A Probabilistic Finite Temporal Logic for Policies and Intentions

Nima Motamed
Utrecht University
Utrecht, The Netherlands
n.motamed@uu.nl

Natasha Alechina
Utrecht University
Utrecht, The Netherlands
n.a.alechina@uu.nl

Mehdi Dastani
Utrecht University
Utrecht, The Netherlands
m.m.dastani@uu.nl

Dragan Doder
Utrecht University
Utrecht, The Netherlands
d.doder@uu.nl

ABSTRACT

We propose a new probabilistic temporal logic for Markov decision processes allowing us to reason about finite histories and policies. We comment on how the logic can express statements important for a theory of intention revision, show that it possesses the finite model property, and show that the model checking and satisfiability problems for the logic are both decidable.

1 INTRODUCTION

Understanding the revision of agents' intentions is an important part of the study of rational agents that has received much interest in recent decades [11, 13, 15–17]. Intention revision under uncertainty has not received the same interest, contrasted with work done on uncertainty within the closely-related field of belief revision [7, 8, 12, 18].

Intention revision is studied by [17] within the AGM framework of belief revision [1], enabling a representation theorem for revision operators like that of [9]. In order to study intention revision using similar methods in a setting where actions have uncertain outcomes, we require an appropriate temporal logic capable of dealing with the probabilistic setting as a basic framework. Though logics like Probabilistic Strategy Logic [2] and other probabilistic variations of CTL [6] are natural contenders, an appropriate logic would ideally have decidable model checking and satisfiability, which said logics are not known to both possess.

We propose a new logic called Finite MDP-PCTL* which possesses these properties, and comment on how it can express statements required for intention revision in the style of [17] under uncertainty.

2 FINITE MDP-PCTL*

We fix a finite set $\mathscr{A}$ of *actions* and a countable set Prop of *propositional variables.* To each action $a \in \mathscr{A}$ we associate a *precondition* $\mathrm{pre}_a \in \mathrm{Prop}$ and a finite nonempty set $\mathrm{Post}_a \subseteq \mathrm{Prop}$ of *possible postconditions.* While we consider pre- and postconditions of actions to be propositional variables for simplicity, for practical usage it will make more sense to consider them to be conjunctions of literals.

We will write $f \colon X \rightharpoonup Y$ to denote that f is a *partial* function from X to Y, and write $f(x)\downarrow$ to denote that f is defined on input $x \in X$. Given a set X, some $x \in X$ and $n \geqslant 1$, we write X^n_x for the set of all sequences in X of length n starting with x, and we write $X^{\leqslant n}_x = \bigcup_{k=1}^n X^k_x$.

Definition 2.1 (MDP). A *Markov decision process* (MDP) is a tuple $\mathbb{M} = \langle S, P, V\rangle$, where S is the set of *states*, $P \colon S \times \mathscr{A} \rightharpoonup [0,1]^S$ is the *partial probabilistic transition function*, and $V \colon S \to 2^{\mathrm{Prop}}$ is the *valuation.* We often abbreviate $P(s,a)$ by $P_{s,a}$. These are required to satisfy the following conditions.

(i) For all $s \in S$, there is some $a \in \mathscr{A}$ such that $\mathrm{pre}_a \in V(s)$ (or by the following condition, equivalently $P_{s,a}\downarrow$).
(ii) $P_{s,a}\downarrow$ iff $\mathrm{pre}_a \in V(s)$.
(iii) Given $P_{s,a}\downarrow$, (iii.i) for all $t \in S$ such that $P_{s,a}(t) > 0$, there is a unique $p \in \mathrm{Post}_a$ such that $p \in V(t)$, and (iii.ii) for all $p \in \mathrm{Post}_a$ there is a unique t such that $P_{s,a}(t) > 0$ and $p \in V(t)$.

Condition (i) states that MDPs have no deadlocks, and conditions (ii) and (iii) ensure that pre- and postconditions are meaningful: an action is executable precisely when the precondition holds, and the possible outcomes of an action are precisely the postconditions.

Our logic will be built around the notion of a *finite policy*, telling the agent how to act for a certain amount of time steps. These allow us to reason about the probabilities of sequences of states. We limit ourselves to *deterministic* policies as they are simple and sufficient for our goals, though we leave the door open for work considering nondeterministic or probabilistic policies.

Definition 2.2 (Finite policies). Given an MDP $\mathbb{M}$ and $n \geqslant 0$, an *n-step $\mathbb{M}$-policy* is a pair $\langle s, \pi\rangle$ where $s \in S$ is referred to as its *initial state* (and we say the policy is *from* s), and $\pi \colon S^{\leqslant n}_s \to \mathscr{A}$ is a function such that $\mathrm{pre}_{\pi(s_1 \cdots s_k)} \in V(s_k)$ (or equivalently $P(s_k, \pi(s_1 \cdots s_k))\downarrow$) for all finite histories $s_1 \cdots s_k \in S^{\leqslant n}_s$ (where $s_1 = s$). A *finite $\mathbb{M}$-policy* is an n-step $\mathbb{M}$-policy for some $n \geqslant 0$.

We drop the $\mathbb{M}$ when the MDP is clear from context. By slight abuse of notation we usually write π^s instead of the pair $\langle s, \pi\rangle$ or the function π in order to make the initial state explicit. We denote the unique trivial 0-step policy by ε.

Given an n-step policy π^s, its *path distribution* is the probability distribution μ^s_π over S^{n+1}_s defined by putting $\mu^s_\pi(s_1 \cdots s_{n+1}) = \prod_{i=1}^n P(s_i, \pi(s_1 \cdots s_i))(s_{i+1})$.

Our policies are defined with respect to finite state histories. For practical purposes it can be useful to consider policies that consider state-action histories as well. Although we do not expand upon this for simplicity of presentation, we note that all of our work extends to such policies as well.

In order to define the temporal part of our logic, we need to be able to shift policies a time step forward, similar to what is usually done in Strategy Logic [3]. (Note that unlike e.g. [4] our logic is *not* stategic and the setting is not adversarial.)

Definition 2.3 (Shifts). Given an n-step policy π^s for $n \geqslant 1$, and a state $t \in S$, the *t-shift of* π^s is the $(n-1)$-step policy $(\pi^s)^t$ from t, defined as $(\pi^s)^t(ts_1 \cdots s_k) = \pi^s(sts_1 \cdots s_k)$. We will write π^{st} for brevity.

Intuitively, the t-shift of a policy is the result of shifting forward one step in time to a state t.

Definition 2.4 (Syntax). The *language of Finite MDP-PCTL** is inductively defined by the grammar

$$\varphi ::= \bot \mid x \mid \mathrm{do}_a \mid \varphi \wedge \varphi \mid \neg\varphi \mid \Diamond^n_{\bowtie r}\Phi^{n+1},$$

$$\Phi^1 ::= \varphi, \quad \Phi^{n+1} ::= \varphi \mid \Phi^{n+1} \wedge \Phi^{n+1} \mid \neg\Phi^{n+1} \mid \mathsf{X}\Phi^n,$$

where $x \in \mathrm{Prop}$, $a \in \mathscr{A}$, $n \geqslant 1$, $r \in \mathbb{Q} \cap [0,1]$, and $\bowtie \in \{<, \leqslant, =, \geqslant, >\}$. Formulas φ are referred to as *state formulas*, and formulas Φ^n are referred to as *n-path formulas* (or more generally, *path formulas*).

The do_a-proposition should be read as "the agent will now execute a" and $\Diamond^n_{\bowtie r}\Phi$ as "the agent can act in the next n steps in such a way that Φ will hold with probability $\bowtie r$".

Definition 2.5 (Semantics). Given an MDP $\mathbb{M}$, the *semantics of Finite MDP-PCTL** is defined via simultaneous induction over state and path formulas as follows. For state formulas, it is defined w.r.t finite policies π^s as

$$\pi^s \models \bot \quad \text{never}, \quad \pi^s \models x \text{ iff } x \in V(s),$$

$$\pi^s \models \mathrm{do}_a \text{ iff } \pi \neq \varepsilon \text{ and } \pi(s) = a,$$

$$\pi^s \models \Diamond^n_{\bowtie r}\Phi \text{ iff } \mu^s_\rho(\{\mathbf{s} \in S^{n+1}_s \mid \mathbf{s}, \rho^s \models \Phi\}) \bowtie r$$
$$\text{for some } n\text{-step policy } \rho^s,$$

with the semantics of Booleans defined as is standard. For path formulas, the semantics is defined w.r.t. nonempty finite sequences $\mathbf{s} = s_1 \cdots s_n$ of states and $(n-1)$-step policies π from s_1 as

$$\mathbf{s}, \pi^{s_1} \models \varphi \text{ iff } \pi^{s_1} \models \varphi, \quad \mathbf{s}, \pi^{s_1} \models \mathsf{X}\Phi \text{ iff } \mathbf{s}_{\mathsf{X}}, \pi^{s_1 s_2} \models \Phi,$$

where $\mathbf{s}_{\mathsf{X}} = s_2 \cdots s_n$. The semantics of Booleans is defined as is standard.

Note that we could also define the semantics of $\Diamond$ by quantifying over policies that either restrict or extend the current policy. This intuitively corresponds to the agent being strongly committed to his intentions; see [10] for a discussion of commitment strategies.

We now comment on basic applications of the logic to intention revision. Following discussions by [13, 17] who argue for intentions as pairs $\langle a, n\rangle$ of an action a and a time step n at which the agent intends to execute the action, we can define, given a set I of intentions and a 'confidence value' $0 < \theta \leqslant 1$, the formula $\mathrm{exec}_{I,\theta} = \Diamond^{n_{\max}}_{\geqslant\theta} \bigwedge_{\langle a,n\rangle \in I} \mathsf{X}^n \mathrm{do}_a$ (where $n_{\max} = \max_{\langle a,n\rangle \in I} n$), denoting the agent's belief (with confidence θ) that he can execute I. The logic's time being finite in the future also facilitates AGM revision without compactness [17].

Using $\mathrm{exec}_{I,\theta}$, we can relax the approach of [17] in regard to 'coherence' w.r.t. an agent's beliefs, which is a condition that is required for rational agents to adopt intentions. We can define the coherence of I with respect to a set Σ of formulas (representing the agent's beliefs) as the satisfiability of $\Sigma \cup \{\mathrm{exec}_{I,\theta}\}$. This satisfies a variant of 'opportunistic planning', with agents needing to consider it to be sufficiently likely that actions' preconditions will hold in order to adopt intentions. It also avoids the so-called 'Little Nell Problem' [14], since $\mathrm{exec}_{I,\theta}$ only entails that the agent *can* act in a way that makes the preconditions of intentions likely.

Now we show the decidability of the logic, which is important to be able to e.g. determine coherence of intentions as just described. Decidable model checking is trivial, even with our (bounded) memoryful policies (unlike the logic of [2]).

Proposition 2.6. *Determining for a given MDP, finite policy π^s, and state formula φ whether $\pi^s \models \varphi$, is decidable.*

To prove that satisfiability is decidable, we require that we can bound MDP and policy size based on the formula.

Definition 2.7 (Policy depth). The *policy depth* of a state formula is inductively defined by putting $\mathrm{pd}(\varphi) = 1$ if φ is atomic, $\mathrm{pd}(\varphi \wedge \psi) = \max\{\mathrm{pd}(\varphi), \mathrm{pd}(\psi)\}$, $\mathrm{pd}(\neg\varphi) = \mathrm{pd}(\varphi)$, and $\mathrm{pd}(\Diamond^n_{\bowtie r}\Phi) = n + \max\{\mathrm{pd}(\varphi) \mid \varphi \text{ appears in } \Phi\}$.

Observing that by definition all distributions in MDPs are finitely supported, the following is easily verified using standard unraveling techniques from modal logic.

Proposition 2.8. *If a state formula φ is satisfiable, then it is also satisfiable in an MDP with at most as many states as a tree with depth $\mathrm{pd}(\varphi)$ and branching factor $\sum_{a \in \mathscr{A}} |\mathrm{Post}_a|$, at a policy of at most $\mathrm{pd}(\varphi)$ steps.*

Theorem 2.9 (Satisfiability). *Determining given a state formula φ whether there exists an MDP and finite policy π^s such that $\pi^s \models \varphi$, is decidable.*

Proof sketch. We describe a decision procedure. Using the bound on the state count from Proposition 2.8 (which we refer to as y), we iterate over all sets S of size at most y (and write $S = \{s_1, \ldots, s_y\}$), and over valuations V on S considering only the variables appearing in φ (plus all preconditions). We require that for all s_i there is some a such that $\mathrm{pre}_a \in V(s_i)$. Given S and V, we now show that we can define some P such that $\pi^s \models \varphi$ for some policy π^s of at most $\mathrm{pd}(\varphi)$ steps in $\langle S, P, V\rangle$, iff some first-order logic (FOL) sentence α_φ holds in the theory of real closed fields (RCF). We refer to [5] for those unfamiliar with RCF. Using the well-known decidability of RCF, we then get a decision procedure.

We encode P through variables $p_{i,j,a}$ for $1 \leqslant i,j \leqslant y$ and $a \in \mathscr{A}$, denoting $P_{s_i,a}(s_j)$ if that is defined, with $p_{i,j,a} = 0$ otherwise. Writing $\mathbf{p}$ for the sequence of all these variables, we consider a FOL formula $\beta(\mathbf{p})$ that holds in RCF iff the P encoded by $\mathbf{p}$ is well-defined (following the conditions of Definition 2.1).

Next, we consider for every $1 \leqslant n \leqslant \mathrm{pd}(\varphi)$, s_i, $X \subseteq S^{n+1}_{s_i}$, n-step π^{s_i}, $\bowtie$ and $r \in \mathbb{Q} \cap [0,1]$, a FOL formula $\gamma_{i,X,\pi,\bowtie,r}(\mathbf{p})$ such that $\beta(\mathbf{p})$ and $\gamma_{i,X,\pi,\bowtie,r}(\mathbf{p})$ hold in RCF iff $\mu^{s_i}_\pi(X) \bowtie r$ in the MDP encoded by $\mathbf{p}$. Note that $\mu^{s_i}_\pi(X) \bowtie r$ is an inequality with a sum of products, expressible in RCF.

We then define FOL formulas $\delta_{\psi,i,\pi}$, $\kappa_{\Phi,\mathbf{w},\pi}$, and $\lambda_{\Phi,i,n,\pi,X}$, respectively expressing "$\pi^{s_i} \models \psi$", "$\pi^{w_1}, \mathbf{w} \models \Phi$", and "$X$ is the set of all $\mathbf{w} \in S^{n+1}_{s_i}$ such that $\pi^{s_i}, \mathbf{w} \models \Phi$". The definitions of these are simple, and almost directly follow the semantics of our logic, with e.g. (i) $\delta_{\Diamond^n_{\bowtie r}\Phi,i,\pi} = \bigvee_{n\text{-step } \rho^{s_i}} \bigwedge_{X \subseteq S^{n+1}_{s_i}} \lambda_{\Phi,i,n,\rho,X} \rightarrow \gamma_{i,X,\rho,\bowtie,r}$, (ii) $\kappa_{\mathsf{X}\Phi,\mathbf{w},\pi} = \kappa_{\Phi,\mathbf{w}_{\mathsf{X}},\pi^{w_1 w_2}}$, and (iii) $\lambda_{\Phi,i,n,\pi,X} = \bigwedge_{\mathbf{w} \in X} \kappa_{\Phi,\mathbf{w},\pi} \wedge \neg \bigvee_{\mathbf{w} \in S^{n+1}_{s_i} \setminus X} \kappa_{\Phi,\mathbf{w},\pi}$.

Finally, we put $\alpha_\varphi = \bigvee_{n \leqslant \mathrm{pd}(\varphi)} \bigvee_{n\text{-step } \pi^{s_i}} \exists\mathbf{p}\big(\beta(\mathbf{p}) \wedge \delta_{\varphi,i,\pi}(\mathbf{p})\big)$, completing the proof. ◻

REFERENCES

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. 1985. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *The Journal of Symbolic Logic* 50, 2 (1985), 510–530. https://doi.org/10.2307/2274239

[2] Benjamin Aminof, Marta Kwiatkowska, Bastien Maubert, Aniello Murano, and Sasha Rubin. 2019. Probabilistic Strategy Logic. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI-19)*, Sarit Kraus (Ed.). 32–38. https://doi.org/10.24963/ijcai.2019/5

[3] Patricia Bouyer, Patrick Gardy, and Nicolas Markey. 2016. On the semantics of Strategy Logic. *Inform. Process. Lett.* 116, 2 (2016), 75–79. https://doi.org/10.1016/j.ipl.2015.10.004

[4] Nils Bulling and Wojciech Jamroga. 2009. What Agents Can Probably Enforce. *Fundamenta Informaticae* 93, 1-3 (2009), 81–96. https://doi.org/10.3233/FI-2009-0089

[5] C. C. Chang and H. Jerome Keisler. 2012. *Model Theory* (3 ed.). Dover Publications.

[6] Hans Hansson and Bengt Jonsson. 1994. A logic for reasoning about time and reliability. *Formal Aspects of Computing* 6, 5 (1994), 512–535. https://doi.org/10.1007/BF01211866

[7] Sven Ove Hansson. 2020. Revising Probabilities and Full Beliefs. *Journal of Philosophical Logic* 49, 5 (2020), 1005–1039. https://doi.org/10.1007/s10992-020-09545-w

[8] Sven Ove Hansson. 2022. A Characterization of Probability-based Dichotomous Belief Revision. *Studia Logica* 110, 2 (2022), 511–543. https://doi.org/10.1007/s11225-021-09961-2

[9] Hirofumi Katsuno and Alberto O. Mendelzon. 1991. Propositional knowledge base revision and minimal change. *Artificial Intelligence* 52, 3 (1991), 263–294. https://doi.org/10.1016/0004-3702(91)90069-V

[10] Anand S. Rao and Michael P. Georgeff. 1991. Modeling Rational Agents within a BDI-Architecture. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, James F. Allen, Richard Fikes, and Erik Sandewall (Eds.). Morgan Kaufmann, 473–484.

[11] Anand S. Rao and Michael P. Georgeff. 1998. Decision Procedures for BDI Logics. *Journal of Logic and Computation* 8, 3 (1998), 293–343. https://doi.org/10.1093/logcom/8.3.293

[12] Gavin Rens, Thomas Meyer, Gabriele Kern-Isberner, and Abhaya Nayak. 2018. Probabilistic Belief Revision via Similarity of Worlds Modulo Evidence. In *Proceedings of the 41st German Conference on AI (KI 2018)*, Frank Trollmann and Anni-Yasmin Turhan (Eds.). Springer, 343–356. https://doi.org/10.1007/978-3-030-00111-7_29

[13] Yoav Shoham. 2009. Logical Theories of Intention and the Database Perspective. *Journal of Philosophical Logic* 38, 6 (2009), 633–647. https://doi.org/10.1007/s10992-009-9116-8

[14] Richmond H. Thomason. 2018. The Little Nell Problem: reasonable and resolute maintenance of agent intentions. *Synthese* 195, 1 (2018), 433–440. https://doi.org/10.1007/s11229-016-1229-3

[15] Wiebe van der Hoek, Wojciech Jamroga, and Michael Wooldridge. 2007. Towards a theory of intention revision. *Synthese* 155, 2 (2007), 265–290. https://doi.org/10.1007/s11229-006-9145-6

[16] Hans van Ditmarsch, Tiago de Lima, and Emiliano Lorini. 2011. Intention Change via Local Assignments. In *Languages, Methodologies, and Development Tools for Multi-Agent Systems*, Mehdi Dastani, Amal El Fallah Seghrouchni, Jomi Hübner, and João Leite (Eds.). Springer, Berlin, Heidelberg, 136–151. https://doi.org/10.1007/978-3-642-22723-3_8

[17] Marc van Zee, Dragan Doder, Leendert van der Torre, Mehdi Dastani, Thomas Icard, and Eric Pacuit. 2020. Intention as commitment toward time. *Artificial Intelligence* 283, Article 103270 (2020). https://doi.org/10.1016/j.artint.2020.103270

[18] Zhiqiang Zhuang, James Delgrande, Abhaya Nayak, and Abdul Sattar. 2017. A Unifying Framework for Probabilistic Belief Revision. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-17)*, Carles Sierra (Ed.). 1370–1376. https://doi.org/10.24963/ijcai.2017/190