \text{Cost}(\alpha, c) \mid \neg\Phi \mid \Phi \wedge \Phi. \end{aligned}$$

As usual, we treat \perp, \vee, \rightarrow and \leftrightarrow as abbreviations.

We shall refer to formulae $\varphi ::= p \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi$ as *static*. If a formula is static, it mentions no actions and no observations.

$[\alpha]_q\varphi$ is read ‘The probability of reaching a world in which φ holds after executing α , is equal to q ’. $[\alpha]$ abbreviates $[\alpha]_1$. $\langle\alpha\rangle\varphi$ abbreviates $\neg[\alpha]\neg\varphi$. $(\varsigma \mid \alpha)_q$ can be read ‘The probability of perceiving ς is equal to q , given α was performed’. $(\varsigma \mid \alpha)^\square$ abbreviates $(\varsigma \mid \alpha)_1$. $(\varsigma \mid \alpha)^\diamond$ is read ‘It is possible to perceive ς ’, given α was performed’.

The definition of a POMDP reward function $\mathcal{R}(a, s)$ may include not only the expected rewards for being in the states reachable from s via a , but it may deduct the cost of performing a in s . To specify rewards and execution costs in SLAOP, we require *Reward* and *Cost* as special predicates. *Reward*(r) can be read ‘The reward for being in the current situation is r units,’ and we read *Cost*(α, c) as ‘The cost for executing α is c units.’

Let $V_{\mathfrak{A}} = \{v_1^\alpha, v_2^\alpha, \dots\}$ be a countable set of *action variables* and $V_{\Omega} = \{v_1^\varsigma, v_2^\varsigma, \dots\}$ a countable set of *observation variables*. Let $\varphi|_{\alpha_1}^{\alpha} \wedge \dots \wedge \varphi|_{\alpha_n}^{\alpha}$ be abbreviated by $(\forall v^\alpha)\varphi$, where $\varphi|_c^v$ means φ with all variables $v \in (V_{\mathfrak{A}} \cup V_{\Omega})$ appearing in it replaced by constant c of the right sort (action or observation). Quantification over observations is similar to that for actions; the symbol \exists is also available for abbreviation, with the usual meaning.

3.2 Semantics

While presenting our semantics, we show how a POMDP, as defined below, can be represented by a SLAOP structure.

A POMDP [Kaelbling *et al.*, 1998] (for our purposes) is a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \Omega, \mathcal{O} \rangle$, where \mathcal{S} is a finite set of states that the agent can be in; \mathcal{A} is a finite set of agent actions; \mathcal{T} is the

state-transition function, representing, for each action, transition probabilities between states; \mathcal{R} is the *reward function*, giving the expected immediate reward gained by the agent, for any state and agent action; Ω is a finite set of observations the agent can experience of its environment; and \mathcal{O} is the *observation function*, giving, for each action and the resulting state, a probability distribution over observations, representing the agent’s ‘trust’ in its observations.

Our semantics follows that of multi-modal logic **K**. However, SLAOP structures are non-standard in that they are extensions of structures with the form $\langle W, R \rangle$, where W is a *finite* set of worlds such that each world assigns a truth value to each atomic proposition, and R is a binary relation on W .

Intuitively, when talking about some world w , we mean a set of features (*fluents*) that the agent understands and that describes a state of affairs in the world or that describes a possible, alternative world. Let $w : \mathfrak{F} \mapsto \{0, 1\}$ be a total function that assigns a truth value to each fluent. Let C be the set of all possible functions w . We call C the *conceivable worlds*.

Definition 3.2 A SLAOP structure is a tuple $\mathfrak{S} = \langle W, R, O, N, Q, U \rangle$ such that

1. $W \subseteq C$: the set of possible worlds (corresponding to \mathcal{S});
2. R : a mapping that provides an accessibility relation $R_\alpha : W \times W \times \mathbb{Q}^\cap$ for each $\alpha \in \mathfrak{A}$ (corresponding to \mathcal{T}); Given some $w^- \in W$, we require that $\sum_{(w^-, w^+, pr) \in R_\alpha} pr = 1$; If $(w^-, w^+, pr), (w^-, w^+, pr') \in R_\alpha$, then $pr = pr'$;
3. O : a nonempty finite set of observations (corresponding to Ω);
4. $N : \Omega \mapsto O$ is a bijection that associates to each name in Ω , a unique observation in O ;
5. Q : a mapping that provides a perceivability relation $Q_\alpha : O \times W \times \mathbb{Q}^\cap$ for each $\alpha \in \mathfrak{A}$ (corresponding to \mathcal{O}); Given some $w^+ \in W$, we require that $\sum_{(o, w^+, pr) \in Q_\alpha} pr = 1$; If $(\varsigma, w^+, pr), (\varsigma, w^+, pr') \in Q_\alpha$, then $pr = pr'$;
6. U : a pair $\langle Re, Co \rangle$ (corresponding to \mathcal{R}), where $Re : W \mapsto \mathbb{Q}$ is a reward function and Co is a mapping that provides a cost function $Co_\alpha : W \mapsto \mathbb{Q}$ for each $\alpha \in \mathfrak{A}$;
7. *Observation-per-action condition*: For all $\alpha \in \mathfrak{A}$, if $(w, w', pr^\alpha) \in R_\alpha$, then there is an $o \in O$ s.t. $(o, w', pr^\alpha) \in Q_\alpha$;
8. *Nothing-for-nothing condition*: For all w , if there exists no w' s.t. $(w, w', pr) \in R_\alpha$ for some pr , then $Co_\alpha(w) = 0$.

\mathfrak{A} corresponds to \mathfrak{A} and Ω to Ω . R_α defines which worlds w^+ are accessible via action α performed in world w^- and the transition probability $pr \in \mathbb{Q}^\cap$. Q_α defines which observations o are perceivable in worlds w^+ accessible via action α and the observation probability $pr \in \mathbb{Q}^\cap$. We prefer to exclude relation elements referring to transitions that cannot occur, hence why $pr \in \mathbb{Q}^\cap$ and not $pr \in \mathbb{Q} \cap [0, 1]$.

Because N is a bijection, it follows that $|O| = |\Omega|$ (we take $|X|$ to be the cardinality of set X). The value of the reward function $Re(w)$ is a rational number representing the reward an agent gets for being in or getting to the world w . It must be defined for each $w \in W$. The value of the cost function $Co(\alpha, w^-)$ is a rational number representing the cost of executing α in the world w^- . It must be defined for each action $\alpha \in \mathcal{A}$ and each $w^- \in W$. Item 7 of Definition 3.2 implies that actions and observations always appear in pairs, even if implicitly. And item 8 seems reasonable; it states that any action that is inexecutable in world w incurs no cost for it in the world w .

Definition 3.3 (Truth Conditions) *Let \mathfrak{S} be a SLAOP structure, with $\alpha, \alpha' \in \mathcal{A}$, $\varsigma, \varsigma' \in \Omega$, $q \in (\Omega \cap (0, 1])$ or \mathbb{Q}^{\cap} as applicable, and $r \in \Omega$ or \mathbb{Q} as applicable. Let $p \in \mathfrak{P}$ and let φ be any sentence in \mathcal{L}_{SLAOP} . We say φ is satisfied at world w in structure \mathfrak{S} (written $\mathfrak{S}, w \models \varphi$) if and only if the following holds:*

1. $\mathfrak{S}, w \models p$ iff $w(p) = 1$ for $w \in W$;
2. $\mathfrak{S}, w \models \top$ for all $w \in W$;
3. $\mathfrak{S}, w \models \neg\varphi$ iff $\mathfrak{S}, w \not\models \varphi$;
4. $\mathfrak{S}, w \models \varphi \wedge \varphi'$ iff $\mathfrak{S}, w \models \varphi$ and $\mathfrak{S}, w \models \varphi'$;
5. $\mathfrak{S}, w \models \alpha = \alpha'$ iff α and α' are identical;
6. $\mathfrak{S}, w \models \varsigma = \varsigma'$ iff ς and ς' are identical;
7. $\mathfrak{S}, w \models [\alpha]_q\varphi$ iff $(\sum_{(w, w', pr) \in R_{\alpha, \mathfrak{S}, w' \models \varphi}} pr) = q$;
8. $\mathfrak{S}, w \models (\varsigma \mid \alpha)_q$ iff $(N(\varsigma), w, q) \in Q_{\alpha}$;
9. $\mathfrak{S}, w \models (\varsigma \mid \alpha)_{\diamond}$ iff a q exists s.t. $(N(\varsigma), w, q) \in Q_{\alpha}$;
10. $\mathfrak{S}, w \models Reward(r)$ iff $Re(w) = r$;
11. $\mathfrak{S}, w \models Cost(\alpha, c)$ iff $Co_{\alpha}(w) = c$.

The definition of item 7 comes from probability theory, which says that the probability of an event (φ) is simply the sum of the probabilities of the atomic events (worlds) where the event (φ) holds.

A formula φ is valid in a SLAOP structure (denoted $\mathfrak{S} \models \varphi$) if $\mathfrak{S}, w \models \varphi$ for every $w \in W$. We define *global logical entailment* (denoted $\mathcal{K} \models_{GS} \varphi$) as follows: for all \mathfrak{S} , if $\mathfrak{S} \models \bigwedge_{\psi \in \mathcal{K}} \psi$ then $\mathfrak{S} \models \varphi$.

4 Some Properties

Remark 4.1 *Item 7 of Definition 3.2, the observation-per-action condition, implies that if $\mathfrak{S}, w \models \langle \alpha \rangle \varphi$ then $\mathfrak{S}, w' \models \varphi \rightarrow (\exists v^{\varsigma})(v^{\varsigma} \mid \alpha)_{\diamond}$, for some $w, w' \in W$.*

Remark 4.2 *Item 8 of Definition 3.2, the nothing-for-nothing condition, implies that $\models_{SLAOP} (\forall v^{\alpha}) \neg(v^{\alpha})\top \rightarrow Cost(v^{\alpha}, 0)$.*

In the terminology of probability theory, a single world would be called an *atomic event*. Probability theory says that the probability of an event e is simply the sum of the probabilities of the atomic events (worlds) where e holds. We are interested in noting the interactions of any two sentences of the form $[\alpha]_q\varphi$ being satisfied in the same world. Given the principle of the sum of atomic events, we get the following properties.

Proposition 4.1 *Assume an arbitrary structure \mathfrak{S} and some w in \mathfrak{S} . Assume $\mathfrak{S}, w \models [\alpha]_q\theta \wedge [\alpha]_{q'}\psi$. Then*

1. if $q = q'$ then no deduction can be made;
2. if $q \neq q'$ then $\mathfrak{S}, w \models \langle \alpha \rangle \neg(\theta \leftrightarrow \psi)$;
3. if $q > q'$ then $\mathfrak{S}, w \models \langle \alpha \rangle \neg(\theta \rightarrow \psi)$;
4. if $q + q' > 1$ then $\mathfrak{S}, w \models \langle \alpha \rangle (\theta \leftrightarrow \psi)$;
5. $\mathfrak{S}, w \models [\alpha] \neg(\theta \wedge \psi) \rightarrow [\alpha]_{q+q'}(\theta \vee \psi)$;
6. if $q = 1$ then $\mathfrak{S}, w \models [\alpha](\psi \rightarrow \theta)$ and $\mathfrak{S}, w \models [\alpha]_{q'}(\theta \wedge \psi)$;
7. $\mathfrak{S}, w \models [\alpha]_q\top$ is a contradiction if $q < 1$;
8. $\mathfrak{S}, w \models [\alpha]_{1-q}\neg\varphi$ iff $\mathfrak{S}, w \models [\alpha]_q\varphi$ and $q \neq 1$;
9. $\mathfrak{S}, w \models \neg[\alpha]_{1-q}\neg\varphi$ iff $\mathfrak{S}, w \models \neg[\alpha]_q\varphi$ and $q \neq 1$.

Proof:

Please refer to our draft report [Rens and Meyer, 2011]. Q.E.D.

It is worth noting that in the case when $q > q'$ (item 3), $\mathfrak{S}, w \models \langle \alpha \rangle \neg(\theta \wedge \psi)$ is also a consequence. But $\langle \alpha \rangle \neg(\theta \rightarrow \psi)$ logically implies $\langle \alpha \rangle \neg(\theta \wedge \psi)$.

Consider item 8 further: Suppose $[\alpha]_{q^*}\varphi$ where $q^* = 1$ (in some structure at some world). Then, in SLAOP, one could represent $\mathfrak{S}, w \models [\alpha]_{1-q^*}\neg\varphi$ as $\neg\langle \alpha \rangle \neg\varphi$. But this is just $[\alpha]\varphi$ ($\equiv [\alpha]_{q^*}\varphi$). The point is that there is no different way to represent $[\alpha]\varphi$ in SLAOP (other than syntactically). Hence, in item 8, we need not cater for the case when $q = 1$.

Proposition 4.2 $\models_{SLAOP} ([\alpha]_q\theta \wedge \neg[\alpha]_q\psi) \rightarrow \neg[\alpha](\theta \leftrightarrow \psi)$.

Proof:

Let \mathfrak{S} be any structure and w a world in \mathfrak{S} . Assume $\mathfrak{S}, w \models [\alpha]_q\theta \wedge \neg[\alpha]_q\psi$. Assume $\mathfrak{S}, w \models [\alpha](\theta \leftrightarrow \psi)$. Then because $\mathfrak{S}, w \models [\alpha]_q\theta$, one can deduce $\mathfrak{S}, w \models [\alpha]_q\psi$. This is a contradiction, therefore $\mathfrak{S}, w \not\models [\alpha](\theta \leftrightarrow \psi)$. Hence, $\mathfrak{S}, w \models ([\alpha]_q\theta \wedge \neg[\alpha]_q\psi) \rightarrow \neg[\alpha](\theta \leftrightarrow \psi)$. Q.E.D.

Proposition 4.3 *Assume an arbitrary structure \mathfrak{S} and an arbitrary world w in \mathfrak{S} . There exists some constant q such that $\mathfrak{S}, w \models [\alpha]_q\varphi$ if and only if $\mathfrak{S}, w \models \langle \alpha \rangle \varphi$.*

Proof:

Assume an arbitrary structure \mathfrak{S} and an arbitrary world w in it. Then

$$\begin{aligned} & \mathfrak{S}, w \models [\alpha]_q\varphi \text{ for some constant } q \\ \Leftrightarrow & \exists q \cdot (\sum_{(w, w', pr) \in R_{\alpha, \mathfrak{S}, w' \models \varphi}} pr) = q \\ \Leftrightarrow & \text{Not: } \exists q \cdot (\sum_{(w, w', pr) \in R_{\alpha, \mathfrak{S}, w' \models \varphi}} pr) = 0 \\ \Leftrightarrow & \text{Not: } \exists q \cdot (\sum_{(w, w', pr) \in R_{\alpha, \mathfrak{S}, w' \models \neg\varphi}} pr) = 1 \\ \Leftrightarrow & \text{Not: } \mathfrak{S}, w \models [\alpha]\neg\varphi \\ \Leftrightarrow & \mathfrak{S}, w \models \langle \alpha \rangle \varphi. \text{ Q.E.D.} \end{aligned}$$

We are also interested in noting the interactions of any two percept events—when sentences of the form $(\varsigma \mid \alpha)_q\varphi$ are satisfied in the same world. Only two consequences could be gleaned, given Definition 3.3, item 8:

Proposition 4.4 *Assume an arbitrary structure \mathfrak{S} and some w in \mathfrak{S} .*

1. If $\mathfrak{S}, w \models (\varsigma \mid \alpha)_q \wedge (\varsigma' \mid \alpha)_{q'}$ and ς is the same observation as ς' , then $q = q'$;

2. If $\mathfrak{S}, w \models (\varsigma \mid \alpha)_q \wedge (\varsigma' \mid \alpha)_{q'}$ and ς is not the same observation as ς' , then $q + q' \leq 1$.

Proof:

Directly from probability theory and algebra. Q.E.D.

Proposition 4.5 Assume an arbitrary structure \mathfrak{S} and an arbitrary world w in it. There exists some constant q such that $\mathfrak{S}, w \models (\varsigma \mid \alpha)_q$ if and only if $\mathfrak{S}, w \models (\varsigma \mid \alpha)^\diamond$.

Proof:

Let $N(\varsigma) = o$. Assume an arbitrary structure \mathfrak{S} and an arbitrary world w in \mathfrak{S} . Then

$\mathfrak{S}, w \models (\varsigma \mid \alpha)_q$ for some constant q

$\Leftrightarrow \exists q. (o, w, q) \in Q_\alpha$

$\Leftrightarrow \mathfrak{S}, w \models (\varsigma \mid \alpha)^\diamond$. Q.E.D.

The following is a direct consequence of Propositions 4.3 and 4.5.

Corollary 4.1 $\models_{SLAOP} [\alpha]_q \varphi \rightarrow \langle \alpha \rangle \varphi$ and $\models_{SLAOP} (\varsigma \mid \alpha)_q \rightarrow (\varsigma \mid \alpha)^\diamond$.

Further Properties of Interest

Recall that $R_\alpha^- = \{(w, w') \mid (w, w', pr) \in R_\alpha\}$. We now justify treating $[\alpha]_1$ as $[\alpha]$ of regular multi-modal logic.

Proposition 4.6 $[\alpha]_1$ is the regular $[\alpha]$. That is, $\mathfrak{S}, w \models [\alpha]_1 \varphi$ if and only if for all w' , if $w R_\alpha^- w'$, then $\mathfrak{S}, w' \models \varphi$, for any structure \mathfrak{S} and any world w in \mathfrak{S} .

Proof:

$\mathfrak{S}, w \models [\alpha]_1 \varphi$

$\Leftrightarrow (\sum_{(w, w', pr) \in R_\alpha, \mathfrak{S}, w' \models \varphi} pr) = 1$

$\Leftrightarrow \forall w'. \text{ if } \exists pr. (w, w', pr) \in R_\alpha \text{ then } \mathfrak{S}, w' \models \varphi$

$\Leftrightarrow \forall w'. \text{ if } w R_\alpha^- w' \text{ then } \mathfrak{S}, w' \models \varphi$. Q.E.D.

Proposition 4.7 $\langle \alpha \rangle$ has normal semantics. That is, $\mathfrak{S}, w \models \langle \alpha \rangle \varphi$ if and only if there exist w', pr such that $(w, w', pr) \in R_\alpha$ and $\mathfrak{S}, w' \models \varphi$.

Proof:

$\mathfrak{S}, w \models \langle \alpha \rangle \varphi$

$\Leftrightarrow \mathfrak{S}, w \models \neg[\alpha] \neg \varphi$

$\Leftrightarrow \mathfrak{S}, w \models \neg[\alpha]_1 \neg \varphi$

$\Leftrightarrow \mathfrak{S}, w \not\models [\alpha]_1 \neg \varphi$

$\Leftrightarrow (\sum_{(w, w', pr) \in R_\alpha, \mathfrak{S}, w' \not\models \neg \varphi} pr) \neq 1$

$\Leftrightarrow \exists w', pr. (w, w', pr) \in R_\alpha \text{ and } \mathfrak{S}, w' \not\models \neg \varphi$

$\Leftrightarrow \exists w', pr. (w, w', pr) \in R_\alpha \text{ and } \mathfrak{S}, w' \models \varphi$. Q.E.D.

5 Specifying Domains with SLAOP

We briefly describe and illustrate a framework to formally specify—in the language of SLAOP—the domain in which an agent or robot is expected to live. Let BK be an agent's *background knowledge* (including non-static formulae) and let IC be its *initial condition*, a static formula describing the world the agent finds itself in when it becomes active. In the context of SLAOP, we are interested in determining $BK \models_{GS} IC \rightarrow \varphi$, where φ is any sentence.

The agent's background knowledge may include *static law axioms* which are facts about the domain that do not change. They have no predictable form, but by definition, they are not dynamic and thus exclude mention of actions. $drank \rightarrow \neg full$ is one static law axiom for the oil-can scenario. The other kinds of axioms in BK are described below.

5.1 The Action Description

In the following discussion, W^φ is the set of worlds in which static formula φ holds (the ‘models’ of φ). A formal description for the construction of *conditional effect axioms* follows. For one action, there is a set of axioms that take the form

$$\begin{aligned} \phi_1 &\rightarrow ([\alpha]_{q_{11}} \varphi_{11} \wedge \dots \wedge [\alpha]_{q_{1n}} \varphi_{1n}); \\ \phi_2 &\rightarrow ([\alpha]_{q_{21}} \varphi_{21} \wedge \dots \wedge [\alpha]_{q_{2n}} \varphi_{2n}); \quad \dots; \\ \phi_j &\rightarrow ([\alpha]_{q_{j1}} \varphi_{j1} \wedge \dots \wedge [\alpha]_{q_{jn}} \varphi_{jn}), \end{aligned}$$

where the ϕ_i and φ_{ik} are static, and where the ϕ_i are conditions for the respective effects to be applicable, and in any one axiom, each φ_{ik} represents a set $W^{\varphi_{ik}}$ of worlds. The number q_{ik} is the probability that the agent will end up in a world in $W^{\varphi_{ik}}$, as the effect of performing α in the right condition ϕ_i . For axioms generated from the effect axioms (later in Sec. 5.1), we shall assume that φ_{ik} is a *minimal disjunctive normal form* characterization of $W^{\varphi_{ik}}$. The following constraints apply.

- There must be a set of effect axioms for each action $\alpha \in \mathfrak{A}$.
- The ϕ_i must be *mutually exclusive*, i.e., the conjunction of any pair of conditions causes a contradiction. However, it is not necessary that $W^{\varphi_{i1}} \cup \dots \cup W^{\varphi_{in}} = C$.
- A set of effects φ_{i1} to φ_{in} in any axiom i must be *mutually exclusive*.
- The *transition probabilities* q_{i1}, \dots, q_{in} of any axiom i must sum to 1.

The following sentence is an effect axiom for the *grab* action: $(full \wedge \neg holding) \rightarrow ([grab]_{0.7}(full \wedge holding) \wedge [grab]_{0.2}(\neg full \wedge \neg holding) \wedge [grab]_{0.1}(full \wedge \neg holding))$.

Executability axioms of the form $\phi_k \rightarrow \langle \alpha \rangle \top$ must be supplied, for each action, where ϕ_k is a precondition conveying physical restrictions in the environment with respect to α . The sentence $\neg holding \rightarrow \langle grab \rangle \top$ states that if the robot is not holding the oil-can, then it is possible to grab the can.

A set of axioms must be *generated* that essentially states that if the effect or executability axioms do not imply executability for some action, then that action is inexecutable. Hence, given α , assume the presence of an *executability closure axiom* of the following form: $\neg(\phi_1 \vee \dots \vee \phi_j \vee \phi_k) \rightarrow \neg \langle \alpha \rangle \top$. The sentence $holding \rightarrow \neg \langle grab \rangle \top$ states that if the robot is holding the oil-can, then it is not possible to grab it.

Now we show the form of sentences that specify what does *not* change under certain conditions—*conditional frame axioms*. Let $\phi_i \rightarrow ([\alpha]_{q_{i1}} \varphi_{i1} \wedge \dots \wedge [\alpha]_{q_{in}} \varphi_{in})$ be the i -th effect axiom for α . For each $\alpha \in \mathfrak{A}$, for each effect axiom i , do: For each fluent $p \in \mathfrak{F}$, if p is not mentioned in φ_{i1} to φ_{in} , then $(\phi_i \wedge p) \rightarrow [\alpha]p$ and $(\phi_i \wedge \neg p) \rightarrow [\alpha]\neg p$ are part of the domain specification.

For our scenario, the *conditional frame axioms* of *grab* are

$$\begin{aligned} (full \wedge \neg holding \wedge drank) &\rightarrow [grab]drank; \\ (full \wedge \neg holding \wedge \neg drank) &\rightarrow [grab]\neg drank; \\ (\neg full \wedge \neg holding \wedge drank) &\rightarrow [grab]drank; \\ (\neg full \wedge \neg holding \wedge \neg drank) &\rightarrow [grab]\neg drank. \end{aligned}$$

Given frame and effect axioms, it may still happen that the probability to some worlds cannot be logically deduced. Suppose (for the purpose of illustration only) that the sentence

$$\begin{aligned} & [grab]_{0.7}(full \wedge holding) \wedge \\ & [grab]_{0.3}(full \wedge \neg holding \wedge drank). \end{aligned} \quad (1)$$

can be logically deduced from the frame and effect axioms in *BK*. Now, according to (1) the following worlds are reachable: $(full \wedge holding \wedge drank)$, $(full \wedge holding \wedge \neg drank)$ and $(full \wedge \neg holding \wedge drank)$. The transition probability to $(full \wedge \neg holding \wedge drank)$ is 0.3, but what are the transition probabilities to $(full \wedge holding \wedge drank)$ and $(full \wedge holding \wedge \neg drank)$? We have devised a process to determine such hidden probabilities via *uniform axioms* [Rens and Meyer, 2011]. Uniform axioms describes how to distribute probabilities of effects uniformly in the case sufficient information is not available. It is very similar to what [Wang and Schmolze, 2005] do to achieve compact representation. A uniform axiom generated for (1) would be

$$\begin{aligned} & [grab]_{0.35}(full \wedge holding \wedge drank) \wedge \\ & [grab]_{0.35}(full \wedge holding \wedge \neg drank) \wedge \\ & [grab]_{0.3}(full \wedge \neg holding \wedge drank). \end{aligned}$$

The following axiom schema represents all the *effect condition closure axioms*. $(\neg(\phi_1 \vee \dots \vee \phi_j) \wedge P) \rightarrow [A]P$, where there is a different axiom for each substitution of $\alpha \in \mathfrak{A}$ for A and each literal for P . For example, $(holding \wedge P) \rightarrow [grab]P$, where P is any $p \in \mathfrak{P}$ or its negation.

5.2 The Perception Description

One can classify actions as either *ontic* (physical) or *sensory*. This classification also facilitates specification of perceivability. Ontic actions have intentional ontic effects, that is, effects on the environment that were the main intention of the agent. *grab*, *drink* and *replace* are ontic actions. Sensory actions—*weigh* in our scenario—result in perception, maybe with (unintended) side-effects.

Perceivability axioms specify what conditions must hold after the applicable action is performed, for the observation to be perceivable. Ontic actions each have perceivability axioms of the form $(obsNil \mid \alpha)^\square$. Sensory actions typically have multiple observations and associated conditions for perceiving them. The probabilities for perceiving the various observations associated with sensory actions must be specified. The following set of perceivability axiom schemata does this:

$$\begin{aligned} & \phi_{11} \rightarrow (\varsigma_1 \mid \alpha)_{q_{11}}; \quad \phi_{12} \rightarrow (\varsigma_1 \mid \alpha)_{q_{12}}; \quad \dots; \\ & \phi_{1j} \rightarrow (\varsigma_1 \mid \alpha)_{q_{1n}}; \quad \phi_{21} \rightarrow (\varsigma_2 \mid \alpha)_{q_{21}}; \quad \dots; \\ & \phi_{nk} \rightarrow (\varsigma_n \mid \alpha)_{q_{kn}}, \end{aligned}$$

where $\{\varsigma_1, \varsigma_2, \dots, \varsigma_n\}$ is the set of first components of all elements in Q_α and the ϕ_i are the conditions expressed as static formulae. The following constraints apply to these axioms.

- There must be a set of perceivability axioms for each action $\alpha \in \mathfrak{A}$.
- In the semantics section, item 7 of the definition of a SLAOP structure states that for every world reachable

via some action, there exists an observation associated with the action, perceivable in that world. The perceivability axioms must adhere to this remark.

- For every pair of perceivability axioms $\phi \rightarrow (\varsigma \mid \alpha)_q$ and $\phi' \rightarrow (\varsigma \mid \alpha)_{q'}$ for the same observation ς , W^ϕ must be disjoint from $W^{\phi'}$.
- For every particular condition ϕ , $\sum_{\phi \rightarrow (\varsigma \mid \alpha)_q} q = 1$. This is so that $\sum_{N(\varsigma): (N(\varsigma), w^+, pr) \in Q_\alpha} pr = 1$.

Some perceivability axioms for the oil-can scenario might be

$$\begin{aligned} & (obsNil \mid grab)^\square; \\ & (\neg full \wedge drank \wedge holding) \rightarrow (obsLight \mid weigh)_{0.7}; \\ & (\neg full \wedge drank \wedge holding) \rightarrow (obsHeavy \mid weigh)_{0.1}; \\ & (\neg full \wedge drank \wedge holding) \rightarrow (obsMedium \mid weigh)_{0.2}. \end{aligned}$$

Perceivability axioms for sensory actions also state *when* the associated observations are *possible*. The following set of axioms states when the associated observations are *impossible* for sensory action *weigh* of our scenario.

$$\begin{aligned} & ((\neg full \wedge drank \wedge \neg holding) \vee (full \wedge \neg drank \wedge \neg holding)) \rightarrow \neg(obsLight \mid weigh)^\diamond; \\ & ((\neg full \wedge drank \wedge \neg holding) \vee (full \wedge \neg drank \wedge \neg holding)) \rightarrow \neg(obsHeavy \mid weigh)^\diamond; \\ & ((\neg full \wedge drank \wedge \neg holding) \vee (full \wedge \neg drank \wedge \neg holding)) \rightarrow \neg(obsMedium \mid weigh)^\diamond. \end{aligned}$$

The *perceivability condition closure axiom* schema is

$$\begin{aligned} & \neg(\phi_{11} \vee \dots \vee \phi_{1j}) \rightarrow \neg(\varsigma_1 \mid \alpha)^\diamond; \\ & \neg(\phi_{21} \vee \dots) \rightarrow \neg(\varsigma_2 \mid \alpha)^\diamond; \quad \dots; \\ & \neg(\dots \vee \phi_{nk}) \rightarrow \neg(\varsigma_n \mid \alpha)^\diamond, \end{aligned}$$

where the ϕ_i are taken from the perceivability axioms. There are no perceivability closure axioms for ontic actions, because they are always tautologies.

Ontic actions each have *unperceivability axioms* of the form $(\forall v^\varsigma)((v^\varsigma \mid \alpha)^\diamond \leftrightarrow v^\varsigma = obsNil)$. The axiom says that no other observation is perceivable given the ontic action. That is, for any instantiation of an observation ς' other than *obsNil*, $\neg(\varsigma' \mid \alpha)^\diamond$ is a logical consequence.

For sensory actions, to state that the observations not associated with action α are always impossible given α was executed, we need an axiom of the form $(\forall v^\varsigma)(v^\varsigma \neq o_1 \wedge v^\varsigma \neq o_2 \wedge \dots \wedge v^\varsigma \neq o_n) \rightarrow \neg(v^\varsigma \mid \alpha)^\diamond$. For the oil-can scenario, they are

$$\begin{aligned} & (\forall v^\varsigma)(v^\varsigma \mid grab)^\diamond \leftrightarrow v^\varsigma = obsNil; \\ & (\forall v^\varsigma)(v^\varsigma \mid drink)^\diamond \leftrightarrow v^\varsigma = obsNil; \\ & (\forall v^\varsigma)(v^\varsigma \mid replace)^\diamond \leftrightarrow v^\varsigma = obsNil; \\ & (\forall v^\varsigma)(v^\varsigma \neq obsHeavy \wedge v^\varsigma \neq obsLight \wedge \\ & \quad v^\varsigma \neq obsMedium) \rightarrow \neg(v^\varsigma \mid weigh)^\diamond. \end{aligned}$$

5.3 The Utility Function

A sufficient set of axioms concerning ‘state rewards’ and ‘action costs’ constitutes a *utility function*.

There must be a means to express the reward an agent will get for performing an action in a world it may find itself—for every action and every possible world. The domain expert must supply a set of *reward axioms* of the form $\phi_i \rightarrow \text{Reward}(r_i)$, where ϕ_i is a condition specifying the world in which the rewards can be got (e.g., $\text{holding} \rightarrow \text{Reward}(5)$ and $\text{drank} \rightarrow \text{Reward}(10)$).

The conditions of the reward axioms must identify worlds that are pairwise disjoint. This holds for cost axioms too:

The domain expert must also supply a set of *cost axioms* of the form $(\phi_i \wedge \langle \alpha \rangle \top) \rightarrow \text{Cost}(\alpha, c_i)$, where ϕ_i is a condition specifying the world in which the cost c_i will be incurred for action α . For example,

$$\begin{aligned} (\text{full} \wedge \langle \text{grab} \rangle \top) &\rightarrow \text{Cost}(\text{grab}, 2); \\ (\neg \text{full} \wedge \langle \text{grab} \rangle \top) &\rightarrow \text{Cost}(\text{grab}, 1); \\ (\text{full} \wedge \langle \text{drink} \rangle \top) &\rightarrow \text{Cost}(\text{drink}, 2); \\ (\neg \text{full} \wedge \langle \text{drink} \rangle \top) &\rightarrow \text{Cost}(\text{drink}, 1); \\ \langle \text{replace} \rangle \top &\rightarrow \text{Cost}(\text{replace}, 0.8). \end{aligned}$$

5.4 A Frame Solution

The method we propose for avoiding generating all the frame and effect closure axioms, is to write the effect and executability axioms, generate the uniform axioms, and then generate a set of a new kind of axioms representing the frame and effect closure axioms much more compactly. By looking at the effect axioms of a domain, one can define for each fluent $p \in \mathfrak{F}$ a set $\text{Cause}^+(p)$ of actions that can (but not necessarily always) cause p (as a positive literal) to flip to $\neg p$, and a set $\text{Cause}^-(p)$ of actions that can (but not necessarily always) causes $\neg p$ (as a negative literal) to flip to p .¹ For instance, $\text{grab} \in \text{Cause}^+(\text{full})$, because in effect axiom

$$\begin{aligned} (\text{full} \wedge \neg \text{holding}) &\rightarrow \\ ([\text{grab}]_{0.7}(\text{full} \wedge \text{holding}) \wedge \\ &[\text{grab}]_{0.2}(\neg \text{full} \wedge \neg \text{holding}) \wedge [\text{grab}]_{0.1}(\text{full} \wedge \neg \text{holding})), \end{aligned}$$

grab flips full to $\neg \text{full}$ (with probability 0.2). The axiom also shows that $\text{grab} \in \text{Cause}^-(\text{holding})$ because it flips $\neg \text{holding}$ to holding (with probability 0.7). The actions mentioned in these sets may have deterministic or stochastic effects on the respective propositions.

Furthermore, by looking at the effects axioms, Cond functions can be defined: For each $\alpha \in \text{Cause}^+(p)$, $\text{Cond}^+(\alpha, p)$ returns a sentence that represents the disjunction of all ϕ_i under which α caused p to be a negative literal. $\text{Cond}^-(\alpha, p)$ is defined similarly.

Suppose that $\text{Cause}^+(p) = \{\alpha_1, \dots, \alpha_m\}$ and $\text{Cause}^-(p) = \{\beta_1, \dots, \beta_n\}$. We propose, for any fluent p , a pair of *compact frame axioms* with schema

$$\begin{aligned} (\forall v^\alpha) p &\rightarrow \\ (v^\alpha = \alpha_1 \wedge \neg \text{Cond}^+(\alpha_1, p)) &\rightarrow [\alpha_1] p \wedge \\ &\vdots \\ (v^\alpha = \alpha_m \wedge \neg \text{Cond}^+(\alpha_m, p)) &\rightarrow [\alpha_m] p \wedge \\ (v^\alpha \neq \alpha_1 \wedge \dots \wedge v^\alpha \neq \alpha_m) &\rightarrow [v^\alpha] p \end{aligned}$$

¹Such sets and functions are also employed by Demolombe, Herzig and Varzinczak [Demolombe *et al.*, 2003].

and

$$\begin{aligned} (\forall v^\alpha) \neg p &\rightarrow \\ (v^\alpha = \beta_1 \wedge \neg \text{Cond}^-(\beta_1, p)) &\rightarrow [\beta_1] \neg p \wedge \\ &\vdots \\ (v^\alpha = \beta_m \wedge \neg \text{Cond}^-(\beta_m, p)) &\rightarrow [\beta_m] \neg p \wedge \\ (v^\alpha \neq \beta_1 \wedge \dots \wedge v^\alpha \neq \beta_m) &\rightarrow [v^\alpha] \neg p. \end{aligned}$$

Claim 5.1 *The collection of pairs of compact frame axioms for each fluent in \mathfrak{F} is logically equivalent to the collection of all conditional frame axioms and effect closure axioms generated with the processes presented above.*

Proof:

Please refer to our draft report [Rens and Meyer, 2011]. Q.E.D.

There are in the order of $|\mathfrak{A}| \cdot 2|\Omega| \cdot D$ frame axioms, where D is the average number of conditions on effects per action (the ϕ_i). Let N be the average size of $|\text{Cause}^+(p)|$ or $|\text{Cause}^-(p)|$ for any $p \in \mathfrak{F}$. With the two compact frame axioms (per fluent), no separate frame or effect closure axioms are required in the action description (AD). If we consider each of the most basic conjuncts and disjuncts as a unit length, then the size of each compact frame axiom is $\mathcal{O}(N)$, and the size of all compact frame axioms in AD is in the order of $N \cdot 2|\mathfrak{F}|$. For reasonable domains, N will be much smaller than $|\mathfrak{A}|$, and the size of all compact frame axioms is thus much smaller than the size of all frame and effect closure axioms ($|\mathfrak{A}| \cdot 2|\mathfrak{F}| \cdot (D + 1)$).

5.5 Some Example Entailment Results

The following entailments have been proven concerning the oil-can scenario [Rens and Meyer, 2011]. BK^{oc} is the background knowledge of an agent in the scenario. To save space and for neater presentation, we abbreviate constants and fluents by their initials.

$$BK^{oc} \models_{GS} (f \wedge d \wedge \neg h) \rightarrow [g]_{0.7}(f \wedge d \wedge h):$$

If the can is full and the oil has been drunk, the probability of successfully grabbing it without spilling oil is 0.7.

$$BK^{oc} \models_{GS} (f \wedge \neg d \wedge h) \rightarrow \neg [d]_{0.2}(f \vee \neg d \vee \neg h):$$

If the robot is in a situation where it is holding the full oil-can (and has not yet attempted drinking), then the probability of having failed to drink the oil is not 0.2.

$$BK^{oc} \models_{GS} (\exists v^s)(v^s \mid \text{drink})^\square:$$

In any world, there always exists an observation after the robot has drunk.

$$BK^{oc} \models_{GS} \langle d \rangle \top \leftrightarrow h:$$

In any world, it is possible to drink the oil if and only if the can is being held.

$$BK^{oc} \models_{GS} (f \wedge \langle d \rangle \top) \rightarrow \neg \text{Cost}(d, 3):$$

Assuming it is possible to drink and the can is full of oil, then the cost of doing the drink action is not 3 units.

6 Concluding Remarks

We introduced a formal language specifically for robots that must deal with uncertainty in affection and perception. It is one step towards a general reasoning system for robots, not the actual system.

POMDP theory is used as an underlying modeling formalism. The formal language is based on multi-modal logic and accepts basic principals of cognitive robotics. We have also included notions of probability to represent the uncertainty, but we have done so ‘minimally’, that is, only as far as is necessary to represent POMDPs for the intended application. Beyond the usual elements of logics for reasoning about action and change, the logic presented here adds *observations* as first-class objects, and a means to represent *utility functions*.

In an associated report [Rens and Meyer, 2011], the *frame problem* is addressed, and we provided a *belief network* approach to domain specification for cases when the required information is available.

The computational complexity of SLAOP was not determined, and is left for future work. Due to the nature of SLAOP structures, we conjecture that entailment in SLAOP is decidable. It’s worth noting that the three latter frameworks discussed in Section 2 [Wang and Schmolze, 2005; Sanner and Kersting, 2010; Poole, 1998] do not mention decidability results either.

The next step is to prove decidability of SLAOP entailment, and then to develop a logic for decision-making in which SLAOP will be employed. Domains specified in SLAOP will be used to make decisions in the ‘meta’ logic, with sentences involving sequences of actions and the epistemic knowledge of an agent. This will also show the significance of SLAOP in a more practical context. Please refer to our extended abstract [Rens, 2011] for an overview of our broader research programme.

References

- [Bacchus *et al.*, 1999] F. Bacchus, J. Y. Halpern, and H. J. Levesque. Reasoning about noisy sensors and effectors in the situation calculus. *Artificial Intelligence*, 111(1–2):171–208, 1999.
- [Bonet and Geffner, 2001] B. Bonet and H. Geffner. Planning and control in artificial intelligence: A unifying perspective. *Applied Intelligence*, 14(3):237–252, 2001.
- [Boutilier and Poole, 1996] C. Boutilier and D. Poole. Computing optimal policies for partially observable decision processes using compact representations. In *Proc. of 13th Natl. Conf. on AI*, pages 1168–1175, 1996.
- [Boutilier *et al.*, 2000] C. Boutilier, R. Reiter, M. Soutchanski, and S. Thrun. Decision-theoretic, high-level agent programming in the situation calculus. In *Proc. of 17th Natl. Conf. on AI*, pages 355–362. AAAI Press, Menlo Park, CA, 2000.
- [Demolombe *et al.*, 2003] R. Demolombe, A. Herzig, and I. Varzinczak. Regression in modal logic. *Journal of Applied Non-Classical Logics*, 13(2):165–185, 2003.
- [Fagin and Halpern, 1994] R. Fagin and J. Y. Halpern. Reasoning about knowledge and probability. *J. of ACM*, 41(2):340–367, 1994.
- [Gabaldon and Lakemeyer, 2007] A. Gabaldon and G. Lakemeyer. *ESP*: A logic of only-knowing, noisy sensing and acting. In *Proc. of 22nd Natl. Conf. on AI*, pages 974–979, 2007.
- [Halpern, 2003] J. Y. Halpern. *Reasoning about Uncertainty*. The MIT Press, Cambridge, MA, 2003.
- [Iocchi *et al.*, 2009] L. Iocchi, T. Lukasiewicz, D. Nardi, and R. Rosati. Reasoning about actions with sensing under qualitative and probabilistic uncertainty. *ACM Transactions on Computational Logic*, 10(1):5:1–5:41, 2009.
- [Kaelbling *et al.*, 1998] L. Kaelbling, M. Littman, and A. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.
- [Levesque and Lakemeyer, 2008] H. Levesque and G. Lakemeyer. Cognitive Robotics. In B. Porter F. Van Harmelen, V. Lifshitz, editor, *Handbook of Knowledge Representation*, pages 869–886. Elsevier Science, 2008.
- [Poole, 1998] D. Poole. Decision theory, the situation calculus and conditional plans. *Linköping Electronic Articles in Computer and Information Science*, 8(3), 1998.
- [Rens and Meyer, 2011] G. Rens and T. Meyer. Logic and utility based agent planning language, Part II: Specifying stochastic domains. Technical Report KRR-10-01, KRR, CSIR Meraka Institute, Pretoria, South Africa, January 2011. url: <http://krr.meraka.org.za/publications/2011>.
- [Rens *et al.*, 2010] G. Rens, I. Varzinczak, T. Meyer, and A. Ferrein. A logic for reasoning about actions and explicit observations. In Jiuyong Li, editor, *Proc. of 23rd Australasian Joint Conf. on AI*, pages 395–404, 2010.
- [Rens, 2010] G. Rens. A belief-desire-intention architecture with a logic-based planner for agents in stochastic domains. Master’s thesis, School of Computing, University of South Africa, 2010.
- [Rens, 2011] G. Rens. From an agent logic to an agent programming language for partially observable stochastic domains. In *Proc. of 22nd Intl. Joint Conf. on AI*, Menlo Park, CA, 2011. AAAI Press. To appear.
- [Sanner and Kersting, 2010] S. Sanner and K. Kersting. Symbolic dynamic programming for first-order POMDPs. In *Proc. of 24th Natl. Conf. on AI*, pages 1140–1146, 2010.
- [Shirazi and Amir, 2007] A. Shirazi and E. Amir. Probabilistic modal logic. In *Proc. of 22nd Natl. Conf. on AI*, pages 489–494. AAAI Press, 2007.
- [Van Benthem *et al.*, 2009] J. Van Benthem, J. Gerbrandy, and B. Kooi. Dynamic update with probabilities. *Studia Logica*, 93(1):67–96, 2009.
- [Van Diggelen, 2002] J. Van Diggelen. Using modal logic in mobile robots. Master’s thesis, Cognitive Artificial Intelligence, Utrecht University, 2002.
- [Wang and Schmolze, 2005] C. Wang and J. Schmolze. Planning with POMDPs using a compact, logic-based representation. In *Proc. of 17th IEEE Intl. Conf. on Tools with AI*, pages 523–530, 2005.
- [Weerdt *et al.*, 1999] M. De Weerdt, F. De Boer, W. Van der Hoek, and J.-J. Meyer. Imprecise observations of mobile robots specified by a modal logic. In *Proc. of ASCI-99*, pages 184–190, 1999.