



Imprecise Probabilistic Model Checking for Stochastic Multi-agent Systems

Alberto Termine^{1,2} · Alessandro Antonucci² · Giuseppe Primiero¹ · Alessandro Facchini²

Received: 4 February 2022 / Accepted: 26 March 2023
© The Author(s) 2023

Abstract

Standard techniques for model checking stochastic multi-agent systems usually assume the transition probabilities describing the system dynamics to be *stationary* and *completely specified*. As a consequence, neither non-stationary systems nor systems whose stochastic behaviour is partially unknown can be treated. So far, most of the approaches proposed to overcome this limitation suffer from complexity issues making them poorly efficient in the case of large state spaces. A fruitful but poorly explored way out is offered by the formalism of *imprecise probabilities* and the related *imprecise Markov models*. The aim of this paper is to show how imprecise probabilities can be fruitfully involved to model-check multi-agent systems characterised by non-stationary behaviours. Specifically, the paper introduces a new class of multi-agent models called *Imprecise Probabilistic Interpreted Systems* and their relative extensions with rewards. It also introduces a proper logical language to specify properties of such models and corresponding model checking algorithms based on iterative procedures to compute probabilistic and epistemic inferences over imprecise Markov models.

Keywords Probabilistic model checking · Imprecise Markov chains · Probabilistic interpreted systems · Imprecise probabilities

Introduction

Probabilistic model checking is a verification technique to analyse the behaviour of stochastic agents and check whether they satisfy some given desirable properties. It relies on: (i) a formal model of the agent; (ii) a formal specification of the desirable properties; and (iii) algorithms to perform an exhaustive model exploration and verify whether the specified properties are met.

Markov models are the reference formalism usually involved in modelling the stochastic behaviour of agents.

These include both discrete and continuous-time Markov chains and their several extensions [1]. Different approaches exist to specify properties of these models: graph-oriented approaches based on non-deterministic state automata [2, 3], algebraic approaches based on formalisms like the *performance evaluation process algebra* [4, 5], and approaches based on logical languages, such as the *Probabilistic Computational Tree Logic* (PCTL) [6] and its various extensions [2]. Here, we limit our focus to logical approaches and the related semantics and algorithms.

Probabilistic model checking is applied to a variety of fields, ranging from software verification [7] and communication protocols [8–10], to service oriented architectures [11–13] and computational biology [14, 15]. Moreover, different frameworks have been proposed so far to cope with stochastic multi-agent systems. The latter are generally based on Kripke-like semantic frames and related epistemic languages [16]. Among them, the *computationally-grounded weighted doxastic logic* proposed in [17] extends the well-known *Computation Tree Logic* (CTL) [2] with a weighted doxastic operator to specify single and multi-agent beliefs. Similarly, the *Probabilistic Computation Tree Logic of Knowledge* (PCTLK) [18] is a PCTL extension including

This article is part of the topical collection “Advances in Multi-Agent Systems Research: EUMAS 2021 Extended Selected Papers” guest edited by Davide Grossi, Ariel Rosenfeld and Nimrod Talmon.

✉ Alberto Termine
alberto.termine@idsia.ch

¹ Logic Uncertainty Information and Computation (LUCI) lab, Department of Philosophy “Piero Martinetti”, University of Milan, Via Festa del Perdono, 7, 20122 Milan, Lombardy, Italy

² Dalle Molle Institute for Artificial Intelligence Research (IDSIA USI-SUPSI), Polo universitario Lugano, Campus Est, Via la Santa, 1, Lugano-Viganello, 6962 Lugano, Ticino, Switzerland

single and multi-agent epistemic modalities to specify different epistemic-probabilistic properties of stochastic multi-agent systems. Both formalisms base their semantics on *probabilistic interpreted systems*, a class of structures obtained by merging Kripke structures [16] with Markov models.

Despite its success, probabilistic model checking suffers from a well-known limitation: it requires all the transition probabilities to be defined by “sharp” (or, “precise”) numerical values. This constraint might be critical in several applications as it prevents from modelling both non-stationary agents and agents characterised by a partial uncertainty on the values of transition probabilities. An appealing way to overcome this limitation is represented by so-called *parametric Markovian models* [19], where precise state transition probabilities are replaced with unknown parameters. This is the solution adopted, for instance, in [20]. However, complexity issues related with the corresponding model checking procedure, based on fraction-free Gaussian elimination, limit its applicability only to models of small size.

A less explored alternative is provided by the formalism of imprecise probabilities [21] and related imprecise Markov models, namely *imprecise Markov chains* (IMC) and their extensions [22–25]. The latter can be seen as the imprecise counterparts of standard Markov chains and are obtained by replacing single-valued probability distributions with so-called *credal sets*, i.e. sets of probability distributions compatible with given constraints [26].

A first attempt to extend probabilistic model checking to the framework of imprecise probabilities has been proposed in [22]. The paper introduces an imprecise version of PCTL based on IMCs and proves that shifting from precise to imprecise models does not increase the time complexity of the relevant model checking tasks, which remains polynomial in the number of states of the models. A first extension of the results in [22] is provided in [25], introducing a specific language and the corresponding model checking procedure for *imprecise Markov reward models* (IMRM) partially based on the recursive procedures outlined in [27].

In both [22, 25], only single-agent systems are considered. A multi-agent extension of this formalism is instead presented in [28]. In this paper, the authors introduce a class of structures to model epistemic-stochastic multi-agent systems called *imprecise probabilistic interpreted systems* (IPIS) together with a language to specify their properties (called EIPCTL) and relative model checking procedure.

The present work introduces a new framework that combines and extends the results in [25, 28]. The framework is based on *Imprecise Probabilistic Interpreted Reward Systems* (IPIRS), a class of imprecise-probabilistic models obtained by combining IPISs presented in [28] with IMRMs introduced in [25]. A language to specify properties of these models, called *Epistemic Imprecise Probabilistic Interpreted*

Reward Systems (EIPRCTL) is consequently introduced. The latter is obtained by extending the EIPCTL introduced in [28] with operators specific for reward properties. Corresponding model checking algorithms are then developed based on iterative scheme exploiting the same *transition operator* used in [22] for IMCs. Furthermore, preliminary computational complexity results included in [22] and [28] are generalised, hence proving that shifting from precise to imprecise models does not increase the time complexity of the relevant model checking tasks with respect to standard (precise) models. The developed formalism and algorithms are finally tested on a case study borrowed from the medical domain.

The paper is structured as follows. In “**Markov Models**”, we provide a general introduction to different kinds of Markov models and related probabilistic inferences. In “**Imprecise Markov Models**”, we introduce different relevant kinds of imprecise Markov models and the corresponding methods to compute probabilistic inferences over them. In “**Epistemic Imprecise PRCTL**”, we define the syntax and semantics of EIPRCTL. In “**Model Checking**”, we develop a model checker for EIPRCTL exploiting procedures for the computation of probabilistic and epistemic inferences in imprecise Markov models introduced in “**Imprecise Markov Models**”. In “**A Case Study on Healthcare Budgeting**”, we offer an example of application based on a case study borrowed from the medical domain. Finally, in “**Conclusions**”, we conclude with some remarks about further potential applications and developments of our framework. The proofs of the theorems are all reported in the Appendix A.

Markov Models

Markov Chains

Let \mathcal{S} be a finite non-empty set of possible states. We are interested in modelling stochastic agents that, at each discrete-time $t \in \mathbb{N}$, shift from a state $s \in \mathcal{S}$ to another, not necessarily different, state $s' \in \mathcal{S}$. We assume the stochastic behaviour of an agent to be time-homogeneous, that is, the probability of a transition from s to s' is independent of the time t at which it occurs, and memory-less, that is, the probability of each transition is independent of the previously occurred transitions. Under these conditions, the behaviour of the agent can be described in terms of a *discrete-time Markov chain* (DTMC).

Definition 1 (*Discrete-time Markov chain*) A discrete-time Markov chain M_{DTMC} is a tuple:

$$M_{\text{DTMC}} := \langle \mathcal{S}, T, \iota \rangle,$$

where:

- \mathcal{S} is a finite non-empty set of states;
- $T : \mathcal{S} \times \mathcal{S} \rightarrow [0, 1]$ is a transition matrix that assigns a probability value to each transition $s, s' \in \mathcal{S} \times \mathcal{S}$;
- $\iota : \mathcal{S} \rightarrow [0, 1]$ is a probability distribution that assigns an initial probability value to each $s \in \mathcal{S}$.

Given a DTMC M_{DTMC} , we call *path* a function $\pi : \mathbb{N} \mapsto \mathcal{S}$ whose values are the states reached by M_{DTMC} at the various time-steps t . Accordingly, each path π describes a possible temporal evolution of the Markov chain and corresponds to an infinite countable sequence of states. In what follows, notation $\pi(t)$ is used to refer to the state of the path π at time t , while $\text{Paths}(s)$ denotes the set of all paths π originating in a given state $s \in \mathcal{S}$ (i.e. such that $\pi(0) = s$). The set of all possible paths π of a given DTMC M_{DTMC} is denoted by $\Pi^{M_{DTMC}}$ and represents the set of all possible outcomes in the temporal evolution of the Markov chain.

To relate paths with probabilities, we endow $\Pi^{M_{DTMC}}$ with a σ -algebra and augment it to a probability space as follows. Given a path π of M_{DTMC} , a *finite prefix* $\hat{\pi}$ of π is any sequence $(\pi(0), \dots, \pi(t))$ originating in $\pi(0)$ and including a finite number of subsequent states of π . The set of all finite prefixes of a given path π is denoted by $\text{pref}(\pi)$, while the set of all finite prefixes $\hat{\pi}$ originating in a given state $s \in \mathcal{S}$ is denoted by $\text{Paths}_{fin}(s)$ (see, [2, Sec. 10.1]).

Definition 2 (Cylinder set) The *cylinder set* $Cyl(\hat{\pi})$ induced by a finite prefix $\hat{\pi}$ is defined as

$$Cyl(\hat{\pi}) := \{ \pi \in \Pi^{M_{DTMC}} \mid \hat{\pi} \in \text{pref}(\pi) \}.$$

That is, $Cyl(\hat{\pi})$ is the set of all paths whose common prefix is $\hat{\pi}$ [2, Def. 10.9].

Definition 3 (σ -algebra of a Markov Chain) The σ -algebra associated with a DTMC M_{DTMC} , denoted $\sigma^{M_{DTMC}}$, is the smallest σ -algebra that contains all cylinder sets $Cyl(\hat{\pi})$, where $\hat{\pi}$ ranges over all finite prefixes of M_{DTMC} .

From basic concepts of probability and Definition 3, it follows that there exists a unique probability measure $P^{M_{DTMC}}$ on the σ -algebra $\sigma^{M_{DTMC}}$ such that (see, [2, Sec. 10.1]):

$$P^{M_{DTMC}}(Cyl(\hat{\pi}(0), \dots, \hat{\pi}(t))) = \iota(\hat{\pi}(0)) \cdot P^{M_{DTMC}}(\hat{\pi}(0), \dots, \hat{\pi}(t)), \tag{1}$$

where:

$$P^{M_{DTMC}}(\hat{\pi}(0), \dots, \hat{\pi}(t)) := \prod_{\tau=0}^{t-1} T(\hat{\pi}(\tau), \hat{\pi}(\tau + 1)), \tag{2}$$

while for finite prefixes composed by just one state (i.e. $\hat{\pi} = \{s\}$), $P_s(\hat{\pi}) = 1$ [2, Sec. 10.1].

Also of interest are the probabilities of a path π , respectively a finite prefix $\hat{\pi}$, *conditional* on a given initial state $s \in \mathcal{S}$, which are henceforth denoted by $P_s^{M_{DTMC}}(\pi)$, respectively $P_s^{M_{DTMC}}(\hat{\pi})$.

Over the probability space $\langle \Pi^{M_{DTMC}}, \sigma^{M_{DTMC}}, P^{M_{DTMC}} \rangle$, we define a family $\{S_t\}_{t \in \mathbb{N}}$ of categorical stochastic variables S_t that range over \mathcal{S} and describe the temporal behaviour of the Markov chain. In this framework, the memory-less condition above mentioned corresponds to the *Markov property*, which establishes that: $P^{M_{DTMC}}(S_{t+1} \mid S_t, \dots, S_0) = P^{M_{DTMC}}(S_{t+1} \mid S_t)$. Time-homogeneity, on the other hand, corresponds to assume $P^{M_{DTMC}}(S_{t+1} \mid S_t)$ to be the same for all t .

Let us now consider the usual definition of (discrete-time) *stochastic process*.

Definition 4 (Stochastic process) Given a finite non-empty set of states \mathcal{S} , a discrete-time *stochastic process* over \mathcal{S} , here denoted M , is a family of categorical stochastic variables $\{S_t\}_{t \in \mathbb{N}}$ ranging over \mathcal{S} and defined over a probability space $\langle \Pi, \sigma(\Pi), P^M \rangle$ such that

- Π is the set of all paths generated by states in \mathcal{S} ,
- $\sigma(\Pi)$ is the cylinder σ -algebra of Π ,
- P^M is a probability measure over $\sigma(\Pi)$.

From Definition 4, it follows that each stochastic process M is uniquely identified by (i.e. it is in one-to-one correspondence with) a probability measure P^M . Accordingly, a DTMC with state-space \mathcal{S} , initial distribution ι and transition matrix T can be alternatively regarded as the (discrete-time) stochastic process over \mathcal{S} uniquely identified by the probability measure P^M generated by ι and T as by Eqs. (1) and (2).

Definition 5 (Labelled DTMC) A *labelled DTMC* is a DTMC extended with a set of atomic propositions AP and a labelling function $l : \mathcal{S} \rightarrow 2^{AP}$ that assigns to each $s \in \mathcal{S}$ a set of atomic propositions $l(s) \subseteq AP$ representing elementary facts (or properties) holding in that state.

In the rest of this article, we consider only *labelled DTMCs*. When not differently specified, we use the term DTMC to refer directly to their labelled extensions.

Inferences in Markov Chains

We next recall two probabilistic inferences that are of central interest in this work, i.e. *marginal* and *hitting* probability.

Definition 6 (Marginal probability) Given an event $B \subseteq \mathcal{S}$ and an initial state $s \in \mathcal{S}$, the marginal probability of B with respect to time t conditional on $S_0 = s$ is defined as follows:

$$P_s^{M_{DTMC}}(S_t \in B) := \sum_{\hat{\pi} := \{\hat{\pi}(0), \dots, \hat{\pi}(t)\} \text{ s.t. } \hat{\pi} \in \text{Paths}_{fin}(s) \wedge \hat{\pi}(t) \in B} P_s^{M_{DTMC}}(\hat{\pi}). \tag{3}$$

Given an event $B \subseteq \mathcal{S}$ and state $s \in \mathcal{S}$, we are also interested in the (conditional) *hitting probability* $h_B(s)$, alternatively called *reachability probability* (see, [2, Sec.10.1.1]), which is the probability of eventually visiting a state $s' \in B$ starting from s .

Consider the event “the process eventually reaches B ” (in model checking literature, this is usually denoted by $\diamond B$). In order to define $h_B(s)$, we need first to characterise the above event as a measurable set of paths. In fact, the latter corresponds to the union of all cylinders $Cyl(\hat{\pi})$ spanned by finite prefixes $\hat{\pi}$ originating in s and such that $\exists t \in \mathbb{N} : \pi(t) \in B \wedge \forall \tau < t, \pi(\tau) \notin B$. Since these sets are pairwise disjoint, the hitting probability of B can be defined as follows [2, Sec.10.1.1.].

Definition 7 (Hitting probability) Given a DTMC M_{DTMC} , a set of states $B \subseteq \mathcal{S}$, and an initial state $s \in \mathcal{S}$,

$$h_B(s) := \sum_{\substack{\hat{\pi} \in \text{Paths}_{fin}(s) : \exists t \in \mathbb{N} \text{ s.t.} \\ \hat{\pi}(t) \in B \wedge \forall \tau < t, \hat{\pi}(\tau) \notin B}} P_s^{M_{DTMC}}(Cyl(\hat{\pi})).$$

Following [22, 24], we compute marginal and hitting probabilities via a *transition operator* \hat{T} defined as follows.

Definition 8 (Transition operator) For any real function f of \mathcal{S} , $\hat{T}f$ is defined as a function $\mathcal{S} \rightarrow \mathbb{R}$ such that

$$\forall s \in \mathcal{S}, (\hat{T}f)(s) := \sum_{s' \in \mathcal{S}} T(s, s') \cdot f(s'). \tag{4}$$

In practice, the transition operator returns the conditional expectation of f , i.e. $\hat{T}f(s) = E[f(S_{t+1}) \mid S_t = s]$ [24]. After t time-steps, the transition operator is obtained as follows:

$$\forall s \in \mathcal{S}, (\hat{T}^t f)(s) := \begin{cases} (\hat{T}f)(s) & \text{if } t = 1, \\ (\hat{T}(\hat{T}^{t-1}f))(s) & \text{if } t > 1, \end{cases} \tag{5}$$

and the respective conditional expectation writes as $E[f(S_t) \mid S_0 = s] = (\hat{T}^t f)(s)$. Since the marginal probability is equivalent to the expectation of the indicator function $\mathbb{1}_B(s)$ that returns one if $s \in B$ and zero otherwise, we can compute it as

$$P_s(S_t \in B) = \hat{T}^t \mathbb{1}_B(s). \tag{6}$$

The hitting probability h_B is obtained instead by computing the minimal non-negative solution¹ of the following system of linear equations [29, Th. 1.3.2]:

$$h_B = \mathbb{1}_B + \mathbb{1}_{B^c} \cdot \hat{T}h_B, \tag{7}$$

where B^c is the complement of B , and sums and products are intended as element-wise operations on arrays. Standard methods solve Eq. (7) in polynomial time with respect to $|\mathcal{S}|$ [2, p.749]. Here we consider an alternative procedure that is easier to be extended to the imprecise-probabilistic framework (see, Section “Imprecise Markov Models”). Let $h_B^t(s)$ be the probability of hitting B from $s \in \mathcal{S}$ within a finite number of time-steps t . For $t = 0$, we trivially have $h_B^t(s) = \mathbb{1}_B(s)$. For $t > 0$, if $s \notin B$, the hitting probability at t is obtained by applying the transition operator to h_B^{t-1} , while if $s \in B$ it is simply set to one. Thus,

$$h_B^t = \mathbb{1}_B + \mathbb{1}_{B^c} \cdot \hat{T}h_B^{t-1}. \tag{8}$$

The hitting probability h_B can thus be computed as the fixed point t^* of the iterative schema in Eq. (8).

The time complexity of the above iterative computation is polynomial with respect to $|\mathcal{S}|t^*$. As each iterative step is based on solving a system of $|\mathcal{S}|$ linear equations, its time complexity results polynomial in $|\mathcal{S}|$. Since t^* further iterations are necessary to reach the fixed point, the overall time complexity will result to be polynomial in $|\mathcal{S}|t^*$.

Markov Reward Models

Among the various extensions of Markov chains, let us consider *Markov reward models* (MRMs) [2]. A MRM is a pair $\langle M, rew \rangle$ composed of a Markov chain M with state space \mathcal{S} and a reward function $rew : \mathcal{S} \mapsto \mathbb{N}$ such that $rew(s)$ represents the reward earned visiting s , for each state $s \in \mathcal{S}$. Given an event $B \subseteq \mathcal{S}$ and a path $\pi \in \Pi^{M_{DTMC}}$, we are interested in the cumulative reward earned along π until visiting an $s \in B$ for the first time. The latter is defined as follows.

Definition 9 (Cumulative reward) Given an event $B \subseteq \mathcal{S}$ and a path $\pi \in \Pi^{M_{DTMC}}$,

$$Rew_B(\pi) := \begin{cases} \sum_{\tau=0}^t rew(\pi(\tau)) & \text{if } \exists t : \pi(t) \in B \wedge \forall \tau < t, \pi(\tau) \notin B \\ \sum_{\tau=0}^{\infty} rew(\pi(\tau)) & \text{otherwise.} \end{cases} \tag{9}$$

Given the above definition, the *expected cumulative reward* earned until reaching B starting from $s \in \mathcal{S}$,

¹ By minimal non-negative solution we mean a solution such that (i) for each $s \in \mathcal{S}$, $h_B(s) \geq 0$, (ii) given any other non-negative solution $h'_B(s)$, then $h_B(s) \leq h'_B(s)$ for each $s \in \mathcal{S}$.

denoted $ExpRew_B(s)$, is now defined as the expectation² of the function Rew_B conditional on initial state $s \in \mathcal{S}$, i.e. $ExpRew_B(s) := E[Rew_B | S_0 = s]$ [2, Def. 10.71]. Let us now discuss more in detail how to compute this value. First of all, we need to recall the following result from [2, Sec. 10.5.1]. Given an event $B \subseteq \mathcal{S}$, let $\mathcal{S}_{=1}^B$ indicates the set of all states $s \in \mathcal{S}$ from which it is possible to reach an $s' \in B$ almost surely, i.e.:

$$\mathcal{S}_{=1}^B := \{s \in \mathcal{S} \mid h_B(s) = 1\}.$$

If $s \notin \mathcal{S}_{=1}^B$, then $ExpRew_B(s)$ may not converge to a finite value. Following [2, sec. 10.5.1], we thus assume for convenience that, by default, $ExpRew_B(s) = \infty$ for all $s \notin \mathcal{S}_{=1}^B$. For all $s \in \mathcal{S}_{=1}^B$, the following result holds [2, Sec. 10.5.1].

Proposition 1 *The values $x_s = E[Rew_B | S_0 = s]$ for each $s \in \mathcal{S}_{=1}^B$ provide the unique solution of the following equations system:*

$$x_s = \begin{cases} rew(s) & \text{if } s \in B, \\ rew(s) + \sum_{s' \in \mathcal{S}_{=1}^B} T(s, s')x_{s'} & \text{otherwise.} \end{cases} \quad (10)$$

There exist several methods to solve the linear system in Eq. (10) (see, [2, Sec.10.5.1]). Here, we adopt a recursive schema similar to that one involved above for hitting probability, which is obtained as follows.

For every $s \in \mathcal{S}_{=1}^B$, let $ExpRew_B^0(s) := rew(s)$, and, for every $t \in \mathbb{N}$, $t \neq 0$, let $ExpRew_B^t(s)$ be defined as

$$ExpRew_B^t(s) := \begin{cases} rew(s) & \text{if } s \in B, \\ rew(s) + \sum_{s' \in \mathcal{S}} T(s, s')ExpRew_B^{t-1}(s') & \text{otherwise.} \end{cases} \quad (11)$$

Notice that the functions $ExpRew_B^t$ are well-defined since if $s \in \mathcal{S}_{=1}^B$, then $s' \in \mathcal{S}_{=1}^B$ for every s' such that $T(s, s') > 0$. Each function $ExpRew_B^t$ can be given a clear interpretation as the expected cumulative reward earned until reaching B from s within a maximum number of time-steps t , as the following result shows.

Theorem 2 *For every $t \in \mathbb{N}$, it holds that*

$$(\forall s \in \mathcal{S}_{=1}^B) \quad ExpRew_B^t(s) = E[Rew_B^t | S_0 = s], \quad (12)$$

where for every $\pi \in \Pi^{M_{DTMC}}$, $Rew_B^0(\pi) := rew(\pi(0))$, and for every $t \in \mathbb{N}$, $t \neq 0$,

$$Rew_B^t(\pi) := \begin{cases} Rew_B(\pi) & \text{if } \exists t^* \leq t : (\forall \tau < t^*) \pi(\tau) \notin B, \pi(t^*) \in B, \\ \sum_{\tau=0}^t rew(\pi(\tau)) & \text{otherwise.} \end{cases} \quad (13)$$

Thanks to Theorem 2, we can now demonstrate the following result proving that the recursive scheme above introduced converges to what we expect.

Theorem 3 *$E[Rew_B | S_0]$ restricted to $\mathcal{S}_{=1}^B$ is a fixed point of the iterative scheme (25).*

As for hitting probability, we can thus compute $ExpRew_B$ by iterating the schema in Eq. (25) over increasing values of t until convergence.

The last MRM inference we consider is the *reward-bounded* hitting probability $h_B^r(s)$, i.e. the probability of reaching B from s before earning a cumulative reward equals to r .

As for standard hitting probability, we proceed by defining the event “the process reaches B before earning a cumulative reward equals to r ” (usually denoted by $\diamond_{\leq r} B$) as a measurable set of paths. Notably, this event corresponds to the union of all cylinder sets $Cyl(\hat{\pi})$ spanned by finite prefixed $\hat{\pi}$ originating in s and such that $\exists t \in \mathbb{N} : \hat{\pi}(t) \in B \wedge \forall \tau < t, \hat{\pi}(\tau) \notin B \wedge rew(\hat{\pi}(0), \dots, \hat{\pi}(\tau)) \leq r$. Since these are pairwise disjoint sets, the reward-bounded hitting probability of B can be defined as follows [2, Sec. 10.5.1]:

Definition 10 (*Reward-bounded hitting probability*) For all $s \in \mathcal{S}$:

$$h_B^r(s) := \sum_{\substack{\hat{\pi} \in \text{Paths}_{fin}(s) : \exists t \in \mathbb{N} \text{ s.t.} \\ \hat{\pi}(t) \in B \wedge \forall \tau < t \hat{\pi}(\tau) \notin B \wedge \\ rew(\hat{\pi}(0), \dots, \hat{\pi}(t)) \leq r}} P_s^{M_{DTMC}}(Cyl(\hat{\pi}))$$

To introduce a method to compute $h_B^r(s)$, let us first recall the following result from [2, Sec. 10.5.1]. Let h_B^ρ denote the vector of reward-bounded hitting probability for a reward-threshold $\rho = 0, \dots, r$. Let $\mathcal{S}_{>0}^B$ be the set of all states $s \in \mathcal{S}$ such that $h_B(s) > 0$, i.e. the set of all states $s \in \mathcal{S}$ such that there exists at least one path π originating in s and reaching B for some t .

Proposition 4 *For each $s \in \mathcal{S}$, the value of $h_B^\rho(s)$ is given by the following system of equations:*

² Notice that Rew_B is not a real-valued function as its value may diverge to ∞ for some $s \in \mathcal{S}$.

$$h_B^\rho(s) = \begin{cases} 1 & \text{if } s \in B \text{ and } rew(s) \leq \rho \\ 0 & \text{if } rew(s) > \rho \text{ or } s \notin \mathcal{S}_{>0}^B \\ \sum_{s' \in \mathcal{S}} T(s, s') h_B^{\rho - rew(s)}(s') & \text{otherwise.} \end{cases} \tag{14}$$

Trivially, $h_B^\rho = 1$ whenever s is already in B and its state-reward does not overcome the desired threshold ρ . Differently, if $s \in B$ but $rew(s) > \rho$, then we have that the cumulative reward earned by the agent already overcomes the specified threshold and, thus, $h_B^\rho(s) = 0$. The same also holds when $s \notin \mathcal{S}_{>0}^B$ as, in this case, we already know that h_B is equal to zero and consequently also $h_B^\rho = 0$. In all the remaining cases, $h_B^\rho(s)$ is computed by iteration over the possible successors of s . That is, we take the sum over all $s' \in \mathcal{S}$ of the probability of reaching s' from s multiplied the probability of reaching B from s' before earning a cumulative reward equal to $\rho - rew(s)$. Remember that every time the agent transits from a state s to one of its successors $s' \in \mathcal{S}$, it earns the reward of s . Consequently, the threshold ρ has to be reduced of a value $rew(s)$ at each further iteration from a state s to its successors s' .

Notice that the system in the above proposition is in fact a linear system with variables (s, ρ) ranging in $\mathcal{S} \times \{0, 1, \dots, \rho\}$. As in the case of standard hitting (see, Eq. (14)), we can solve the system by standard methods [2, Sec. 10.5.1]. Here, we follow a different strategy based on a recursive schema that iterates both over *times* and *rewards*. Let $h_B^{t,\rho}(s)$ denote the probability of hitting B from s before earning cumulative reward ρ and within time-step t . For each $t \in \mathbb{N}$, the values of $h_B^{t,\rho}$ computed for $\rho = 0, \dots, r$ are collected in $\mathcal{S} \times r$ matrix that we denote by $\mathbf{h}_B^{t,\rho_0:r}$.

For $t = 0$, we generate $\mathbf{h}_B^{t=0,\rho_0:r}$ by computing the vectors $h_B^{t=0,\rho}$ for $\rho = 0, \dots, r$ and for each $s \in \mathcal{S}$ as follows:

$$h_B^{t=0,\rho}(s) = \begin{cases} 1 & \text{if } s \in B \text{ and } rew(s) \leq \rho \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Consider that when $t = 0$ no transition occurs. Hence, we clearly have that, for each $s \in \mathcal{A}$, $h_B^{t=0,\rho}(s)$ is equals to *one* if s belongs to the hitting event B and its reward $rew(s)$ does not exceed the specified threshold ρ , otherwise $h_B^{t=0,\rho}(s)$ is *zero*.

For $t > 0$, $\mathbf{h}_B^{t,\rho_0:r}$ is generated by computing the various vectors $h_B^{t,\rho}$ for $\rho = 0, 1, \dots, r$ via the following recursive schema applied to each $s \in \mathcal{S}$:

$$h_B^{t,\rho}(s) = \begin{cases} 1 & \text{if } s \in B \text{ and } rew(s) \leq \rho \\ 0 & \text{if } h_B(s) = 0 \text{ or } rew(s) > \rho \\ \sum_{s' \in \mathcal{S}} T(s, s') h_B^{t-1,\rho - rew(s)}(s') & \text{otherwise.} \end{cases} \tag{16}$$

The first two cases follow from considerations analogous to those leading to Eq. (15). In the third case, we obtain the value of $h_B^{t,\rho}(s)$ from $h_B^{t-1,\rho - rew(s)}$ via one-step application

of the transition operator. As the reward is cumulative, the threshold ρ is decreased by the reward of the current state. The values of $h_B^{t-1,\rho - rew(s)}$ are provided by the matrix $\mathbf{h}_B^{t-1,\rho_0:r}$ that we generate at time-step $t - 1$. Specifically, $h_B^{t-1,\rho - rew(s)}(s')$ corresponds to the cell identified by the $|s'|$ -row and the $|\rho - rew(s)|$ -column of the matrix $\mathbf{h}_B^{t-1,\rho_0:r}$. For all $s \in \mathcal{S}$, the procedure converges to a finite value, as proved by the following theorem.

Theorem 5 *Let $\langle M, rew \rangle$ be a MRM, $B \subseteq [S]$, and $\rho \in \mathbb{N}$. There is a $t^* \in \mathbb{N}$ such that for all $\tau \geq 0$:*

$$\mathbf{h}_B^{t^* + \tau, \rho_0:r} = \mathbf{h}_B^{t^*, \rho_0:r}. \tag{17}$$

As for standard hitting, we can, therefore, compute h_B^r simply by iterating $\mathbf{h}_B^{t,\rho_0:r}$ over increasing values of t until convergence.

Probabilistic Interpreted Systems

Probabilistic Interpreted Systems (PISs) [18] are another MC extension we consider. PISs are a class of semantic frames used in computational logic for modelling epistemic and probabilistic properties of stochastic multi-agent systems. Consider a finite non-empty set of agents \mathcal{A} . The possible configurations of each agent $a \in \mathcal{A}$ are described by a finite non-empty set of *local* states \mathcal{S}^a . The set of *global* states \mathcal{S} describing the possible configurations of the whole multi-agent system is obtained as the Cartesian product $\mathcal{S} := \times_{a \in \mathcal{A}} \mathcal{S}^a$. Accordingly, each $s \in \mathcal{S}$ is a tuple $\langle s^{a_1}, s^{a_2}, \dots, s^{a_n} \rangle$ of $|\mathcal{A}|$ local states. Hence, formally, we have the following definition.

Definition 11 A PIS is defined as a tuple:

$$M_{\text{PIS}} := \langle \mathcal{A}, \mathcal{S}, \{T^a\}_{a \in \mathcal{A}}, \{P^a\}_{a \in \mathcal{A}}, AP, l(s) \rangle, \tag{18}$$

where:

- \mathcal{A} is a finite non-empty set of agents;
- \mathcal{S} is a finite non-empty set of global states;
- $\{T^a\}_{a \in \mathcal{A}}$ is a family of transition matrices $T^a : \mathcal{S} \times \mathcal{S} \rightarrow [0, 1]$;
- $\{P^a\}_{a \in \mathcal{A}}$ is a family of initial probability distributions $P^a : \mathcal{S} \mapsto [0, 1]$;
- AP is a set of atomic propositions;
- $l : \mathcal{S} \rightarrow 2^{AP}$ is the labelling function.

For each agent $a \in \mathcal{A}$, we also introduce an *epistemic equivalence relation* $\sim^a \subseteq \mathcal{S} \times \mathcal{S}$ such that

$$s \sim^a s' \text{ iff } s^a = s'^a.$$

The latter denotes that two global states s, s' are *epistemically indistinguishable* by an agent a if and only if they are identical as far as the agent knows. The equivalence relation \sim^a induces a partition Eq^{\sim^a} over \mathcal{S} . The elements of this partition are denoted by eq^{\sim^a} , and are called *epistemic equivalence classes* (EEC). They consist of sets of global states that are epistemically indistinguishable among each others by a . Specific equivalence relations can be also defined to model different kinds of multi-agent knowledge in a group of agents $\Gamma \subseteq \mathcal{A}$, including:

- *Everybody Knows* $\sim_E^\Gamma := \bigcup_{a \in \Gamma} \sim^a$;
- *Common Knowledge* $\sim_C^\Gamma := it(\bigcup_{a \in \Gamma} \sim^a)$, where *it* denotes the *iterative closure*;
- *Distributed Knowledge* $\sim_D^\Gamma := \bigcap_{a \in \Gamma} \sim^a$.

Each relation induces a partition whose elements are the respective EECs $eq^{\sim_E^\Gamma}$, $eq^{\sim_C^\Gamma}$, and $eq^{\sim_D^\Gamma}$ for groups of agents $\Gamma \subseteq \mathcal{A}$.

From the transition matrices associated with each agent, a global transition matrix T_{PIS} describing the stochastic behaviour of the whole multi-agent system can be obtained by *logarithmic pooling* as follows:

$$T_{PIS}(s, s') := \eta \prod_{a \in \mathcal{A}} T^a(s, s'), \tag{19}$$

where η is a normalising constant.

A *global* initial probability distribution P_{PIS} is similarly defined:

$$P_{PIS}(s) := \eta' \prod_{a \in \mathcal{A}} P^a(s). \tag{20}$$

Furthermore, the global transition matrix T_{PIS} and the global initial probability distribution identify a particular MC $\langle \mathcal{S}, T_{PIS}, P_{PIS}, AP, l \rangle$ called the *embedded* MC of the PIS. The latter is used to compute probabilistic inferences concerning the overall stochastic behaviour of the multi-agent system.

A pair $\langle M_{PIS}, rew \rangle$ of a PIS M_{PIS} and a reward function $rew : \mathcal{S} \mapsto \mathbb{N}$, is called a *Probabilistic Interpreted Reward System* (PIRS).

Imprecise Markov Models

In this section, we provide imprecise-probabilistic counterparts for the Markov models presented in the previous section. This basically corresponds to replace the (sharp) specifications of the probabilistic parameters with set-valued ones. We also show how the efficient inference algorithms described in the previous section can be extended to such generalised setup, thus allowing for the computation of the bounds with respect to the set-valued specification, without increased computational costs. These results partially rely on

recent works about imprecise Markov models [24, 27]. Note that the imprecise Markov models we consider follow the so-called *measure-theoretic* interpretation [24] and relies on the formalism of *credal sets* [30]. The alternative *game-theoretic* formalisation [31] is briefly mentioned without going into details insofar the two formalisms are equivalent for the inference tasks relevant for this work, as proved in [32].

Imprecise Transition Matrices

Given a variable S , a *Credal Set* (CS) $K(S)$ is a set of probability mass functions over S . The upper expectation of a real-valued function f of S with respect to CS $K(S)$ is intended as $\bar{E}[f(S)] := \sup_{P(S) \in K(S)} \sum_{s \in S} f(s) \cdot P(s)$ (the lower expectation $\underline{E}[f(S)]$ is analogously defined). Here we only consider closed and convex CSs induced by a finite number of linear constraints. These are polytopes in the probability simplex with a finite number of extreme points collected in a set $Ext[K(S)]$. For these CSs, upper (lower) expectations can be equivalently obtained by taking the maximum (minimum) with respect to the precise expectations computed on the extreme points. Conditional CSs might be defined analogously [33].

In this framework, an *imprecise* transition matrix \mathcal{T} is defined as a collection of conditional CSs $\{K(S'|s)\}_{s \in \mathcal{S}}$, each one representing a separately specified row of the matrix. This allows for defining precise transition matrices whose rows are obtained by taking a $P(S'|s) \in K(S'|s)$ for each $s \in \mathcal{S}$. Each one of these matrices represents a stochastic behaviour compatible with the “imprecise” specification given by \mathcal{T} .

Imprecise Markov Chains

As a first example of imprecise Markov model, we consider (discrete-time) *imprecise Markov chains* (IMCs), thus providing an imprecise-probabilistic version of the models introduced in Section “Markov Chains”. Note that there exist two main ways of formalising IMCs in the literature. On the one hand, the *measure-theoretic* characterisation defines an IMC as a family of (discrete-time) Markov models compatible with beliefs about initial and transition probabilities. On the other hand, the *game-theoretic* characterisation is grounded on the game-theoretic view of probability popularised in [31] that, applied to the theory of stochastic processes directly leads to imprecise models.³ The two characterisations are different but have been recently proved to coincide for all expectations on the following domains: (i) monotone pointwise limits of finitary real-valued functions, and (ii) bounded below Borel-measurable variables [32]. In this work we focus on measure-theoretic IMCs only.

³ See [34, 35] for an in-depth discussion on the relation between measure-theoretic and game-theoretic IMCs.

However, all the inferences we consider fall under (i) and thus, for the purposes of our work, the two characterisations can be considered equivalent.

Given a CS $K(S_0)$ and an imprecise transition matrix \mathcal{T} , both defined over \mathcal{S} , the (discrete-time) IMC \mathcal{M} induced by $K(S_0)$ and \mathcal{T} can be defined as the largest set of (discrete-time) stochastic processes that are *compatible* with $K(S_0)$ and \mathcal{T} .

The term “compatible” here deserves an exact characterisation. In literature, indeed there exist at least two different criteria for establishing compatibility, which depends on the imprecise interpretation of the notion of stochastic irrelevance, and, consequently, of the Markov property one considers [36]. The two notions of irrelevance typically involved for IMCs are *strong independence* and *epistemic irrelevance*. The former is defined via product-independence of the CS extreme points: we say that $K(S)$ and $K(S')$ are strong-independent if and only if, for all $P(S) \in \text{Ext}[K(S)]$ and all $P(S') \in \text{Ext}[K(S')]$, it holds that $P(S, S') = P(S) \cdot P(S')$. The latter is defined via conditioning: we say that $K(S)$ is epistemically irrelevant for $K(S')$ if and only if $K(S' | s) = K(S')$ for each $s \in \mathcal{S}$.⁴ Notice that, unlike strong independence, epistemic irrelevance is asymmetric, i.e. the irrelevance of $K(S)$ for $K(S')$ does not entail the irrelevance of $K(S')$ for $K(S)$.

Following [27, 38], and also the early work of [22], in this paper we focus on epistemic irrelevance. Notably, by exploiting the results in [30], the imprecise-probabilistic inferences considered in the rest of this paper can easily proved to be independent of the specific characterisation we adopt. Epistemic irrelevance leads to an imprecise characterisation of the Markov property practically corresponding to assume that “whenever the agent knows the current state, then her beliefs about future states are not altered upon learning what states were visited in the past” [36, p. 265]. A formal definition of IMC can thus be given as follows:⁵

Definition 12 (*Imprecise Markov chain under epistemic irrelevance*) Given $K(S_0)$ and \mathcal{T} , an IMC \mathcal{M} (under epistemic irrelevance) is defined as the largest set of all, potentially non-Markov, non-homogeneous, stochastic processes for which, for all $t \in \mathbb{N}$ and all $s_0, \dots, s_t \in \mathcal{S}^t$, there is some $T \in \mathcal{T}$ such that $P(S_{t+1} = s' | S_{0:t} = s_{0:t}) = T(s_{0:t}, s')$ for all $s' \in \mathcal{S}$.

Furthermore, each IMC is uniquely identified by a set of probability measures $P^M : \sigma(\Pi) \rightarrow [0, 1]$ that we denote by K^M . Each $P^M \in K^M$ uniquely identifies a (potentially

non-Markov and non-homogeneous) stochastic process compatible with the IMC identified by K^M . The IMC identified by K^M is also uniquely identified by K_s^M , which is the credal set including all the *conditional* probability distribution P_s^M obtained by conditioning the various $P^M \in K^M$ on a given initial state $s \in \mathcal{S}$. As detailed in the next section, inferences in IMCs are consequently intended as the computation of lower and upper expectations with respect to such credal set.

Before moving on, notice that, as in the precise case, we are interested here in *labelled* IMCs, i.e. IMCs augmented with a finite set of atomic propositions AP and a labelling function $l : \mathcal{S} \rightarrow 2^{AP}$. In what follows, when using the term IMC, we always refer to labelled IMCs.

Inference in Imprecise Markov Chains

To compute inferences in IMCs, let us first introduce the analogous of the transition operator in Eq. (4). This is obtained by taking the bounds with respect to all the possible (precise) specifications of transition probabilities consistent with the imprecise transition matrix of the IMC. For upper bounds, this corresponds to the following non-linear upper operator:

$$\left(\overline{\mathcal{T}}f\right)(s) := \max_{T(s,S') \in \mathcal{T}(s,S')} \sum_{s' \in \mathcal{S}} T(s,s') \cdot f(s'), \tag{21}$$

while an analogous definition, with the minimum replacing the maximum, holds for the lower operator $\underline{\mathcal{T}}$ [27, Eq. 1]. Equation (21) can be computed by solving $|\mathcal{S}|$ linear programming tasks whose feasible regions are the conditional CSs in the definition of \mathcal{T} . This is possible, in particular, because we assume (Section “Imprecise Transition Matrices”) that each row of \mathcal{T} is separately specified and consists of a conditional CS $K(S' | s)$ described by a finite number of linear constraints.

An iterated application of the above operators can be used to compute the bounds of the probability of reaching a given set of states after a number of time-steps t , as shown by the following result.

Theorem 6 *Given an event $B \subseteq \mathcal{S}$ and a time $t \in \mathbb{N}$, let $\overline{P}_s(s_t \in B)$ denote the upper bound for the probability of reaching B after t time-steps when starting from s . It holds that:*

$$\overline{P}_s(s_t \in B) = \overline{\mathcal{T}}^t \mathbb{1}_B(s). \tag{22}$$

A similar result allows to obtain the lower probability by means of the lower operator. For the sake of conciseness, in the rest of the paper, we only report the results for upper probabilities, expectations and transition operators. The lower bounds can always be obtained by replacing the upper transition operator with its lower analogous.

⁴ For a more detailed characterisation of the difference between these notions, we refer to [37].

⁵ This definition is as in [27]. An analogous definition is given in [24].

For what concerns the upper hitting probability, the latter can be regarded as the upper bound of the h_B defined in Eq. (7) with respect to the set K^M , as detailed in the following definition.

Definition 13 (*Upper hitting probability*) Given a IMC \mathcal{M} , a set of states $B \subseteq \mathcal{S}$, and an initial state $s \in \mathcal{S}$:

$$h_B(s) := \max_{P_s^M \in K^M} \sum_{\hat{\pi} \in \text{Paths}_{\text{fin}}(s) : \exists t \in \mathbb{N} \text{ s.t. } \hat{\pi}(t) \in B \wedge \forall \tau < t, \hat{\pi}(\tau) \notin B} P_s^M(\text{Cyl}(\hat{\pi}))$$

The latter can be compute as the minimal solution of the following system of equations [24, Corollary 19] (See also [27]).

$$\bar{h}_B = \mathbb{1}_B + \mathbb{1}_{B^c} \bar{T} \bar{h}_B. \tag{23}$$

Differently from the precise case, the system in Eq. (23) is non-linear and cannot be solved by the standard methods typically used in the precise case. Nevertheless, as we show below, it is possible to apply a schema analogous to that in Eq. (8) and compute \bar{h}_B by recursion over increasing values of t (see, Ref. [24]). Let \bar{h}_B^t denote the upper hitting probability of B for a finite number of time-steps $t \in \mathbb{N}$ conditional on $S_0 = s$. For $t = 0$, we trivially have that $\bar{h}_B^{t=0} = \mathbb{1}_B$. For $t > 0$, we have instead the following recursion:

$$\bar{h}_B^t = \mathbb{1}_B + \mathbb{1}_{B^c} \bar{T} \bar{h}_B^{t-1}. \tag{24}$$

In practice, the procedure consists of t iterated applications of the transition operator \bar{T} and, consequently, requires the solution of $|\mathcal{S}| \cdot t$ linear programming tasks. The time complexity of the procedure is, therefore, polynomial in $|\mathcal{S}| \cdot t$, exactly as in the precise case.

As for standard DTMCs, it is proved that the least fixed point of Eq. (24) is the minimal non-negative solution of the schema in Eq. (23) (see, [24, Prop 16]). We can thus compute \bar{h}_B simply by iterating the schema in Eq. (24) over increasing values of t until converge. The overall time complexity results polynomial in $|\mathcal{S}|t^*$ as in the precise case. Each iteration step is based on a one-step application of the upper transition operator \bar{T} and requires the solution of $|\mathcal{S}|$ linear optimisation tasks: its time complexity is, therefore, polynomial in $|\mathcal{S}|$. As t^* further iterations are necessary to reach convergence, the overall time complexity results polynomial in $|\mathcal{S}|t^*$.

Imprecise Markov Reward Models

The *imprecise Markov reward models* (IMRMs) are the first extension of IMCs we consider. The latter can be defined as the imprecise counterpart of a MRM and

consists of a pair $\langle \mathcal{M}, \text{rew} \rangle$ of a IMC \mathcal{M} and a reward function $\text{rew} : \mathcal{S} \rightarrow \mathbb{R}$. For IMRMs, we characterise the expected cumulative reward ExpRew_B by its upper and lower bounds, respectively denoted $\overline{\text{ExpRew}}_B$ and $\underline{\text{ExpRew}}_B$. As in the precise case, we restrict the latter to only $s \in S_{=1}^B$, where $S_{=1}^B$ is now defined as the set of all $s \in \mathcal{S}$ such that $\bar{h}_B(s) = 1$.

Given an event $B \subseteq \mathcal{S}$ and a path $\pi \in \Pi^{M_{\text{DTIMC}}}$, let us consider the cumulative reward $\text{Rew}_B(\pi)$ earned along π until visiting an $s \in B$ for the first time, as in Definition 9. The *upper $\text{ExpRew}_B(s)$* expected cumulative rewards earned until reaching B starting from $s \in \mathcal{S}$ can be defined as the upper expectation of Rew_B conditional on the initial state $s \in \mathcal{S}$, i.e. $\overline{\text{ExpRew}}_B(s) := \bar{E}[\text{Rew}_B | s]$.

Inspired by Theorem 3 valid for the precise case, we introduce an imprecise version of the recursive scheme presented in Eq. (10). To this end, let $\overline{\text{ExpRew}}_B^0(s) := \text{rew}(s)$ for every $s \in S_{=1}^B$. Instead, for each $t \in \mathbb{N}$, $t \neq 0$, let $\overline{\text{ExpRew}}_B^t(s)$ be defined as follows:

$$\overline{\text{ExpRew}}_B^t(s) := \begin{cases} \text{rew}(s) & \text{if } s \in B, \\ \text{rew}(s) + \left(\bar{T} \overline{\text{ExpRew}}_B^{t-1} \right)(s) & \text{otherwise,} \end{cases} \tag{25}$$

Similarly to the precise case, $\overline{\text{ExpRew}}_B^t$ can be given a clear interpretation via the following theorem.

Theorem 7 For every $t \in \mathbb{N}$, it holds that

$$(\forall s \in S_{=1}^B) \overline{\text{ExpRew}}_B^t(s) = \bar{E}[\text{Rew}_B^t | S_0 = s], \tag{26}$$

where for each $\pi \in \Pi^{M_{\text{DTIMC}}}$, $\text{Rew}_B^0(\pi) := \text{rew}(\pi(0))$, and for each $t \in \mathbb{N}$, $t \neq 0$,

$$\text{Rew}_B^t(\pi) := \begin{cases} \text{Rew}_B(\pi) & \text{if } \exists t^* \leq t : (\forall \tau < t^*) \\ & \pi(\tau) \notin B, \pi(t^*) \in B, \\ \sum_{\tau=0}^t \text{rew}(\pi(\tau)) & \text{otherwise.} \end{cases} \tag{27}$$

By exploiting Theorem 7, we can now demonstrate the following result proving that the recursive schema above introduced converges to what expected.

Theorem 8 $\bar{E}[\text{Rew}_B | S_0]$ restricted to $S_{=1}^B$ is a fixed point of the iterative scheme (25).

To conclude, let us focus on the imprecise counterpart of reward-bounded hitting probability h_B^r and its upper bound \bar{h}_B^r , which we defined as follows.

Definition 14 (*Upper hitting probability*) Given a IMC \mathcal{M} , a set of states $B \subseteq \mathcal{S}$, a reward-threshold r , and an initial state $s \in \mathcal{S}$:

$$\bar{h}_B^r(s) := \max_{P_s^M \in \mathcal{K}^M} \sum_{\substack{\hat{\pi} \in \text{Paths}_{\text{fin}}(s) : \exists t \in \mathbb{N} \text{ s.t.} \\ \hat{\pi}(t) \in B \wedge \forall \tau < t \hat{\pi}(\tau) \notin B \wedge \\ \text{rew}(\hat{\pi}(0), \dots, \hat{\pi}(t)) \leq r}} P_s^{M_{\text{DTMC}}}(\text{Cyl}(\hat{\pi})).$$

Similarly to the precise case (see Proposition 4), the values of $\bar{h}_B^\rho(s)$ for all $s \in \mathcal{S}$ and $\rho := 0, 1, \dots, r$ provide a solution to the following system of equations:

$$\bar{h}_B^\rho(s) = \begin{cases} 1 & \text{if } s \in B \text{ and } \text{rew}(s) \leq \rho \\ 0 & \text{if } \text{rew}(s) > \rho \text{ or } s \notin \mathcal{S}_{>0} \\ \max_{T(s,S') \in \mathcal{T}(s,S')} \sum_{s' \in \mathcal{S}} T(s,s') \bar{h}_B^{\rho - \text{rew}(s)}(s') & \text{otherwise.} \end{cases} \tag{28}$$

To compute \bar{h}_B^r , we can, therefore, use a recursive schema analogous to that presented in “Markov Reward Models”.

First, we define a matrix $\bar{\mathbf{h}}_B^{t,\rho_0:r}$ whose cells are the values of $\bar{h}_B^{t,\rho}(s)$ computed for each $s \in \mathcal{S}$ and for $\rho = 0, \dots, r$.

For $t = 0$, we generate $\bar{\mathbf{h}}_B^{t,\rho_0:r}$ by computing the vectors $\bar{h}_B^{t=0,\rho}$ for $\rho = 0, \dots, r$ as in Eq. (15).

For increasing values of t , $\bar{\mathbf{h}}_B^{t,\rho_0:r}$ is generated by computing the vectors $\bar{h}_B^{t,\rho}$ for $\rho = 0, \dots, r$ as follows:

$$\bar{h}_B^{t,\rho}(s) = \begin{cases} 1 & \text{if } s \in B \text{ and } \text{rew}(s) \leq \rho \\ 0 & \text{if } \text{rew}(s) > \rho \text{ or } s \notin \mathcal{S}_{>0} \\ \max_{T(s,S') \in \mathcal{T}(s,S')} \sum_{s' \in \mathcal{S}} T(s,s') \bar{h}_B^{t-1,\rho - \text{rew}(s)}(s') & \text{otherwise.} \end{cases} \tag{29}$$

As in the precise case, the values of $\bar{h}_B^{t-1,\rho - \text{rew}(s)}(s')$ for all $s' \in \mathcal{S}$ are provided by the matrix $\bar{\mathbf{h}}_B^{t-1,\rho_0:r}$ that we generate at time-step $t - 1$. The recursion is based on iterated applications of the upper transition operator \bar{T} , each one based on solving $|\mathcal{S}|$ linear programming tasks.

To compute $\bar{h}_B^r(s)$ for all $s \in \mathcal{S}$, we proceed as in the precise case, that is, we iterate $\bar{\mathbf{h}}_B^{t,\rho_0:r}$ over increasing values of t until convergence. Hence, for each $s \in \mathcal{S}$, $\bar{h}_B^r(s)$ is given by the s -cell of the r -vector of the matrix $\bar{\mathbf{h}}_B^{t^*,\rho_0:r}$, where t^* is the convergence time-step. The convergence of the procedure is granted by the following theorem:

Theorem 9 *Let $(\mathcal{M}, \text{rew})$ be a IMRM, $B \subseteq [\mathcal{S}]$, and $\rho \in \mathbb{N}$, there is a $t^* \in \mathbb{N}$ such that for all $\tau \geq 0$:*

$$\bar{\mathbf{h}}_B^{t^* + \tau, \rho_0:r} = \bar{\mathbf{h}}_B^{t^*, \rho_0:r}. \tag{30}$$

The overall time complexity of the procedure is polynomial in $|\mathcal{S}|t^*$ for a reasoning analogous to that stated for Eq. (24).

Imprecise Probabilistic Interpreted Systems

The second IMC extension we consider are *imprecise probabilistic interpreted systems* (IPISs) [28]. The latter are defined as multi-agent systems composed by agents whose stochastic behaviour is described in terms of IMCs. An IPIS under this interpretation is constructed as follows. For each agent $a \in \mathcal{A}$, let $\{K^a(S' | s)\}_{s \in \mathcal{S}}$ denote a family of CSs including, for each

$s \in \mathcal{S}$, all the transition probability mass functions $P^a(S' | s)$ that are compatible with some agent’s probabilistic beliefs. To obtain an IPIS, we replace all the transition matrices T^a , $a \in \mathcal{A}$ in a standard PIS with corresponding (row-stochastic) imprecise transition matrices $\mathcal{T}^a := \{K^a(S' | s)\}_{s \in \mathcal{S}}$, whose rows correspond to the transition CSs $K^a(S' | s)$ for all $s \in \mathcal{S}$. The overall stochastic behaviour of the entire multi-agent system is then described by a global *imprecise* transition matrix $\mathcal{T}_{\text{IPIS}}$ which can be obtained following different approaches. The

most conservative approach consists of defining $\mathcal{T}_{\text{IPIS}}$ as a collection of $|\mathcal{S}|$ conditional CSs $K_{\text{IPIS}}(S' | s)$, each one being defined as $\bigcup_{a \in \mathcal{A}} K^a(S' | s)$. While natural, this approach always implies an increase of the degree of imprecision, defined in terms of the size of the credal sets.

Another approach to obtain $\mathcal{T}_{\text{IPIS}}$ consists of computing, for each transition $s, s' \in \mathcal{S} \times \mathcal{S}$, the credal version of the *logarithmic pooling* of the family of conditional CSs $\{K^a(S' | s) : a \in \mathcal{A}\}$. In general, this is defined as the element-wise application of the standard logarithmic pooling to the elements of the credal sets. This element-wise approach, however, might comport exponential complexity with respect to the number of agents in the model. A similar problem also occurs when considering alternative strategies, such as the one proposed in [39] within the framework of general credal networks. An efficient way to overcome the problem we adopt here consists in considering an outer approximation of the lower and upper bounds of the credal logarithmic pooling achieved as follows:

$$\bar{\mathcal{T}}_{\text{IPIS}}(s, s') := \frac{\prod_{a \in \mathcal{A}} \bar{T}^a(s, s')}{\prod_{a \in \mathcal{A}} \bar{T}^a(s, s') + \sum_{s'' \neq s'} \prod_{a \in \mathcal{A}} \underline{T}^a(s, s'')} \tag{31}$$

The lower bound is analogously computed. The obtained global matrix $\mathcal{T}_{\text{IPIS}}$ consists of an imprecise transition matrix $\mathcal{T}_{\text{IPIS}}$ whose entries are intervals $(m, n) \subseteq [0, 1]$ whose extremes are given by the lower $\underline{\mathcal{T}}_{\text{IPIS}}(s, s')$ and the upper $\overline{\mathcal{T}}_{\text{IPIS}}(s, s')$ bounds of the transition probabilities. As in the precise case, the global matrix describes the *embedded* IMC of the IPIS that is used to compute probabilistic inferences arising with the overall stochastic behaviour of the multi-agent system.

Imprecise Probabilistic Interpreted Reward Systems

The last class of IMC-based structures we consider are *imprecise-probabilistic interpreted reward systems* (IPIRSs). The latter are defined as pairs $\langle \mathcal{M}_{\text{IPIS}}, \text{rew} \rangle$ of a IPIS $\mathcal{M}_{\text{IPIS}}$ and a reward function $\text{rew} : \mathcal{S} \mapsto \mathbb{N}$. As for IPISs, the global transition matrix $\mathcal{T}_{\text{IPIRS}}$ of an IPIRS is obtained by combining the credal transition matrices \mathcal{T}^a of the various agents $i \in \mathcal{A}$.

Notice that all the various imprecise Markov models above introduced can be seen as components of an IPIRS. Specifically, an IPIRS $(\mathcal{M}_{\text{IPIS}}, \text{rew})$ includes an IPIS $\mathcal{M}_{\text{IPIS}}$ that is composed by several IMCs \mathcal{M} , one for each agent of the system. On the other hand, an IPIRS $(\mathcal{M}_{\text{IPIS}}, \text{rew})$ includes several IMRMs, one for each agent in the system, composed by an IMC \mathcal{M} and the reward function rew . This work focuses on only IPIRS and their properties. Nevertheless, the various results obtained for IPIRSs can be easily transferred to IMCs, IMRMs, and IPISs.

Epistemic Imprecise PRCTL

This section presents *Epistemic Imprecise Probabilistic Reward Computation Tree Logic* (EIPRCTL), an epistemic and imprecise-probabilistic extension of the PCTL introduced in [6] suitable to specify epistemic, probabilistic, and reward properties of non-stationary (or not fully specifiable) stochastic multi-agent systems. The language is targeted on IPIRSs but can be also applied to the other kinds of imprecise Markov models previously introduced. Notably, properties of IMCs, IMRMs, and IPISs can be specified via specific languages, such as IPCTL [22], IPRCTL [25], and EIPCTL [28], which can be obtained as EIPRCTL fragments.

EIPRCTL Syntax

The EIPRCTL syntax is recursively defined as follows:

$$\phi := \left\{ \top \mid p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid \underline{P}_{\nabla b}\psi \mid \overline{P}_{\nabla b}\psi \mid P_J\psi \mid \underline{E}_{\nabla r}\phi \mid \overline{E}_{\nabla r}\phi \mid E_J\phi \mid K^a\phi \mid E^\Gamma\phi \mid C^\Gamma\phi \mid D^\Gamma\phi \right\},$$

$$\psi := \bigcirc\phi \mid \phi_1 \cup \phi_2 \mid \phi_1 \cup^{\leq t} \phi_2 \mid \phi_1 \cup_{\leq r} \phi_2,$$

$$\epsilon := B_{\nabla b}^a\phi \mid B_{\nabla b}^a\phi,$$

where $p \in AP$, $b \in [0, 1]$, $J \subseteq [0, 1]$, $a \in \mathcal{A}$, $\Gamma \subseteq \mathcal{A}$, and ∇ is a short notation for $\{<, \leq, =, \geq, >\}$.

The language is composed by ϕ , ψ and ϵ -formulae. The former extend classical propositional logic with usual operators for single-agent knowledge K^a , *common knowledge* C^Γ , and *distributed knowledge* D^Γ , and with the following probabilistic modalities:

- $\underline{P}_{\nabla b}\psi$: the lower probability of reaching a path that satisfies ψ is ∇b ;
- $\overline{P}_{\nabla b}\psi$: the upper probability of reaching a path that satisfies ψ is ∇b ;
- $P_J\psi$: the probability of reaching a path that satisfies ψ belongs to the closed interval $J \subseteq [0, 1]$;
- $\underline{E}_{\nabla r}\phi$: the lower bound of the expected cumulative reward earned by the system until reaching a state that satisfies ϕ is ∇r ;
- $\overline{E}_{\nabla r}\phi$: the upper bound of the expected cumulative reward earned by the system until reaching a state that satisfies ϕ is ∇r ;
- $E_J\phi$: the expected cumulative reward earned by the system until reaching a state that satisfies ϕ belongs to the closed interval $J \subseteq [0, 1]$;

The ψ -formulae are standard CTL path-formulae [2, p. 313] used to represent properties of paths:

- $\bigcirc\phi$: in the next state of the path ϕ holds;
- $\phi_1 \cup \phi_2$: ϕ_1 holds along the path until ϕ_2 holds;
- $\phi_1 \cup^{\leq t} \phi_2$: there exists a time-step $\tau \leq t$ such that ϕ_2 holds in the τ -step of the path and ϕ_1 holds in all the previous time-steps;
- $\phi_1 \cup_{\leq r} \phi_2$: ϕ_1 holds in all states of the path until a cumulative reward lower then or equals to r is earned then ϕ_2 holds.

Finally, ϵ -formulae include the two following *weighted-belief* modalities:⁶

- $B_{\nabla b}^a\phi$: agent a believes that the lower probability of hitting ϕ eventually in the future is ∇b ;
- $\overline{B}_{\nabla b}^a\phi$: agent a believes that the upper probability of hitting ϕ eventually in the future is ∇b .

EIPRCTL Semantics

Let us introduce a proper semantics for EIPRCTL formulae based on IPIRSs. This can be seen as a generalisation of the

⁶ In the rest of this work, we refer to such doxastic formulae as *imprecise-probabilistic beliefs*.

semantics proposed in [22] for IMCs and those proposed in [18] for standard (precise) PISs.

Semantics of State-Formulae

We start by presenting satisfiability conditions for Boolean, probabilistic, and epistemic ϕ -formulae separately.

Definition 15 (*Semantics of Boolean formulae*) Given an IPIRS $\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle$ and a state $s \in \mathcal{S}$, the following conditions hold:

$$\begin{aligned} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models p & \quad \text{iff } p \in I(s), \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \neg \phi & \quad \text{iff } \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \not\models \phi, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \phi_1 \wedge \phi_2 & \quad \text{iff } \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \phi_1 \\ & \quad \text{and } \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \phi_2. \end{aligned}$$

Definition 16 (*Semantics of probabilistic formulae*) Given an IPIRS $\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle$ and a state $s \in \mathcal{S}$, the following conditions hold:

$$\begin{aligned} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models P_{\geq b} \psi & \quad \text{iff } P_{\text{IPIRS}}(s \models \psi) \geq b, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models P_{\leq b} \psi & \quad \text{iff } P_{\text{IPIRS}}(s \models \psi) \leq b, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models P_J \psi & \quad \text{iff } \begin{cases} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models P_{\geq \min J} \psi \text{ and} \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models P_{\leq \max J} \psi. \end{cases} \end{aligned}$$

$$\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi \models \bigcirc \phi \quad \text{iff } \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(1) \models \phi,$$

$$\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi \models \phi_1 \bigcup^{\leq t} \phi_2 \quad \text{iff } \exists \tau \leq t : \begin{cases} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(\tau) \models \phi_2 \text{ and} \\ \forall \tau' < \tau : \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(\tau') \models \phi_1, \end{cases}$$

$$\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi \models \phi_1 \bigcup \phi_2 \quad \text{iff } \exists t \geq 0 : \begin{cases} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(t) \models \phi_2 \text{ and} \\ \forall \tau : 0 \leq \tau < t : \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(\tau) \models \phi_1, \end{cases}$$

$$\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi \models \phi_1 \bigcup_{\leq r} \phi_2 \quad \text{iff } \exists t \in \mathbb{N} : \begin{cases} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(t) \models \phi_2 \text{ and} \\ \forall \tau < t : \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, \pi(\tau) \models \phi_1 \\ \text{and } \text{Rew}(\pi, t) \leq r. \end{cases}$$

For the probabilistic formulae, the satisfiability conditions refer to the lower and upper bounds of $P_{\text{IPIRS}}(s \models \psi)$, where $P_{\text{IPIRS}}(s \models \psi)$ is the probability that a path $\pi \models \psi$ belongs to the set of paths originating in s conditional to $S_0 = s$.⁷ Similarly to standard PCTL [2], the computation of the lower and upper bounds of $P_{\text{IPIRS}}(s \models \psi)$ varies depending on ψ .

⁷ Here the subscript *IPIRS* denotes the fact that this probability is computed through the global transition matrix T_{IPIRS} describing the stochastic behaviour of the whole multi-agent system.

We further analyse this point in Section “[Model Checking](#)” dedicated to model checking procedures.

Definition 17 (*Semantics of expected reward formulae*) Let $\text{Sat}(\phi)$ be the set of all states that satisfy ϕ . Given an IPIRS $\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle$ and state $s \in \mathcal{S}$, the following condition holds:

$$\begin{aligned} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \overline{E}_{\nabla r} \phi & \quad \text{iff } \overline{\text{ExpRew}}_{\text{Sat}(\phi)}(s) \nabla r, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \underline{E}_{\nabla r} \phi & \quad \text{iff } \underline{\text{ExpRew}}_{\text{Sat}(\phi)}(s) \nabla r, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models E_J \phi & \quad \text{iff } \begin{cases} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \underline{E}_{\geq \min J} \phi \text{ and} \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \underline{E}_{\leq \max J} \phi. \end{cases} \end{aligned}$$

Definition 18 (*Semantics of epistemic formulae*) Given an IPIRS $\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle$, an agent $i \in \mathcal{A}$ or a group of agents $\Gamma \subseteq \mathcal{A}$, and a state $s \in \mathcal{S}$, the following conditions hold:

$$\begin{aligned} \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models K^a \phi & \quad \text{iff } \forall s', s \sim^a s' : s' \models \phi, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models E^\Gamma \phi & \quad \text{iff } \forall s', s \sim^\Gamma s' : s' \models \phi, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models C^\Gamma \phi & \quad \text{iff } \forall s', s \sim_C^\Gamma s' : s' \models \phi, \\ \langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models D^\Gamma \phi & \quad \text{iff } \forall s', s \sim_D^\Gamma s' : s' \models \phi. \end{aligned}$$

Semantics of Path-Formulae

Definition 19 (*Semantics of ψ -formulae*) Given an IPIRS $\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle$ and a path π , the following conditions hold:

Semantics of Weighted-Belief Formulae

The ϵ -formulae are doxastic formulae that specify the probabilistic beliefs of a specific agent concerning the overall behaviour of the multi-agent system. They are expressed in terms of the lower and upper bounds of $P^a(s \models \top \bigcup \phi)$, that is, the probability according to the agent a that the multi-agent system eventually visit a state $s' \models \phi$ starting from $s \in \mathcal{S}$. This probability is computed analogously to $P_{\text{IPIRS}}(s \models \pi \models \top \bigcup \phi)$ in Definition 16, but replacing the global transition matrix T_{IPIRS} with a local transition matrix T^a that describes the specific stochastic behaviour of the

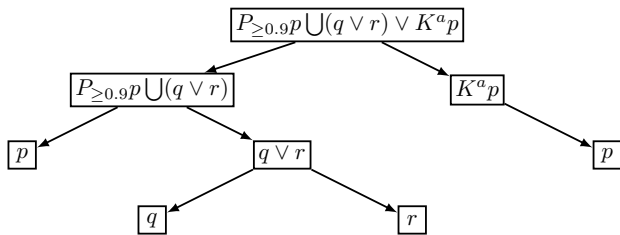


Fig. 1 Parse-tree of a formula

single-agent $a \in \mathcal{A}$. Obviously, since we are considering imprecise models, we are interested in computing the lower and upper bounds of $P^a(s \models \top \cup \phi)$. These are denoted, respectively, by $\underline{P}^a(s \models \top \cup \phi)$ and $\overline{P}^a(s \models \top \cup \phi)$. The specific procedure to compute those bounds is further detailed in Section “Model Checking”. Here we focus only on the satisfiability conditions for ϵ formulae, defined as follows:

Definition 20 (Satisfiability of ϵ -formulae) Given an IPIRS $(\mathcal{M}_{\text{IPIRS}}, \text{rew})$, and a state $s \in \mathcal{S}$,

$$\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models B_{\nabla b}^a \phi \text{ iff } \forall s' : s \sim^a s', \underline{P}^a(s \models \top \cup \phi) \nabla b,$$

$$\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \overline{B}_{\nabla b}^a \phi \text{ iff } \forall s' : s \sim^a s', \overline{P}^a(s \models \top \cup \phi) \nabla b.$$

Model Checking

In the present section, we explain how to check systems modelled by IPIRSs against properties specified in the EIPRCTL language. The procedure we present is obtained by extending the well-known *parsing-tree* algorithm originally introduced for model checking with CTL formulae [2]. We start by briefly recalling the structure and functioning of the parsing-tree. Then, we extend the algorithm introducing a series of new sub-routines to solve specific model checking tasks related to the different kinds of EIPRCTL formulae.

Parsing Tree

Let Λ be a short notation for either a ϕ -formula or an ϵ -formula. Given an IPIRS $(\mathcal{M}_{\text{IPIRS}}, \text{rew})$, a state $s \in \mathcal{S}$, and a formula Λ , a model checking algorithm is an automatic procedure to verify whether $\langle \mathcal{M}_{\text{IPIRS}}, \text{rew} \rangle, s \models \Lambda$ holds. The standard algorithm for CTL and its extension exploits the *parse tree* of Λ generated by decomposing Λ in its various sub-formulae as in Fig. 1 [2, p. 336]. The algorithm works as follows:

1. Generate the *parse tree* of Λ , recursively decomposing Λ in its sub-formulae λ until only atoms remain.

2. Traverse the parse tree of Λ visiting all the sub-formulae λ , starting from the leaves and working backwards to the roots,
3. At each sub-formula λ , calculate the set of states that satisfy λ , denoted $Sat(\lambda)$, by checking whether $s \models \lambda$ for all $s \in \mathcal{S}$,
4. Calculate $Sat(\Lambda)$ by composition of the various $Sat(\lambda)$,
5. Check whether $s \in Sat(\Lambda)$.

The algorithm includes a specific sub-routine to compute $Sat(\Lambda)$ for each specific kind of sub-formula λ .

Boolean Formulae

For Boolean formulae, $Sat(\lambda)$ is computed as follows:

$$Sat(\top) := \mathcal{S},$$

$$Sat(p) := \{s \in \mathcal{S} : p \in l(s)\},$$

$$Sat(\phi_1 \wedge \phi_2) := Sat(\phi_1) \cap Sat(\phi_2),$$

$$Sat(\neg \phi) := \mathcal{S} \setminus Sat(\phi).$$

Probabilistic Formulae

For formulae of the kind $\underline{P}_{\nabla b} \psi$, $\overline{P}_{\nabla b} \psi$ and $P_J \psi$, the set $Sat(\lambda)$ is obtained by computing the lower and upper bounds of $P_{\text{IPIRS}}(s \models \psi)$ for each $s \in \mathcal{S}$ and then check whether they satisfy the specified threshold ∇b . The specific procedure to compute the lower and upper bounds of $P_{\text{IPIRS}}(s \models \psi)$ varies depending on the specification of ψ .

Next If $\psi := \bigcirc \phi$, then $P_{\text{IPIRS}}(s \models \psi)$ corresponds to $P_s(s_1 \in Sat(\phi))$ and its upper (lower) bound can be computed as in Eq. (22).

Time-Bounded Until If $\psi = \phi_1 \bigcup^{\leq t} \phi_2$, then $P_{\text{IPIRS}}(s \models \psi)$ corresponds to the probability of hitting $Sat(\phi_2)$ within a finite number of time-steps t conditional on $S_0 = s$ and with the additional condition that all the states visited before reaching $Sat(\phi_2)$ are in $Sat(\phi_1)$. For each $s \in \mathcal{S}$, we denote such probability by $h_{Sat(\phi_2)|Sat(\phi_1)}^t(s)$. A recursive schema analogous to that in Eq. (24) can be formulated to compute $\overline{h}_{Sat(\phi_2)|Sat(\phi_1)}^t$. Let $\mathbb{1}_{Sat(\phi_1) \setminus Sat(\phi_2)}$ denote the indicator vector whose values are one for all $s \in Sat(\phi_1) \setminus Sat(\phi_2)$ and 0 otherwise. A slightly modified version of the algorithm in (24) for computing $\overline{h}_{Sat(\phi_2)|Sat(\phi_1)}^t$ by recursion over increasing values of t is achieved as follows:

$$\overline{h}_{Sat(\phi_2)|Sat(\phi_1)}^t := \mathbb{1}_{Sat(\phi_2)} + \mathbb{1}_{Sat(\phi_1) \setminus Sat(\phi_2)} \left(\overline{T}_{\text{IPIRS}} \overline{h}_{Sat(\phi_2)|Sat(\phi_1)}^{t-1} \right). \tag{32}$$

As in Eq. (24), the initialisation is given by the indicator function of $Sat(\phi_2)$ while the recursive steps consist of

Fig. 2 Model-checking algorithm for epistemic formulae

Algorithm 4: $Sat(\kappa\phi)$

```

Input: A IPIRS  $\mathcal{M}_{IPIRS}$  and a formula  $\kappa\phi$ 
Output:  $Sat(\kappa\phi)$ 
1  $Sat(\kappa\phi) \leftarrow \{\}$ 
2 Compute  $Sat(\phi) := \{s \in \mathcal{S} \mid \mathcal{M}_{IPIRS}, s \models \phi\}$ 
3 foreach  $eq^{\sim\kappa} \in Eq^{\sim\kappa}$  do
4   if  $eq^{\sim\kappa} \subseteq Sat(\phi)$  then
5      $Sat(\kappa\phi) \leftarrow \{eq^{\sim\kappa}\}$ 
6   end
7 end
8 return  $Sat(\kappa\phi)$ 
    
```

iterated applications of the upper transition operator to the hitting vector computed at the precedent time-step $t - 1$. The only relevant difference with the analogous scheme presented in Section “Imprecise Markov Chains” consists of the indicator vector $\mathbb{1}_{Sat(\phi_1) \setminus Sat(\phi_2)}$ that replaces the indicator vector $\mathbb{1}_{B^c}$ of the complement of the hitting event B . In the general scheme, $\mathbb{1}_{B^c}$ limits the iteration considering only paths that have not already visited an $s \in B$. Here, by using $\mathbb{1}_{Sat(\phi_1) \setminus Sat(\phi_2)}$, we limit the iteration to only those paths whose actual and previous states are all in $Sat(\phi_1)$ and that have not already reached a state $s \in Sat(\phi_2)$. Notice that

$$\bar{h}_{Sat(\phi_2)|Sat(\phi_1)}^{-t,\rho}(s) = \begin{cases} 1 & \text{if } s \in Sat(\phi_2) \text{ and } rew(s) \leq \rho, \\ 0 & \text{if } h(s) = 0 \text{ or } s \notin Sat(\phi_1) \setminus Sat(\phi_2) \text{ or } rew(s) > \rho, \\ \bar{\mathcal{T}}h_B^{t-1,\rho-rew(s)}(s) & \text{otherwise.} \end{cases} \tag{33}$$

Eq. (32) is the imprecise analogous of the system of linear equations used to compute in $P(s \models \phi_1 \cup \phi_2)$ in the precise case, as reported in [2, Sec. 10.2.1]. Finally, as for Eq. (24), the computation of Eq. (32) is based on solving $|\mathcal{S}|t$ linear programming tasks and its time complexity is, thus, polynomial in $|\mathcal{S}|t$.

Until If $\psi := \phi_1 \cup \phi_2$, then $P_{IPIRS}(s \models \psi)$ corresponds to the probability of hitting $Sat(\phi_2)$ conditional on $S_0 = s$ and with the additional requirement that all the states visited before reaching $Sat(\phi_2)$ are in $Sat(\phi_1)$. To compute the lower and upper bounds of this probability, we simply iterate the schema in Equation (32) over increasing values of t until convergence.

Reward-Bounded Until If $\psi := \phi_1 \cup_{\leq r} \phi_2$, then $P_{IPIRS}(s \models \phi_1 \cup_{\leq r} \phi_2)$ corresponds to the reward-bounded hitting probability of $Sat(\phi_2)$ with the additional condition that all the states visited before reaching $Sat(\phi_2)$ are in $Sat(\phi_1)$. We denote this probability by $h_{Sat(\phi_2)|Sat(\phi_1)}^r$. To compute the upper (lower) bound of the latter, we involve a

slightly modified version of the procedure introduced in Eqs. (28) and (29)

Let $\bar{h}_{Sat(\phi_2)|Sat(\phi_1)}^{-t,\rho_0:r}$ be a matrix whose cells are the values of $\bar{h}_{Sat(\phi_2)|Sat(\phi_1)}^{-t,\rho}(s)$ computed for each $s \in \mathcal{S}$ and for $\rho : 0, 1, \dots, r$.

For $t = 0$, we generate $\bar{h}_{Sat(\phi_2)|Sat(\phi_1)}^{-t,\rho_0:r}$ by computing the vectors $\bar{h}_{Sat(\phi_2)|Sat(\phi_1)}^{-t=0,\rho}$ for $\rho = 0, 1, \dots, r$, as in Eq. (15).

For $t > 0$, we generate $\bar{h}_{Sat(\phi_2)|Sat(\phi_1)}^{-t,\rho_0:r}$ by computing the various vectors $\bar{h}_B^{-t,\rho}$ for $\rho = 0, 1, \dots, r$ as follows:

The schema is analogous to that in Eq. (29). The only relevant difference is the additional clause prescribing that $\bar{h}_B^{-t,\rho}(s) = 0$ also for all $s \in Sat(\phi_1) \setminus Sat(\phi_2)$, whereas in Eq. (29) $\bar{h}_B^{-t,\rho}(s) = 0$ only for $s \in \mathcal{S}$ such that either $h_B(s') = 0$ or $rew(s) > r$. The additional clause blocks the recursion for the successors of the initial state that do not belong to $Sat(\phi_1)$. Indeed, if a certain successor $s' \notin Sat(\phi_1) \setminus Sat(\phi_2)$ is reached at a certain time-step τ of the iteration, then $h_{Sat(\phi_2)|Sat(\phi_1)}^r(s')$ takes value zero and the recursion from that state is stopped. In such a way, it is possible to account for the additional requirement that all states visited before reaching the hitting event $Sat(\phi_2)$ are in $Sat(\phi_1)$. Notice that the slightly modification does not alter the general results about \bar{h}_B^r reported above. In particular, the time complexity of the procedure remains polynomial in $|\mathcal{S}|t^*$ (with t^* denoting the convergence time-steps) as it practically depends on the iterative step $\bar{\mathcal{T}}h_B^{t-1,\rho-rew(s)}(s)$, which is the same both in Eq. (29) and (33).

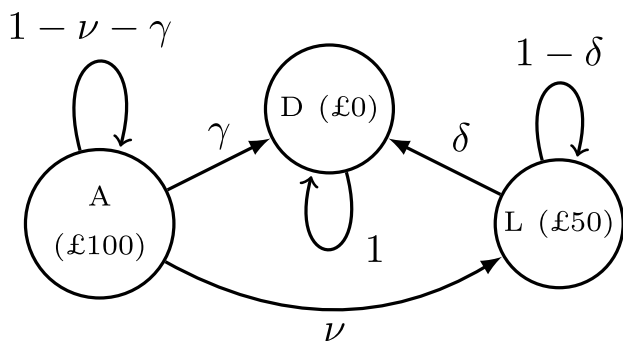


Fig. 3 Transitions in a three-state MRM

Expected Reward Formulae

For formulae of the kind $E_{\nabla_r} \phi \mid \bar{E}_{\nabla_r} \phi \mid E_j \phi$, the procedure to determine $Sat(\lambda)$ is based on computing the lower or the upper bounds of $ExpRew_{Sat(\lambda)}(s)$, with $Sat(\lambda) \subseteq \mathcal{S}$ following the procedure in Eq. (25).

Epistemic Formulae

For formulae of the kind $K^i \phi \mid C^\Gamma \phi \mid D^\Gamma \phi$, the sub-routine to compute $Sat(\lambda)$ is based on the algorithm reported in Fig. 2.

Let κ be a short notation for $K^i \mid E^\Gamma \mid C^\Gamma \mid D^\Gamma$, and let \sim^κ be a short notation for $\sim^i \mid \sim_E^\Gamma \mid \sim_C^\Gamma \mid \sim_D^\Gamma$. Given an epistemic equivalence relation \sim^κ , we denote by Eq^{\sim^κ} the partition induced on \mathcal{S} by \sim^κ . Each element of Eq^{\sim^κ} is an EEC that we denote by eq^{\sim^κ} . The algorithm in Fig. 2, works as follows:

1. It takes in input an IPIRS M_{IPIRS} and a formula $\kappa \phi$;
2. It computes $Sat(\phi)$ by recursively calling the respective sub-routine;
3. For each $eq^{\sim^\kappa} \in Eq^{\sim^\kappa}$, it checks whether $eq^{\sim^\kappa} \subseteq Sat(\phi)$. If this is the case, then the algorithm adds the whole equivalence class eq^{\sim^κ} to $Sat(\kappa \phi)$.

The main advantage of this procedure is that it does not consider single states $s \in \mathcal{S}$ but works directly on EECs. This strategy drastically reduces the time necessary for the execution of the procedure, which ultimately results polynomial in $|\mathcal{S}|n$, with n denoting the nesting-depth of $\kappa \phi$, i.e. the number of nested instances of epistemic operators occurring in the formula. For more details on this procedure, we refer to [40, Sec. 2.1].

Imprecise Weighted-Belief Formulae

For ϵ -formulae, the procedure to compute $Sat(\lambda)$ requires to compute the upper (lower) bounds of $P^a(s \models \top \cup \phi)$ for each $s' \in eq^a(s)$. In practice, these correspond to the lower and upper bounds of $h_{Sat(\phi_2)}^a$, that is, the hitting probability of

Table 1 Conversion and dismissing rates

Rate (%)	Department 1	Department 2	Department 3
γ	1.750	3.540	2.810
ν	0.031	0.187	0.149
δ	0.120	0.130	0.180

event $Sat(\phi_2)$ computed through the local transition matrix T^a of the agent a . The procedure to compute $P^a(s \models \top \cup \phi)$ is totally analogous to that for computing $P_{IPIRS}(s \models \top \cup \phi)$ reported above, with the relevant difference that we use the local transition matrix T^a instead of the global transition matrix T_{IPIRS} .

A Case Study on Healthcare Budgeting

We present a first validation of EIPRCTL based on a slightly modified version of the MRM originally proposed in [41]. MRMs are used in that work to estimate the recovery costs for patients in geriatric departments. Two kinds of recovery are considered: short-term recoveries for acute cares have a daily cost estimated as £100, while long-term recoveries cost £50 per day. From a cumulative perspective, long-term recoveries are more expensive, since those patients typically remain in the hospitals for longer periods.

The evolution across time of health conditions of a patient can be described through a MC with three states: acute care A , long-term care L , and discharge or death D . Transitions from L to A are considered impossible, while D acts as an absorbing state. A parametrized version of the transition matrix for this model is in Fig. 3. The parameters have the following interpretation: the *conversion rate* ν corresponds to the probability of passing from a short-term to a long-term recovery, while the *dismissing rates* γ and δ correspond to the probability of being discharged/die, respectively, in a short and long-term recovery. Rates γ , ν and δ vary depending on the patient and disease. An assessment of these parameters for three different departments is in Table 1.

The reward rew associated with each state represents the daily cost per patient. In a scale where one is assumed to correspond to one pound, the daily costs per patient are described by a function such that $rew(A) = 100$, $rew(L) = 50$, and $rew(D) = 0$. When a patient is dismissed or death she no longer has a cost for the hospital. Following these specifications, it is possible to construct a (precise) MRM $\langle M, rew \rangle$ able to predict the expected cumulative cost incurred by the hospital for each patient up to the time of the patient’s discharge or death. Suppose that the total amount of financial resources per patient available to the hospital is $\rho := \pounds 40,000$. We are interested in verifying whether the

expected cumulative cost per patient is sustainable, i.e. it does not exceed the amount of resources available. This corresponds to check whether:

$$\text{ExpRew}_D(s) \leq \rho,$$

for both $s = A$, i.e. for a patient initially recovered in *acute care*, and $s = L$, i.e. for a patient initially recovered in *long-term care*.

The MRM in [41] presents an important limitation, that is, it considers precise values for the transition rates ν , γ and δ , which is tantamount to assuming that the probability of a patient to change her health-condition is always the same independently from time. This assumption is clearly problematic. For example, it is obvious that the probability to die for patients in long-term cares increases with time. Imprecise probabilistic models allow us to overcome this limitation by considering, instead of precise values, probability intervals within which the transition rates can fluctuate over time.

In practice, we define an imprecise transition matrix for each department obtained by a perturbation of the values of each column of Table 1. As a perturbation method for converting a probability mass function into a credal set, we simply adopt a *linear-vacuous* contamination [21]. The method works as follows. If $P(S)$ is a PMF over S , the CS $K(S)$ (see Section “Imprecise Transition Matrices”) becomes the CS that includes all the PMFs obtained as a mixture $(1 - \alpha)P(S) + \alpha P'(S)$, where $P'(S)$ is any probability mass function and the parameter $\alpha \in [0, 1]$ defines the level of imprecision in the CS (e.g. $K(S) = \{P(S)\}$ for $\alpha = 0$, while for $\alpha = 1$ we get the vacuous CS). In our example, we assume a perturbation level $\alpha = 0.03$. Hence, by applying the above contamination model to each row of the three precise transition matrices described in Table 1, we obtain three imprecise transition matrices \mathcal{T}^a , $a \in \{1, 2, 3\}$ (where a is the number of the department to which the matrix refers).

Given that the costs associated with each state in the model remain the same as above, we have now obtained three different IMRMs, each describing the scenario related to one of the three departments. We are still interested in checking whether, given a patient, the expected cumulative cost incurred until the patient is dismissed or dies does not overcome the available resources. However, we neither know from which one of the three departments the patient comes nor whether it is recovered in acute (A) or long-term (L) care. We can model this scenario as follows.

First, instead of considering a specific transition matrix \mathcal{T}^a , $a \in \{1, 2, 3\}$, we consider an aggregated model where

the global imprecise transition matrix is a pooling of the imprecise transition matrices of each department.

Such a global imprecise transition matrix $\mathcal{T}_{\text{IPIS}}$ can be obtained by the logarithmic pooling as in Eq.(31). As an alternative, more cautious, estimate, we also consider instead a *conservative* pooling consisting in taking as aggregated model the union of the probability intervals of the imprecise transition matrices of the different departments.

Finally, the fact that, independently from which is the department the patient comes from, we cannot know whether the patient is recovered in acute or long-term cares corresponds to assume that $A \sim^a L, \forall a \in \{1, 2, 3\}$, i.e. states labelled by A and L are indistinguishable. We have now obtained a description of the considered scenario in terms of an IPISR $\langle \mathcal{M}_{\text{IPIS}}, \text{rew} \rangle$.

Checking whether the upper maximum expected cumulative cost incurred by the hospital until a patient is dismissed or dies is sustainable corresponds, therefore, to verify whether or not the formula $\overline{E}_{\leq \rho} D$ holds in the model $\langle \mathcal{M}_{\text{IPIS}}, \text{rew} \rangle$ for each state s in the equivalence class $\{A, L\}$. To check this formula, we apply the procedure discussed in Section “Model Checking”. The algorithms described in Section “Imprecise Markov Models” are finally used to compute the upper bounds of $\text{ExpRew}_D(s)$ for both $s = A$ and $s = L$.⁸ The most cautious bounds returned by the conservative pooling are:

$$\overline{\text{ExpRew}}_D(A) = \pounds 29'561, \quad (34)$$

$$\overline{\text{ExpRew}}_D(L) = \pounds 42'343. \quad (35)$$

As expected, the cumulative costs for patients initially admitted in acute care are lower and not exceeding the resources available to the hospital. The same does not happen for patients initially admitted in long-term care.

The management of the hospital might consequently need to check how likely is the fact that the cumulative costs for an hospitalised patient are exceeding the available resources. We do that by checking that the probability of a patient to be dismissed/died before the cumulative cost overcomes the available resources is sufficiently high, e.g. greater than or equal to a threshold $\pi := 0.95$. This corresponds to check whether the formula $\underline{P}_{\geq \pi} \top \bigcup_{\leq \rho} D$ holds in the model $\langle \mathcal{M}_{\text{IPIS}}, \text{rew} \rangle$ and for each state s in the equivalence class $\{A, L\}$. Remarkably, by means of the algorithms described in Section “Imprecise Markov Models” we obtain that the formula is satisfied for both initial states. The resources overrun is, therefore, a relatively unlikely event for the hospital.

⁸ Code available here <https://github.com/IDSIA-papers/2021-ISIPTA-IPRCTL>.

Conclusions

An intrinsic limitation of probabilistic model checking is related to its fundamental reliance on the use of standard Markov models, which can notoriously model only stationary agents whose transition probabilities are all specified by known numerical values. To overcome this limitation, we have presented a novel framework based on the theory of imprecise probabilities and the related imprecise Markov models. More specifically, we have explained how the use of imprecise Markov models allows us to apply probabilistic model checking methods to both non-stationary agents and agents whose transition probabilities are not fully known without compounding computational complexity issues. The key point is that both probabilistic and reward inferences in imprecise Markov models can be computed by iteratively solving linear programming tasks. This allows us to solve relevant model checking tasks without increasing the computational complexity of the relative procedures, which always remains polynomial in the number of states of the models. The paper focuses specifically on stochastic multi-agent systems, but the framework it introduces is useful also for single stochastic agents. The main limitation is that, so far, we considered only discrete-time models. Recent developments [38] in the study of imprecise *continuous-time* Markov chains (CTMC) strongly suggest that an analogous framework can be introduced for continuous-time models, which are of fundamental relevance for applications in fields like computational and systems biology [14, 15]. In the model checking community, some works concerning non-stationarity issues in continuous-time Markov models have been recently proposed. In [42], for example, non-stationary agents are modelled via *uncertain* CTMCs, which are CTMCs whose transition probabilities vary non-deterministically in time within bounded continuous intervals. Although uncertain and imprecise CTMCs are similar, they are not equivalent formalisms. Notably, while an uncertain CTMC can be regarded as the largest family of precise CTMCs compatible with the bounds of the intervals, an imprecise CTMC is the largest family of all, potentially non-Markov, non-homogeneous, processes compatible with given constraints. In other words, imprecise CTMCs are more expressive and potentially useful for a wider range of applications. So far, however, there are no works specifically concerning model checking with imprecise CTMCs.

To conclude, another important development to consider might concern the development of an imprecise-probabilistic framework for Markov decision-processes, notably for the relevance of the latter and their natural connection with the field of Reinforcement Learning [43].

Appendix A

Proof of Theorem 2 The proof is analogous to the one of [24, Lemma 9] given for a similar result. The statement is proven by induction. First, we prove the induction base:

$$(\forall s \in S_{=1}^B) E[Rew_B^0 | S_0 = s] = rew(s) =: ExpRew_B^0(s). \tag{A1}$$

Next, note that for every $t \in \mathbb{N}$, $t \neq 0$, the value of Rew_B^t on any π is completely determined by the values $\pi(0), \dots, \pi(t) \in S$. Thus, we can alternatively interpret Rew_B^t as a function on S^{t+1} . Moreover, for any $t \in \mathbb{N}$, $t \neq 0$ and $s_0, \dots, s_t \in S^{t+1}$, we observe that

$$Rew_B^t([s_i]_{i=0}^t) = rew(s_0) + I_{B^c}(s_0)Rew_B^{t-1}([s_i]_{i=1}^t), \tag{A2}$$

with $Rew_B^0(s_0) = rew(s_0)$. By exploiting this observation, we can proceed with the induction step. Assuming that the statement is true for $t - 1$, with $t \in \mathbb{N}$, $t \neq 0$, we prove that it is also true for t . First, for every $s \in S_{=1}^B$, we have that

$$\begin{aligned} E[Rew_B^t([S_i]_{i=0}^t) | S_0 = s] &= E[rew(s) + I_{B^c}(s)Rew_B^{t-1}([S_i]_{i=1}^t) | S_0 = s] \\ &= rew(s) + I_{B^c}(s)E[Rew_B^{t-1}([S_i]_{i=1}^t) | S_0 = s] \\ &= rew(s) + I_{B^c}(s)E[E[Rew_B^{t-1}([S_i]_{i=1}^t) | [S_i]_{i=0}^1] | S_0 = s], \end{aligned}$$

where in the last step we use the law of iterated expectations (or [24, Prop. 39] and [24, Prop. 4]). Clearly, $E[Rew_B^{t-1}([S_i]_{i=1}^t) | [S_i]_{i=0}^1]$ does not depend on the initial state S_0 , hence

$$\begin{aligned} E[Rew_B^{t-1}([S_i]_{i=1}^t) | [S_i]_{i=0}^1] &= E[Rew_B^{t-1}([S_i]_{i=1}^t) | S_1] \\ &= E[Rew_B^{t-1}([S_i]_{i=0}^{t-1}) | S_0]. \end{aligned}$$

Now, $E[Rew_B^{t-1}([S_i]_{i=0}^{t-1}) | S_0 = s] = ExpRew_B^{t-1}(s)$ for every $s \in S_{=1}^B$ by the inductive hypothesis. Hence, plugging this back into the expression we obtained for $E[Rew_B^t([S_i]_{i=0}^t) | S_0 = s]$, we have

$$\begin{aligned} E[Rew_B^t([S_i]_{i=0}^t) | S_0 = s] &= rew(s) + I_{B^c}(s)E[ExpRew_B^{t-1}(S_1) | S_0 = s] \\ &= rew(s) + I_{B^c}(s)[\hat{T}ExpRew_B^{t-1}](s) =: ExpRew_B^t(s), \end{aligned}$$

for all $s \in S_{=1}^B$, by using the fact that a transition matrix for an homogeneous Markov chain encodes 1 time-step expectation [24, Sec. 2.1]. □

Proof of Theorem 3 First, notice that $\lim_{t \rightarrow +\infty} Rew_B^t = Rew_B$. Hence, for every $s \in \mathcal{S}_{=1}^B$, $\lim_{t \rightarrow +\infty} E[Rew_B^t | S_0 = s] = E[Rew_B | S_0 = s]$ by the monotone convergence theorem since $0 \leq Rew_B^0 \leq Rew_B^1 \leq \dots \leq Rew_B$ (or [24, Prop. 7] and [24, Prop. 4]). From Th. 2, it follows that $ExpRew_B^t$ are also non decreasing and non-negative. Hence, by using as well the continuity of \hat{T} with respect to non decreasing and non negative sequences, we find that

$$\begin{aligned} E[Rew_B | S_0] |_{\mathcal{S}_{=1}^B} &= \lim_{t \rightarrow +\infty} E[Rew_B^t | S_0] |_{\mathcal{S}_{=1}^B} = \lim_{t \rightarrow +\infty} ExpRew_B^t \\ &= \lim_{t \rightarrow +\infty} \left(rew_{\mathcal{S}_{=1}^B} + I_{B^c} \hat{T} ExpRew_B^{t-1} \right) \\ &= rew_{\mathcal{S}_{=1}^B} + I_{B^c} \hat{T} E[Rew_B | S_0] |_{\mathcal{S}_{=1}^B}, \end{aligned}$$

where the third and fifth step follow from Theorem 2, while the fourth step follows from the definition of $ExpRew_B^t$. \square

Proof of Theorem 5 It is actually enough to verify that, given a MRM (M, rew) and $B \subseteq [S]$, for all $s \in \mathcal{S}$ and all $r \in \mathbb{N}$, it holds that

$$\exists t^* \in \mathbb{N} : \forall \tau \geq 0, h_B^{t^* + \tau, r}(s) = h_{B^*}^{t^*, r}(s). \tag{A3}$$

Let $s \in \mathcal{S}$, $B \subseteq \mathcal{S}$ and $r \in \mathbb{N}$. We show that, for all $t \in \mathbb{N}$, the time-bounded hitting probability $h_B^{t, r}(s)$ is equivalent to the time-bounded hitting probability $h_{B^*}^{t, r}(s)$ of an alternative event B^* computed on the alternative MRM (M^*, rew^*) , with $M^* := \langle \mathcal{S}^*, T^* \rangle$. Such structure is construct as follows. Let $r^* := r + \sup_{s' \in \mathcal{S}} rew(s')$. We first inductively define \mathcal{S}^* . Set $X_0 := \{(s, rew(s))\}$, and for every $n > 0$,

$$\begin{aligned} X_{n+1} &:= X_n \cup \{(s'', k) \in \mathcal{S} \times \mathbb{N} \mid \exists (s', m) \in X_n \text{ s.t.} \\ &\quad m \leq r \text{ and } k = rew(s') + m\}. \end{aligned}$$

The idea is that each set $X_{n+1} \setminus X_n$ contains pairs (s', k) where s' is the last node belonging to a path of length $n - 1$ starting in s and whose cumulated reward is $k \leq r^*$.

Notice that the sequence $(X_\ell : \ell \in \mathbb{N})$ is increasing monotone. Moreover for every n , $X_n \subseteq \{(s'', k) \in \mathcal{S} \times \mathbb{N} \mid k \leq r^*\}$, where the latter is a finite set. Hence, there is ℓ such that $X_\ell = X_{\ell+1}$, and define $\mathcal{S}^* := X_\ell \cup \{\perp\}$, where \perp is some new ‘‘sink’’ state. For the remaining functions, we define, for $(s'', m), (s', k) \neq \perp$, $rew^*((s', k)) := rew(s')$ and

$$T^*((s', k), (s'', m)) := \begin{cases} T(s', s'') & \text{if } m = rew(s') + k, \\ 0 & \text{else;} \end{cases}$$

whereas we set $rew^*(\perp) := 0$,

$$T^*(\perp, \circ) := \begin{cases} 1 & \text{if } \circ = \perp, \\ 0 & \text{else;} \end{cases}$$

and $T^*((s', k), \perp) := 1 - \sum_{(s'', m) \neq \perp} T^*((s', k), (s'', m))$. Thus, (M^*, rew^*) is essentially the tree unravelling around s of (M, rew) up to a cumulation of rewards bounded by r^* . That is, every finite path $\hat{\pi}$ in M starting in s such that $rew(\hat{\pi}) \leq r$ corresponds to a finite path $\hat{\pi}^*$ in M^* starting in $(s, rew(s))$ with $P_s(\hat{\pi}) = P_{(s, rew(s))}(\hat{\pi}^*)$ and $rew(\hat{\pi}) = rew^*(\hat{\pi}^*)$. Vice versa, every finite path $\hat{\pi}^*$ in M^* starting in $(s, rew(s))$ such that $rew^*(\hat{\pi}^*) \leq r$ corresponds to a finite path $\hat{\pi}$ in M starting in s with $P_s(\hat{\pi}) = P_{(s, rew(s))}(\hat{\pi}^*)$ and $rew(\hat{\pi}) = rew^*(\hat{\pi}^*)$. Given this observation, let $B^* := \{(s'', k) \in \mathcal{S}^* \mid s'' \in B, k \leq r\}$. We first notice that for every $t \in \mathbb{N}$:

$$h_B^{t, r}(s) = h_{B^*}^{t, r}((s, rew(s))) = h_{B^*}^t((s, rew(s))). \tag{A4}$$

Claim: Let $s' \in \mathcal{S}$. Assume there is a finite path $\hat{\pi}(0) \dots \hat{\pi}(\ell)$ with $\hat{\pi}(0) = s, \hat{\pi}(\ell) = s', k := rew(\hat{\pi}) \leq r^*$, then Eq. (A4) is actually a consequence of the following list of facts:

1. If $h_B(s') = 0$ or $rew(s') > r$ then $h_B^{t, \rho}(s') = h_{B^*}^{t, \rho}((s', k)) = h_{B^*}^t((s', k)) = 0$, for every $\rho \leq r$ and every $t \in \mathbb{N}$,
2. Else, if $k > r$, then $h_B^{t, \rho}((s', k)) = h_{B^*}^{t, \rho}((s', k)) = 0$, for every $\rho, t \in \mathbb{N}$,
3. Else for every $t \in \mathbb{N}$

$$h_B^{t, (r+rew(s')-k)}(s') = h_{B^*}^{t, (r+rew(s')-k)}((s', k)) = h_{B^*}^t((s', k)). \tag{A5}$$

Proof of Claim. The first two items being clear (e.g. the second is simply due to the fact that $(s', k) \notin B^*$ and $h_{B^*}((s', k)) = 0$), we verify the last one. First notice that $k = rew(\hat{\pi}) \leq r$. This means that $s' \in B$ if and only if $(s', k) \in B^*$. We now reason by induction on t . For $t = 0$, we have that all terms are either equals to 1 or 0. However, given that $k \leq r$ and $rew(s') = rew^*((s', k)) \leq r + rew(s') - k$, the identities hold. We now verify the induction step, and therefore, suppose that for $t \in \mathbb{N}$, $h_B^{t, r+rew(s')-m}(s'') = h_{B^*}^{t, r+rew(s')-m}((s'', m)) = h_{B^*}^t((s'', m))$, for every $(s'', m) \in \mathcal{S}^*$ such that $h_B(s'') \neq 0, rew(s'') > r$ and $m \leq r$. Let $\theta := r + rew(s') - k$. Hence $rew(s') \leq \theta$. Also, remember that by hypothesis $h_B(s') > 0$, meaning that

$$h_B^{t+1, \theta}(s') = \begin{cases} 1 & \text{if } s' \in B \\ \sum_{s'' \in \mathcal{S}} T(s', s'') h_B^{t, \theta-rew(s')} (s'') & \text{otherwise.} \end{cases}$$

If $s' \in B$, reasoning as before, Eq. (A4) holds. We thus consider the case when $s' \notin B$, and terms $h_B^{t, \theta-rew(s')} (s'')$. Let $(s'', m) \in \mathcal{S}^*$, with $m - rew(s'') = k$, meaning that

$h_B^{t,\theta-rew(s'')}(s'') = h_B^{t,r-k}(s'') = h_B^{t,r+rew(s'')-m}(s'')$. Notice that by definition $T(s', s'') = T^*((s', k), (s'', m))$. Therefore, if we prove that $h_B^{t,r+rew(s'')-m}(s'') = h_{B^*}^{t,(r+rew(s'')-m)}((s'', m)) = h_{B^*}^t((s'', m))$ we are done. Now, let $\epsilon := r - k \geq 0$, as $k \leq r$. Therefore, $r + rew(s'') - m = r - k = \epsilon$. We reason by cases. If $h_B(s'') = 0$ or $rew(s'') > r \geq \epsilon$, then $h_B^{\epsilon}(s'') = h_{B^*}^{\epsilon}((s'', m)) = h_{B^*}^t((s'', m)) = 0$. Else, assume $m = rew(s'') + k > r$. This means $rew(s'') > r - k = \epsilon$, and therefore, $h_B^{\epsilon}(s'') = h_{B^*}^{\epsilon}((s'', m)) = h_{B^*}^t((s'', m)) = 0$. Else, $m = rew(s'') + k \leq r$, and conclude by induction hypothesis.

We finally end the main proof by applying Eq. (A4) to the fact that, by the convergence of standard hitting probability, there is $t^* \in \mathbb{N}$ such that, for every $\tau \geq 0$, $h_{B^*}^{t^*+\tau}((s, rew(s))) = h_{B^*}^{t^*}((s, rew(s)))$. \square

Proof of Theorem 6 We have $\bar{P}_s(S_t \in B) := \bar{E}[\mathbb{1}_B(s) \mid S_0 = s]$ by definition of upper probability. By definition of upper conditional expectation, we have instead $\bar{E}[\mathbb{1}_B(s_t) \mid S_0 = s] := \bar{T}^t \mathbb{1}_B(s_0)$. Hence, $\bar{P}_s(s_t \in B) = \bar{E}[\mathbb{1}_B(s) \mid S_0 = s] = \bar{T}^t \mathbb{1}_B(s_0)$. \square

Proof of Theorem 7 The proof is analogous to the one of [24, Lemma 9] given for a similar result. Notice that results and lemmas from [24] that we will use in the following are originally formulated for game-theoretic IMCs. As from results proved in [32], for the domain of functions we consider in what follows, the measure-theoretic and the game-theoretic formulations can be considered equivalent. We give the proof for \overline{ExpRew}_B^t , the one for \underline{ExpRew}_B^t will follow an analogous reasoning.

The statement is proven by induction and we first prove the induction base. To this end, note that Rew_B^0 only depends on the state S_0 , so, it holds that

$$(\forall s \in S_{=1}^B) \bar{E}[Rew_B^0 \mid S_0 = s] = rew(s) =: \overline{ExpRew}_B^0(s). \tag{A6}$$

Next, as in the proof of Theorem 2, we interpret Rew_B^t as a function on S^{t+1} . Moreover, for any $t \in \mathbb{N}$, $t \neq 0$ and $s_0, \dots, s_t \in S^{t+1}$, we rewrite $Rew_B^t([s_i]_{i=0}^t)$ as

$$Rew_B^t([s_i]_{i=0}^t) = rew(s_0) + I_{B^c}(s_0)Rew_B^{t-1}([s_i]_{i=1}^t), \tag{A7}$$

with $Rew_B^0(s_0) = rew(s_0)$. Exploiting this observation, we can proceed with the induction step. Assuming that the statement is true for $t - 1$, with $t \in \mathbb{N}$, $t \neq 0$, we prove that it is also true for t . First, for every $s \in S_{=1}^B$, we have that

$$\begin{aligned} &\bar{E}[Rew_B^t([S_i]_{i=0}^t) \mid S_0 = s] \\ &= \bar{E}[rew(s) + I_{B^c}(s)Rew_B^{t-1}([S_i]_{i=1}^t) \mid S_0 = s] \\ &= rew(s) + I_{B^c}(s)\bar{E}[Rew_B^{t-1}([S_i]_{i=1}^t) \mid S_0 = s] \\ &= rew(s) + I_{B^c}(s)\bar{E}[\bar{E}[Rew_B^{t-1}([S_i]_{i=1}^t) \mid [S_i]_{i=0}^1] \mid S_0 = s], \end{aligned}$$

where the last step is based on [24, Prop. 38].

Now, $\bar{E}[Rew_B^{t-1}([S_i]_{i=1}^t) \mid [S_i]_{i=0}^1]$ does not depend on the initial state S_0 , hence

$$\begin{aligned} &\bar{E}[Rew_B^{t-1}([S_i]_{i=1}^t) \mid [S_i]_{i=0}^1] \\ &= \bar{E}[Rew_B^{t-1}([S_i]_{i=1}^t) \mid S_1] = \\ &= \bar{E}[Rew_B^{t-1}([S_i]_{i=0}^{t-1}) \mid S_0]. \end{aligned}$$

Now, $\bar{E}[Rew_B^{t-1}([S_i]_{i=0}^{t-1}) \mid S_0 = s] = \overline{ExpRew}_B^{t-1}(s)$ for every $s \in S_{=1}^B$ by the inductive hypothesis. Plugging this back into the expression we obtained for $\bar{E}[Rew_B^t([S_i]_{i=0}^t) \mid S_0 = s]$, we have

$$\begin{aligned} &\bar{E}[Rew_B^t([S_i]_{i=0}^t) \mid S_0 = s] \\ &= rew(s) + I_{B^c}(s)\bar{E}[\overline{ExpRew}_B^{t-1}(S_1) \mid S_0 = s] \\ &= rew(s) + I_{B^c}(s)[\bar{T} \overline{ExpRew}_B^{t-1}](s) =: \overline{ExpRew}_B^t(s), \end{aligned}$$

for all $s \in S_{=1}^B$, where the second step uses [24, Eq. 6]. \square

Proof of Theorem 8 The proof follows the one of [24, Prop. 10] and considerations below.

As we know from the proof of Theorem 3, $\lim_{t \rightarrow +\infty} Rew_B^t = Rew_B$. Hence, for every $s \in S_{=1}^B$, $\lim_{t \rightarrow +\infty} \bar{E}[Rew_B^t \mid S_0 = s] = \bar{E}[Rew_B \mid S_0 = s]$ by [24, Prop. 7], since $0 \leq Rew_B^0 \leq Rew_B^1 \leq \dots \leq Rew_B$. From Theorem 7 it also follows that \overline{ExpRew}_B^t are also non decreasing and non-negative. Hence, by using as well the continuity of \bar{T} with respect to non decreasing and non-negative sequences, see [24, Lemma 1], we find that

$$\begin{aligned} &\bar{E}[Rew_B \mid S_0] \Big|_{S_{=1}^B} \\ &= \lim_{t \rightarrow +\infty} \bar{E}[Rew_B^t \mid S_0] \Big|_{S_{=1}^B} = \lim_{t \rightarrow +\infty} \overline{ExpRew}_B^t \\ &= \lim_{t \rightarrow +\infty} \left(rew_{S_{=1}^B} + I_{B^c} \bar{T} \overline{ExpRew}_B^{t-1} \right) \\ &= rew_{S_{=1}^B} + I_{B^c} \bar{T} \bar{E}[Rew_B \mid S_0] \Big|_{S_{=1}^B}, \end{aligned}$$

where in the last step we use again Theorem 7. \square

Proof of Theorem 9 The proof is completely analogous to that of Theorem 5. \square

Funding Open access funding provided by SUPSI - University of Applied Sciences and Arts of Southern Switzerland.

Declarations

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Lawler GF. Introduction to stochastic processes. London: Chapman and Hall/CRC; 2018.
2. Baier C, Katoen JP. Principles of model checking. Cambridge: MIT Press; 2008.
3. Huang M, Fu H, Katoen J. Deciding probabilistic simulation between probabilistic pushdown automata and finite-state systems. *Inf Comput*. 2019. <https://doi.org/10.1016/j.ic.2019.05.004>.
4. Hillston J. A compositional approach to performance modelling. PhD thesis, University of Edinburgh, UK, Edinburgh, UK. 1994. <https://hdl.handle.net/1842/15027>. Accessed 01 Mar 2016.
5. Hillston J, Marin A, Rossi S, Piazza C. Contextual lumpability. In: Horváth A, Buchholz P, Cortellessa V, Muscariello L, Squillante MS, editors. 7th International Conference on Performance Evaluation Methodologies and Tools, ValueTools '13, Torino, Italy, December 10–12, 2013; pp. 194–203. ICST/ACM, New York, USA (2013). <https://doi.org/10.4108/icst.valuetools.2013.254408>.
6. Hansson H, Jonsson B. A logic for reasoning about time and reliability. *Formal Aspects Comput*. 1994;6(5):512–35.
7. Bérard B, Bidoit M, Finkel A, Laroussinie F, Petit A, Petrucci L, Schnoebelen P, McKenzie P. Systems and software verification. Cham: Model-Checking Techniques and Tools. Springer; 2001.
8. Bentahar J, Moulin B, Meyer JC. A new model checking approach for verifying agent communication protocols. In: Proceedings of the Canadian Conference on Electrical and Computer Engineering, CCECE 2006, May 7–10, 2006, Ottawa Congress Centre, Ottawa, Canada, 2006; pp. 1586–1590. IEEE, Ottawa. <https://doi.org/10.1109/CCECE.2006.277640>.
9. Gallina L, Hamadou S, Marin A, Rossi S. A probabilistic energy-aware model for mobile ad-hoc networks. In: Al-Begain K, Balsamo S, Fiems D, Marin A, editors. Analytical and stochastic modeling techniques and applications—18th International Conference, ASMTA 2011, Venice, Italy, June 20–22, 2011. Proceedings. Lecture Notes in Computer Science, 2011; vol. 6751, pp. 316–330. Springer, Cham. https://doi.org/10.1007/978-3-642-21713-5_23.
10. Gallina L, Marin A, Rossi S. Connectivity and energy-aware preorders for mobile ad-hoc networks. *Telecommun Syst*. 2016;63(2):307–33. <https://doi.org/10.1007/s11235-015-0122-6>.
11. Rossi, S. Model checking adaptive multilevel service compositions. In: Barbosa LS, Lumpe M, editors. Formal Aspects of Component Software—7th International Workshop, FACS 2010, Guimarães, Portugal, October 14–16, 2010, Revised Selected Papers. Lecture Notes in Computer Science, 2010; vol. 6921, pp. 106–124. Springer, Cham. https://doi.org/10.1007/978-3-642-27269-1_7.
12. Bernardi G, Bugliesi M, Macedonio D, Rossi S. A theory of adaptable contract-based service composition. In: Negru V, Jebelean T, Petcu D, Zaharie D, editors. SYNASC 2008, 10th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, Romania, 26–29 September 2008, pp. 327–334. IEEE Computer Society, New York, USA, 2008. <https://doi.org/10.1109/SYNASC.2008.38>.
13. Bugliesi M, Macedonio D, Pino L, Rossi S. Compliance preorders for web services. In: Laneve C, Su J, editors. Web Services and Formal Methods, 6th International Workshop, WS-FM 2009, Bologna, Italy, September 4–5, 2009, Revised Selected Papers. Lecture Notes in Computer Science, 2009; vol. 6194, pp. 76–91. Springer, Cham. https://doi.org/10.1007/978-3-642-14458-5_5.
14. Brim L, Ceska M, Safránek D. Model checking of biological systems. In: Bernardo M, de Vink EP, Pierro AD, Wiklicky H, editors. Formal methods for dynamical systems—13th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2013, Bertinoro, Italy, June 17–22, 2013. Advanced Lectures. Lecture Notes in Computer Science, 2013; vol. 7938, pp. 63–112. Springer, Cham.
15. Benes N, Brim L, Pastva S, Safránek D. Model checking approach to the analysis of biological systems. In: Liò P, Zuliani P, editors. Automated reasoning for systems biology and medicine. Computational biology, vol. 30. Cham: Springer; 2019. p. 3–35.
16. van Ditmarsch H, van der Hoek W, Halpern J, Kooi B, editors. Handbook of epistemic logic. London: College Publications; 2015.
17. Chen T, Primiero G, Raimondi F, Rungta N. A computationally grounded, weighted doxastic logic. *Stud Log*. 2016;104(4):679–703.
18. Wan W, Bentahar J, Hamza AB. Model checking epistemic-probabilistic logic using probabilistic interpreted systems. *Knowl Based Syst*. 2013;50:279–95. <https://doi.org/10.1016/j.knosys.2013.06.017>.
19. Daws C. Symbolic and parametric model checking of discrete-time Markov chains. In: Zhiming L, Araki K, editors. Theoretical aspects of computing—ICTAC 2004, First International Colloquium, Guiyang, China, September 20–24, 2004, Revised Selected Papers, vol. 3407. Lecture Notes in Computer Science. Cham: Springer; 2004. p. 280–94.
20. Baier C, Hensel C, Hutschenreiter L, Junges S, Katoen JP, Klein J. Parametric Markov chains: PCTL complexity and fraction-free gaussian elimination. *Inf Comput*. 2020;272: 104504. <https://doi.org/10.1016/j.ic.2019.104504>.
21. Walley P. Statistical reasoning with imprecise probabilities. New York: Chapman and Hall; 1991.
22. Troffaes MCM, Skulj D. Model checking for imprecise markov chains. In: Cozman F, Denoeux T, Destercke S, Seidenfeld T, editors. ISIPTA '13 : Proceedings of the Eighth International Symposium on Imprecise Probability : Theories and Applications July 2–5 2013, Compiègne, France., pp. 337–344. Society for Imprecise Probability: Theories and Applications (SIPTA), Compiègne. 2013.
23. Delgado KV, De Barros LN, Dias DB, Sanner S. Real-time dynamic programming for Markov decision processes with imprecise probabilities. *Artif Intell*. 2016;230:192–223.
24. Krak TE, T'Joens N, Bock JD. Hitting times and probabilities for imprecise markov chains. In: Bock JD, de Campos CP, de Cooman G, Quaeghebeur E, Wheeler GR, editors. International

- Symposium on Imprecise Probabilities: Theories and Applications, ISIPTA 2019, 3–6 July 2019, Thagaste, Ghent, Belgium, vol. 103. Proceedings of Machine Learning Research. Ghent: PMLR; 2019. p. 265–75.
25. Termine A, Antonucci A, Facchini A, Primiero G. Robust model checking with imprecise Markov reward models. In: International Symposium on imprecise probabilities: theories and applications, ISIPTA 2021, July 6–9 2021, Granada, Spain. Proceedings of Machine Learning Research. PMLR, Granada; 2021.
 26. De Cooman G, De Bock J, Lopatzidis S. Imprecise stochastic processes in discrete time: global models, imprecise Markov chains and ergodic theorems. *Int J Approx Reason.* 2016;76:18–46.
 27. T'Joens N, Krak TE, De Bock J, De Cooman G. A recursive algorithm for computing inferences in imprecise markov chains. In: Kern-Isberner G, Ognjanovic Z, editors. Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 15th European Conference, ECSQARU 2019, Belgrade, Serbia, September 18–20, 2019, Proceedings, vol. 11726. Lecture Notes in Computer Science. Cham: Springer; 2019. p. 455–65.
 28. Termine A, Antonucci A, Primiero G, Facchini A. Logic and model checking by imprecise probabilistic interpreted systems. In: European Conference on multi-agent systems, 2021; pp. 211–227. Springer.
 29. Norris JR. Markov chains, vol. 2. Cambridge University Press; 1998.
 30. Mauá DD, de Campos CP, Benavoli A, Antonucci A. Probabilistic inference in Credal networks: New complexity results. *J Artif Intell Res.* 2014;50:603–37.
 31. Shafer G, Vovk V. Probability and finance: it's only a game!, vol. 491. London: Wiley; 2005.
 32. T'Joens N, De Bock J. Global upper expectations for discrete-time stochastic processes: in practice, they are all the same! In: International Symposium on imprecise probability: theories and applications, 2021; p. 310–319. PMLR.
 33. De Cooman G, Hermans F, Quaeghebeur E. Imprecise Markov chains and their limit behavior. *Probab Eng Inf Sci.* 2009;23(4):597–635.
 34. De Cooman G, Hermans F. Imprecise probability trees: bridging two theories of imprecise probability. *Artif Intell.* 2008;172(11):1400–27.
 35. Destercke S, Cooman Gd. Relating epistemic irrelevance to event trees. In: Soft methods for handling variability and imprecision. Cham: Springer, Cham; 2008, , p. 66–73.
 36. Hermans F, Škulj D. Stochastic processes. In: Augustin T., Coolen FPA, De Cooman G, and Troffaes MCM, editors. Introduction to imprecise probabilities. London: Wiley, 2014; p. 258–78.
 37. Miranda E, De Cooman G. Structural judgements. In: Augustin T., Coolen F.P.A., De Cooman G., and Troffaes M.C.M. Introduction to imprecise probabilities. London: Wiley, Introduction to imprecise probabilities, 2014; p. 56–78.
 38. Krak T, De Bock J, Siebes A. Imprecise continuous-time Markov chains. *Int J Approx Reason.* 2017;88:452–528.
 39. Antonucci A, Huber D, Zaffalon M, Luginbühl P, Chapman I, Ladouceur R. Credo: a military decision-support system based on Credal networks. In: Proceedings of the 16th Conference on Information Fusion (FUSION 2013), Istanbul, Turkey; 2013.
 40. Lomuscio A, Raimondi F. MCMAS: a model checker for multi-agent systems. In: Hermanns H, Palsberg J, editors. Tools and algorithms for the construction and analysis of systems, 12th International Conference, TACAS 2006 Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2006, Vienna, Austria, March 25–April 2, 2006, Proceedings, vol. 3920. Lecture Notes in Computer Science. Cham: Springer; 2006. p. 450–4.
 41. McClean SI, McAlea B, Millard PH. Using a Markov reward model to estimate spend-down costs for a geriatric department. *J Oper Res Soc.* 1998;49(10):1021–5.
 42. Cardelli L, Grosu R, Larsen KG, Tribastone M, Tschaikowski M, Vandin A. Lumpability for uncertain continuous-time markov chains. In: International Conference on quantitative evaluation of systems, Cham, 2021; p. 391–409. Springer.
 43. Sutton RS, Barto AG. Reinforcement learning—an introduction. Adaptive computation and machine learning. Cambridge: MIT Press; 1998.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.