Model Checking of Open Interval Markov Chains

Souymodip Chakraborty(✉) and Joost-Pieter Katoen

RWTH Aachen University, 52056 Aachen, Germany
souymodip@gmail.com

Abstract. We consider the model checking problem for interval Markov chains with open intervals. Interval Markov chains are generalizations of discrete time Markov chains where the transition probabilities are intervals, instead of constant values. We focus on the case where the intervals are open. At first sight, open intervals present technical challenges, as optimal (min, max) value for reachability may not exist. We show that, as far as model checking (and reachability) is concerned, open intervals does not cause any problem, and with minor modification existing algorithms can be used for model checking interval Markov chains against PCTL formulas.

1 Introduction

Discrete time Markov chains (DTMCs) are useful models for analyzing the reliability and performance of computer systems. A DTMC is defined as a weighted directed graph where the weights on the outgoing transitions define a probability distribution. In general, the precise values of these probabilities may not be always available [9,11,12]. This is precisely the case when transition probabilities are obtained by statistical methods.

Interval Markov chains [9,13] are useful in modeling and verifying probabilistic systems where the value of the transition probabilities are not known precisely. IMCs generalize discrete time Markov chains by allowing intervals of possible probabilities on the state transitions in order to capture the system uncertainty more faithfully. For example, instead of specifying that the probability of moving from state s to t is 0.5, one can specify an interval $[0.3, 0.7]$ which captures the uncertainty in the probability of moving from state s to t. Uncertainty in the model may occur due to various reasons [12]. In some cases, the transition probabilities may depend on an unknown environment, and are approximately known, in other cases the interval may be introduced to make the model more robust.

There are two prevalent semantics of interval Markov chains. *Uncertain Markov Chains* (UMC) [9,11] is an interpretation of interval Markov chains as set of (possibly uncountably many) discrete time Markov chains where each element of the set is a DTMC whose transition probabilities lie within the interval range defined by the IMC. In the other semantics, called *Interval Markov Decision Processes* (IMDP) [11], the uncertainty of the transition probabilities are resolved non-deterministically. It requires the notion of *scheduler*, which chooses

M. Gribaudo et al. (Eds.): ASMTA 2015, LNCS 9081, pp. 30–42, 2015.
DOI: 10.1007/978-3-319-18579-8_3

a distribution, each time a state is visited in an execution, from a (possibly uncountable) set of distributions defined by the intervals on the transitions.

The logic *probabilistic computation tree logic* (PCTL) [8], extends the temporal logic CTL [7] with probabilities. This allows us to express properties like "after a request for a service, there is 99% chance of fulfilling the request". PCTL formulas are interpreted over DTMCs and model checking on DTMCs can be done in PTIME. The problem of model checking PCTL properties for IMCs was studied in [11], it provides a PSPACE algorithms for both UMC and IMDP semantics for interval Markov chains. Furthermore, NP and co-NP hardness was shown for model checking in UMC semantics and PTIME hardness for IMDP semantics which follows from PTIME hardness of model checking PCTL formulas on DTMCs. [4] improved the upper bound and showed that model checking problem for IMDP semantics is in co-NP. This result is shown for a richer class of logic, called ω-*PCTL*, which allow Büchi and co-Büchi properties in the formula.

In the literature, the intervals of IMCs are always assumed to be closed. This assumption is sensible from the model checking perspective in IMDP semantics as models with open interval may not have an optimal value of satisfying a temporal property. The focus of this paper is to study IMDP semantics of IMCs with open intervals. We will later contrast it with the UMC semantics, and will see that the existing algorithm is applicable for IMCs with open intervals, but its outcome may vary with the model at hand. The main intuition is that the value of reachability property in a IMC with open intervals can be made arbitrarily close to the value of the property obtained by *closing* the intervals. We use this observation to show the equivalence between model checking IMCs with open interval and IMCs with closed intervals.

2 Interval Markov Chains

Definition 1. *Let $\mathcal{I}$ be the set of intervals (open or closed) in the range $[0,1]$. The subsets $\mathcal{I}_0 \triangleq \{(a,b] \mid 0 \leq a < b \leq 1\}$, $\mathcal{I}_1 \triangleq \{(a,b) \mid 0 \leq a < b \leq 1\}$, $\mathcal{I}_2 \triangleq \{[a,b) \mid 0 \leq a < b \leq 1\}$ and $\mathcal{I}_3 \triangleq \{[a,b] \mid 0 \leq a \leq b \leq 1\}$. $\mathcal{I} = \bigcup_{i\in\{0,1,2,3\}} \mathcal{I}_i$.*

Let $I \triangleq \langle a,b\rangle$ be an interval in $\mathcal{I}$, where $\langle \in \{(,[\}$ and $\rangle \in \{),]\}$. The lower bound $I{\downarrow} = a$ and upper bound is $I{\uparrow} = b$. Point intervals ($[a,a]$) are closed intervals where the upper and lower bounds are equal. The closure of an interval I, denoted by $\bar{I}$, is the smallest closed interval that includes I.

Definition 2. *A discrete time Markov chain (DTMC) is a tuple $M = (S, L, \delta)$ where S is a finite set of states, $L : S \rightarrow 2^{AP}$ is a labeling function (AP is the set of atomic propositions), $\delta : S \rightarrow S \rightarrow [0,1]$ is a transition probability matrix, such that for all $s \in S$, $\sum_{t\in S} \delta(s)(t) = 1$.*

For simplicity of notation we will use the un-Curry notation $\delta(s,t)$ for $\delta(s)(t)$. A path π of a DTMC M is an infinite sequence of states $\pi = s_0 s_1 \ldots$ such that for all $i \geq 0$, $\delta(s_i, s_{i+1}) > 0$. The i^{th} state of the path π is denoted by $\pi_i = s_i$. Let Ω_s be the set of paths starting from state s. The cylinder (open) set $Cyl(\rho)$ is the set of all paths with ρ as prefix. Let $\mathcal{B}$ be the smallest Borel σ-algebra defined on the cylinder sets. Let ρ be a finite sequence of states $s_0 s_1 \dots s_n$ such that $\delta(s_i, s_{i+1}) > 0$ for all $0 \leq i < n$. The unique measure μ is thus induced from δ as, $\mu(Cyl(\rho)) = \delta(s_0, s_1)\cdot\delta(s_1, s_2) \dots \cdot\delta(s_{n-1}, s_n)$.

Definition 3. *An Interval Markov chain (IMC) is a tuple $\mathcal{M} \triangleq (S, L, \delta)$, where S is a (finite) set of states and L is a labeling function $L : S \to 2^{AP}$, where AP is the set of atomic propositions. δ is a function $\delta : S \to \mathcal{D}$, where $\mathcal{D}$ is the set of functions from the set of states to the set of intervals $\mathcal{I}$, i.e., $\mathcal{D} = S \to \mathcal{I}$.*

As before, we will use the un-Curry notation $\delta(s, t)$ for $\delta(s)(t)$. For a state s, the probability of a single step from s to t lies in the interval $\delta(s, t)$. Thus an IMC defines a collection of Markov chains, where the single step transition probability of moving from state s to t lies in the interval $\delta(s, t)$. Not every IMC defines a collection of Markov chains. Thus, we have the notion of realizability.

Definition 4. *Let $\mathcal{M} = (S, L, \delta)$ be an IMC with states $S = \{s_1, \dots, s_m\}$. Let $D^{\mathcal{M}}$ be the set of $m \times 1$ vectors $\boldsymbol{d}$, such that $\boldsymbol{d}^T \cdot \boldsymbol{1} = 1$, which represents the set of distributions on states of $\mathcal{M}$. Where $\mathcal{M}$ is fixed we denote the set as D.*

$\mathcal{M}$ is said to be realizable *if for each set of intervals defined by $\delta(s)$, there exists a distribution $\boldsymbol{d}$ such that for all $i \in [1, m]$ $\boldsymbol{d}_i$ (the i^{th} component of $\boldsymbol{d}$) is in $\delta(s, s_i)$. The distribution $\boldsymbol{d}$ is said to be a solution of $\delta(s)$. Let $sol(s)$ be the set of solutions of $\delta(s)$.*

Next we give two semantics of IMCs: 1) *Uncertain Markov chains* (UMC), 2) *Interval Markov decision process* (IMDP).

Definition 5. *(**Uncertain Markov chain semantics**) An IMC $\mathcal{M} = (S, L, \delta)$ represents a set of DTMCs, denoted by $[\mathcal{M}]_u$, such that for each DTMC $M = (S, L, \delta_M)$ in $[\mathcal{M}]_u$, $\delta_M(s)$ is a solution of $\delta(s)$ for every state $s \in S$. In UMC semantics, we assume that nature non-deterministically picks a solution of $\delta(s)$ for each state $s \in S$, and then all transitions behave according to the chosen transition probability matrix.*

To define interval Markov decision process semantics, we need the notion of *schedulers*. The schedulers resolve the non-determinism at each state s by choosing a particular distribution from $sol(s)$.

Definition 6. *A* scheduler *of an IMC $\mathcal{M} = (S, L, \delta)$ is a function $\eta : S^+ \to D^{\mathcal{M}}$, such that for every finite sequence of states $\pi \cdot s$ of $\mathcal{M}$, $\eta(\pi \cdot s)$ is a solution of $\delta(s)$.*

A path $w = s_0 s_1 s_2 \dots$ of an IMC $\mathcal{M}$ is an infinite sequence of states. A path w starting from a state s (i.e., $w_0 = s$) is said to be *according* to the scheduler η if for all $i \geq 0$, $\eta(w_0, \dots, w_i)(w_{i+1}) > 0$. A scheduler is *memoryless* if the choice of the distribution depends solely on the current state, that is, $\eta : S \to D^{\mathcal{M}}$.

Definition 7. *(**Interval Markov decision process semantics**) In IMDP semantics, before every transition from a state s of a IMC $\mathcal{M} = (S, L, \delta)$, nature chooses a solution of $\delta(s)$ and then takes a one-step probabilistic transition* *according to the chosen distribution. In other words, nature chooses a scheduler η which then defines a DTMC M. The set of all DTMC in this semantics is denoted by $[\mathcal{M}]_d$.*

Obviously, for any IMC $\mathcal{M}$ we have:

$$[\mathcal{M}]_u \subseteq [\mathcal{M}]_d.$$

Given an IMC $\mathcal{M}$ and a state s, let σ-algebra $(\Omega_s, \mathcal{F})$ be the smallest σ-algebra on the cylinder sets of Ω_s, where Ω_s is the set of infinite paths starting from s. For each scheduler η we have a probability measure Pr^η (also denoted by $\mu^\eta_{\mathcal{M}}$) on the events in $\mathcal{F}$.

3 Probabilitic Computation Tree Logic

Probabilistic computation tree logic (PCTL) [8] replaces the path quantifiers in CTL by probabilistic operators. It has the following syntax:

$$\begin{aligned} f &::= a \mid {\sim}f \mid f \wedge f \mid \mathsf{P}_{\bowtie p} g \\ g &::= \mathsf{X} f \mid f \;\mathsf{U}\; f \end{aligned}$$

where $a \in AP$, f is called a state formula, g is called a path formula, $\bowtie \in \{<, \leq, >, \geq\}$ and p is a rational number in $[0, 1]$. The PCTL semantics is define on DTMCs. A DTMC M satisfies a state formula f at a state s if:

$$\begin{array}{ll} M, s \models a & \text{iff } a \in L(s) \\ M, s \models {\sim}f & \text{iff } M, s \not\models f \\ M, s \models f_1 \wedge f_2 & \text{iff } M, s \models f_1 \text{ and } M, s \models f_2 \\ M, s \models \mathsf{P}_{\bowtie p} g & \text{iff } Pr\{s \models g\} \bowtie p, \end{array}$$

where $\{s \models g\} = \{w \mid w_0 = s \text{ and } M, w \models g\}$. A path formula g is true for a path w of M if:

$$\begin{array}{ll} M, w \models \mathsf{X} f & \text{iff } M, w_1 \models f \\ M, w \models f_1 \;\mathsf{U}\; f_2 & \text{iff } \exists i : M, w_i \models f_2 \text{ and } \forall j < i : M, w_j \models f_1 \end{array}$$

We will denote the satisfaction relation by $s \models f$ (and $w \models g$) when M is fixed. Next we define the satisfaction relation of a PCTL formula f for an IMC $\mathcal{M}$ for the two semantics. In UMC semantics, $\mathcal{M}, s \models_u f$ iff for every DTMC $M \in [\mathcal{M}]_u$, $M, s \models f$. Note that for a PCTL formula f, $\mathcal{M}, s \models_u f$ does not imply $\mathcal{M}, s \not\models_u {\sim}f$. In IMDP semantics, the satisfaction of a PCTL formula f by a state s of $\mathcal{M}$ ($\mathcal{M}, s \models_d f$) is the same as for a DTMC except the formula with probabilistic operator, which is as follows:

$$\mathcal{M}, s \models \mathsf{P}_{\bowtie p} g \text{ iff } \forall \eta : Pr^\eta_{\mathcal{M}}\{w \mid w_0 = s \text{ and } M, w \models g\} \bowtie p$$

Particularly,

$$\begin{aligned} s &\models \mathrm{Pr}_{\leq c}\, g \text{ iff } \sup\nolimits_\eta Pr^\eta(s \models g) \leq c \\ s &\models \mathrm{Pr}_{< c}\, g \text{ iff } \sup\nolimits_\eta Pr^\eta(s \models g) < c \\ s &\models \mathrm{Pr}_{\geq c}\, g \text{ iff } \inf\nolimits_\eta Pr^\eta(s \models g) \geq c \\ s &\models \mathrm{Pr}_{> c}\, g \text{ iff } \inf\nolimits_\eta Pr^\eta(s \models g) > c \end{aligned} \tag{1}$$

Thus if event $E \in \mathcal{F}$ defines a set of paths, we are interested in the values

$$\inf_{\eta} Pr^{\eta}_{\mathcal{M}}(E) \quad \text{and} \quad \sup_{\eta} Pr^{\eta}_{\mathcal{M}}(E)$$

Open intervals present a problem for model checking in IMDP semantics. There might not exist a scheduler that gives the optimal values. Consider the reachability problem for IMCs in the following example:

Example 1. It is possible that an optimal scheduler may not exist for IMCs with open intervals. Consider the following example Figure 1, E is the set of paths that eventually reach the state s_1 from s_0. $\inf_\eta Pr^\eta(E) = 0.6$, but no scheduler gives the probability of reaching s_1 from s_0 as 0.6. The reason for this is the open lower bound of $(0.3, 1]$.

4 ϵ-Approximate Scheduler for Reachability

In this section we consider the reachability problem in IMDP semantics for IMCs with open intervals. As observed in the previous example, an optimal scheduler may not exists, thus we will construct ϵ-approximate schedulers.

An IMC is called a *closed* IMC if the probability interval of every transition is closed. We can obtain a closed IMC from an arbitrary IMC by taking the closure of the probability intervals.

Definition 8. *Given an IMDP $\mathcal{M} \triangleq (S, L, \delta)$, a* closed *IMDP $\bar{\mathcal{M}}$ is defined as (S, L, δ'), where for every s, t, $\delta'(s, t) = \bar{\delta}(s, t)$.*

Example 2. The closed IMC $\bar{\mathcal{M}}$ for $\mathcal{M}$ in the example 1 is shown below:

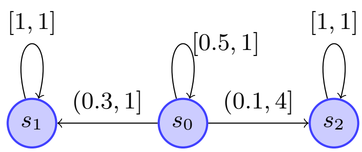


Fig. 1. A interval Markov chain

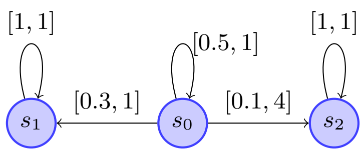


Fig. 2. A closed interval Markov chain

Evidently, if an IMC $\mathcal{M}$ is realizable then $\bar{\mathcal{M}}$ is also realizable.

Definition 9. *Basic feasible solution (BFS). Given a set of closed intervals $R \triangleq \{I_1, \ldots, I_m\}$ a basic feasible solution $\boldsymbol{d}$ is an $m \times 1$ vector, such that there exists a set $H \subseteq R$ with $|H| \geq |R| - 1$ and for all $I_i \in H$, $\boldsymbol{d}_i = I_i{\downarrow}$ or $\boldsymbol{d}_i = I_i{\uparrow}$, and $\boldsymbol{d}^T \cdot \boldsymbol{1} = 1$.*

BFSs of a set of intervals $\mathcal{J}$ that contains open intervals are the BFSs of the set of closed intervals $\bar{\mathcal{J}}$. We have the following observation.

Proposition 1. *Every solution of a set of (open or closed) intervals, can be represented as the convex combination of the BFSs.*

Proposition 2 ([4]). *Let $\mathcal{M}$ be a closed IMC, and E be an event defining the reachability of some set of states $T \subseteq S$. There exists a memoryless scheduler η such that the probability of the event E is optimal.*

The proposition says that, if $\mathcal{M}$ is closed then we have a scheduler $\eta : S \to D^{\mathcal{M}}$ such that $Pr^{\eta}(E) = \inf_{\eta'} Pr^{\eta'}(E)$ (or $\sup_{\eta'} Pr^{\eta'}(E)$), and η chooses at each state s one of the BFSs of $\delta(s)$ (pure scheduler). The proposition follows directly from the existence of an optimal scheduler for reachability in Markov Decision Processes [2].

The main theorem of this paper is as follows:

Theorem 1. *Let E be the event describing the set of paths of an IMC $\mathcal{M}$ starting from a state s and eventually reaching some goal states T. Then:*

$$\forall \varepsilon > 0 \;\; \exists \hat{\eta} : |\min_{\eta} Pr^{\eta}_{\bar{\mathcal{M}}}(E) - Pr^{\hat{\eta}}_{\mathcal{M}}(E)| \leq \varepsilon$$

and

$$\forall \varepsilon > 0 \;\; \exists \hat{\eta} : |\max_{\eta} Pr^{\eta}_{\bar{\mathcal{M}}}(E) - Pr^{\hat{\eta}}_{\mathcal{M}}(E)| \leq \varepsilon$$

Proof. Let $\mathcal{M} \triangleq (S, L, \delta)$ and $\bar{\mathcal{M}} \triangleq (S, L, \delta')$. $\bar{\mathcal{M}}$ is closed, thus by Prop. 2 an optimal scheduler exists. Let $\overset{*}{\eta}$ be an optimal scheduler that minimizes $\mathrm{Pr}^{\eta}_{\bar{\mathcal{M}}}(E)$. Furthermore, $\overset{*}{\eta}$ is memoryless, deterministic and chooses one of the BFS of $\delta'(s)$ at each state s. Hence, $\overset{*}{\eta}$ induces a DTMC on $\bar{\mathcal{M}}$, and $\overset{*}{\eta}(s,t)$ defines the single step transition probability from a state s to a state t.

Let the stochastic matrix $\overset{*}{P}$ be such that each row is identified with a state of $\bar{\mathcal{M}}$. We have :

$$\overset{*}{P}(s,t) = \overset{*}{\eta}(s,t) \quad \text{if } s \notin T \quad \text{and} \quad \overset{*}{P}(s,s) = 1 \quad \text{if } s \in T \tag{2}$$

Let $A = (1 + \overset{*}{P} + (\overset{*}{P})^2 + (\overset{*}{P})^3 \ldots)$, A is well-defined stochastic matrix as the series converges. Let $\gamma = \|A\|_{\infty}$.

Now we are in a position to define a scheduler $\hat{\eta}$ for the IMC $\mathcal{M}$. The scheduler $\hat{\eta}$ is a function, $\hat{\eta} : S \times \mathbb{N} \to D^{\mathcal{M}}$. We assume that there are no positive point intervals. (We can set the value of $\hat{\eta}$ if point intervals are present.) Define the following:

$$\begin{aligned}
Q_s &= \{t \mid \overset{*}{\eta}(s,t) > 0,\ \overset{*}{\eta}(s,t) \notin \delta(s)\} \\
L_s &= \{t \mid \overset{*}{\eta}(s,t) \in \delta(s,t),\ \overset{*}{\eta}(s,t) = \delta(s,t)\!\downarrow\} \\
R_s &= \{t \mid \overset{*}{\eta}(s,t) \in \delta(s,t),\ \overset{*}{\eta}(s,t) = \delta(s,t)\!\uparrow\} \\
I_s &= \{t \mid \overset{*}{\eta}(s,t) \in \delta(s,t),\ \overset{*}{\eta}(s,t) \neq \delta(s,t)\!\uparrow,\ \overset{*}{\eta}(s,t) \neq \delta(s,t)\!\downarrow\} \\
\rho &= \min\{\{x \mid \exists s, \exists t \in L_s \cup I_s : x = \overset{*}{\eta}(s,t) - \delta(s,t)\!\downarrow\}, \\
&\qquad \{x \mid \exists s, \exists t \in R_s \cup I_s : x = \delta(s,t)\!\uparrow - \overset{*}{\eta}(s,t)\}, \\
&\qquad \{x \mid \exists s, \exists t \in Q_s : x = \delta(s,t)\!\uparrow = \delta(s,t)\!\downarrow\}\}
\end{aligned}$$

Observe that ρ is a constant of the model $\mathcal{M}$. Let $\hat{\eta}$ be defined as follows:

– Let $t \in Q_s$. This implies $\overset{*}{\eta}(s,t) = \delta(s,t){\uparrow}$ or $\overset{*}{\eta}(s,t) = \delta(s,t){\downarrow}$. If $\overset{*}{\eta}(s,t) = \delta(s,t){\uparrow}$ then $\delta(s,t)$ is open from above and $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t) - 2^{-n}\frac{\kappa\rho}{|Q_s|}$, where $\kappa = \frac{\varepsilon}{1+\gamma}$. Similarly, if $\overset{*}{\eta}(s,t) = \delta(s,t){\downarrow}$ then $\delta(s,t)$ is open from below and $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t)+2^{-n}\frac{\kappa\rho}{|Q_s|}$.
– Let $t \in R_s$ and $\alpha \triangleq \sum_{t\in Q_s} \hat{\eta}(s,n,t) - \overset{*}{\eta}(s,t)$. If $\alpha < 0$ then for all $t \in R_s \cup I_s$, $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t)+\frac{\alpha}{|R_s \cup I_s|}$ and for $t \in L_s$, $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t)$. If $\alpha > 0$ then for all $t \in L_s \cup I_s$, $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t) + \frac{\alpha}{|L_s \cup I_s|}$ and for $t \in R_s$, $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t)$. If $\alpha = 0$ then for all $t \in L_s \cup I_s \cup R_s$, $\hat{\eta}(s,n,t) = \overset{*}{\eta}(s,t)$.

It remains to prove that $\boldsymbol{d} = \overset{*}{\eta}(s,n)$, defined above, is a solution to $\delta(s)$. From the construction it follows that $\sum_{t\in S} \boldsymbol{d}_t = 1$ and hence it is a valid distribution on the states of the IMC $\mathcal{M}$. Consider the following cases: $t \in Q_s$ and $\overset{*}{\eta}(s,t) = \delta(s,t){\uparrow}$, the upper bound of $\delta(s,t)$ is open. The lower bound of $\delta(s,t)$ is strictly smaller than $2^{-n}\kappa\rho$ for any $n \in \mathbb{N}$ i.e., $\delta(s,t){\downarrow} < \kappa\rho$ since ρ is at the most as large as the smallest interval in $\mathcal{M}$. Thus $\boldsymbol{d}_t \in \delta(s,t)$. Similarly, for every $t \in Q_s$, $\boldsymbol{d}_t \in \delta(s,t)$. Suppose $\alpha < 0$, then $R_s \cup I_s$ is not empty, else $\delta(s)$ will not be realizable. The changes to the probability for a transition s to t, where $t \in R_s \cup I_s$ is small enough so that $\boldsymbol{d}_t \in \delta(s,t)$. Thus, for every t, $\boldsymbol{d}_t \in \delta(s,t)$, or equivalently $\boldsymbol{d}$ is a solution to $\delta(s,t)$. Identical argument holds when $\alpha > 0$.

Let $\hat{P}_n$ be a sub-stochastic matrix defined as follows: $\hat{P}_n(s,t) = \hat{\eta}(s,t)$ if $\overset{*}{P}(s,t) > 0$ else $\hat{P}_n(s,t) = 0$. In other words, $\hat{P}_n(s,t) > 0$ if the state t is in $support(\overset{*}{\eta}(s))$.

$$\hat{P}_n = \overset{*}{P} + P_n \tag{3}$$

where $|P_n(s,t)| \leq 2^{-n}\kappa\rho$ for every (s,t).

Let $\overset{*}{\eta}$ and $\hat{\eta}$ induce DTMCs M' and M on the IMCs $\bar{\mathcal{M}}$ and $\mathcal{M}$, respectively. Let the corresponding σ-algebra be $\mathcal{S} \triangleq (\Omega_s, \mathcal{F}, \overset{*}{\mu})$ and $\mathcal{S}' \triangleq (\Omega_s, \mathcal{F}, \hat{\mu})$, where s is some state of $\mathcal{M}$ and Ω_s is the set of paths starting from state s. Define $\overset{*}{R} \triangleq \{w \in \Omega_s \mid w \text{ is according to } \overset{*}{\eta}\}$ and $\hat{R} \triangleq \{w \in \Omega_s \mid w \text{ is according to } \hat{\eta}\}$, i.e., $\overset{*}{R}$ and R are set of paths in M' and M, respectively. Let $B \in \mathcal{F}$ be the event of reaching the goal states T, and $E = \overset{*}{R} \cap B$ and $E' = \hat{R} \cap B$. It follows from the construction that $E \subseteq E'$. Define $A_i \triangleq \{w \mid \exists u \in E : w_0 \dots w_i = u_0 \dots u_i \text{ and } \overset{*}{\eta}(w_i, w_{i+1}) = 0, \hat{\eta}(w_i, i, w_{i+1}) > 0\}$. Let $A = \bigcup_i A_i$. It is easy to see that, $E' \cap \bar{A} = E$. We will first show that the event A has a very small probability measure in $\mathcal{S}'$:

$$\hat{\mu}(A) = Pr^{\hat{\eta}}_M(A) = \sum_{i=0} Pr^{\hat{\eta}}_M(A_i)$$

If $w \in A_i$ then $\delta(w_i, w_{i+1}){\uparrow} > 0$ and $\overset{*}{\eta}(s,t) = 0$. Thus,

$$Pr^{\hat{\eta}}_M(A_i) \leq 2^{-i}\kappa\rho \;\text{ or }\; Pr^{\hat{\eta}}_M(A) \leq \kappa\rho$$

Thus,

$$\hat{\mu}(A) \leq \kappa\rho \tag{4}$$

We will now show that the probability of E' can be made infinitesimally close to the probability of E. Formally, we will show, $|\hat{\mu}(E') - \bar{\mu}(E)| \leq \varepsilon$. The left hand side can be written as:

$$\begin{aligned} |\hat{\mu}(E') - \bar{\mu}(E)| &= |\hat{\mu}(E' \cap A) + \hat{\mu}(E' \cup \bar{A}) - \bar{\mu}(E)| \\ &\leq |\hat{\mu}(E) - \bar{\mu}(E)| + \kappa\rho \end{aligned} \tag{5}$$

That is, we restrict to the paths that belong to E. Let x_s^n denote the probability of reaching the goal states T at the n^{th} step in M' from the state s. Let E_n be the event of reaching the goal states T at the n^{th} step in the Markov chain M such that $E_n \subseteq E$ and thus $\bigcup_n E_n = E$. Let $y_s^n = \hat{\mu}(E_n)$. Thus, we can write the following:

$$\begin{aligned} x_s^{n+1} &= \sum_{t \in support(\overset{*}{\eta}(s))} \overset{*}{P}(s,t) x_t^n, \\ y_s^{n+1} &= \sum_{t \in support(\overset{*}{\eta}(s))} \hat{P}_n(s,t) y_t^n. \end{aligned}$$

Or, using vector notation, $\boldsymbol{x}_{n+1} = \overset{*}{P}\boldsymbol{x}_n$ and $\boldsymbol{y}_{n+1} = \hat{P}_n\boldsymbol{y}_n$. Therefore:

$$\begin{aligned} \boldsymbol{y}_{n+1} - \boldsymbol{x}_{n+1} &= \overset{*}{P}(\boldsymbol{y}_n - \boldsymbol{x}_n) + P_n\boldsymbol{y}_n \qquad \text{from equation (3)} \\ &\leq \overset{*}{P}(\boldsymbol{y}_n - \boldsymbol{x}_n) + 2^{-n}\kappa\rho\boldsymbol{1} \\ &\leq 2^{-n}\kappa\rho(1 + \overset{*}{P} + \overset{*}{P}{}^2 + \ldots)\boldsymbol{1} \\ \text{Thus, } \|\boldsymbol{y}_{n+1} - \boldsymbol{x}_{n+1}\|_\infty &\leq 2^{-n}\kappa\rho\gamma. \end{aligned}$$

We have,

$$|\hat{\mu}(E) - \bar{\mu}(E)| \;\leq\; |\sum_n (y_s^n - x_s^n)| \;\leq\; \sum_n 2^{-n}\kappa\rho\gamma \;\leq\; \kappa\rho\gamma$$

Combining this with equation (5) we can conclude:

$$|\hat{\mu}(E') - \bar{\mu}(E)| \;\leq\; (1+\gamma)\kappa\rho \;\leq\; \varepsilon$$

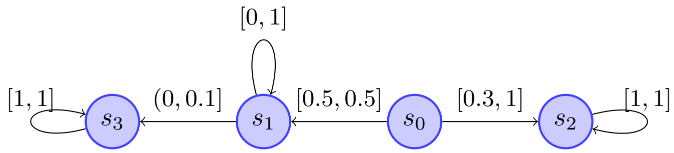
By similar argument we conclude $\forall \varepsilon > 0\ \exists \hat{\eta} : |\max_\eta Pr^\eta_{M'}(E) - Pr^{\hat{\eta}}_M(E)| \leq \kappa$.

Corollary 1. *Let E be the set of paths that reach some goal states T of IMC $\mathcal{M}$. Then:*

$$\min_\eta Pr^\eta_{\bar{\mathcal{M}}}(E) = \inf_\eta Pr^\eta_{\mathcal{M}}(E) \textit{ and } \max_\eta Pr^\eta_{\bar{\mathcal{M}}}(E) = \sup_\eta Pr^\eta_{\mathcal{M}}(E).$$

Proof. We need to show $\forall \kappa > 0\ \exists \hat{\eta} : |\min_\eta Pr^\eta_{M'}(E) - Pr^{\hat{\eta}}_M(E)| \leq \kappa$. Observe that, $\hat{\eta}$ is also a scheduler of M', thus, $Pr^{\hat{\eta}}_M(E) - \min_\eta Pr^\eta_{M'}(E) \leq \kappa$. Similarly, for all $\kappa > 0$ there exists a scheduler $\hat{\eta}$ of M such that $\max_\eta Pr^\eta_{M'}(E) - Pr^{\hat{\eta}}_M(E) \leq \kappa$.

Example 3. In UMC semantics, the nature picks the probability transition matrix and the model behaves according to it. The infimum (or supremum) probability of reaching some state is different than the infimum probability in IMDP semantics. This becomes apparent in the following IMC with an open interval:



The minimum and maximum probability of reaching state s_3 from s_0 in the UMC semantics is 0.5. But for any $\epsilon > 0$ there exists a scheduler for which the probability of reaching s_3 is smaller than ϵ. That is, the infimum of the probability of reaching state s_3 is 0.

5 PCTL Model Checking

In this section we briefly recall PCTL model checking on DTMC and IMCs with closed intervals (for the two semantics), and then show how to use the result of previous section to do model checking for IMCs with open intervals.

Model checking of PCTL [1,6] formula f on DTMC M proceeds much like the CTL model checking on Kripke structures [5]. The satisfiability of a (state) sub-formula f' of f for a state s of M is iteratively calculated and the labeling functions are updated accordingly. For example, for the until formula $f = \mathsf{P}_{\bowtie p}(f_1 \mathsf{U} f_2)$ and a state s, the formula f is added to the label of s iff the probability of reaching states with label f_2, via states with label f_1 satisfies $\bowtie p$. This can be done in polynomial time by solving linear constraints. Finally, a state $s \models f$ if $f \in L(s)$ and the model checking problem can be solved in polynomial time.

Model checking in UMC semantics uses the *existential* theory of reals [10]. An IMC $\mathcal{M}, s \models_u f$ in UMC semantics iff for all DTMC $M \in [\mathcal{M}]_u$, $M, s \models f$, or equivalently, $\mathcal{M}, s \not\models_u f$ iff there exists a $M \in [\mathcal{M}]_u$ such that $M, s \models \sim f$. Basically, we use parameters to encode the transition probabilities which are constrained by the intervals and construct a formula Γ in existential theory of reals such that Γ is satisfiable iff there exists a $M \in [\mathcal{M}]_u$ such that $M, s \models \sim f$ [4]. Observe, that the presence (or absence) of open intervals does not affect the algorithm and the algorithm operates in PSPACE.

Model checking in IMDP semantics is done by first transforming the IMC into an Markov decision process (MPD) and then doing model checking on the MDP [2]. Let $\mathcal{M} = (S, L, \delta)$ be a *closed* IMC and for each state $s \in S$, let B_s be the set of basic feasible solution of $\delta(s)$. Let $D_{\mathcal{M}} = (S, L, \mu)$ be the MDP with $\mu : S \to S \to [0, 1]$, where $\mu(s) = B_s$. From Proposition 1, we can deduce that, a DTMC $M \in [\mathcal{M}]_d$ iff M is induced by some scheduler η of $D_{\mathcal{M}}$. Model checking of MDP proceeds the same way as model checking of DTMC. We iteratively update the labels of the state with (state) sub-formulas. Conjunctions and disjunctions are handled as in the DTMC model checking. Interesting cases are formulas with probabilistic operator and negations. Let g be a path formula and $\mathsf{P}_{\succ p} g$ (or $\mathsf{P}_{\prec p} g$) is added to the label of a state $s \in S$, iff

$$\min_{\eta} Pr^{\eta}_{D_{\mathcal{M}}}(s \models g) \succ p \text{ (or } \max_{\eta} Pr^{\eta}_{D_{\mathcal{M}}}(s \models g) \prec p)$$

where $\succ \in \{\geq, >\}$ ($\prec \in \{\leq, <\}$). This is done by solving a linear optimization problem. We use the following proposition to handle formulas with negations.

Proposition 3. *For any $E \in \mathcal{F}$ of $(\Omega_s, \mathcal{F})$ on MDP M,*

$$\inf_{\eta} Pr^{\eta}(E) \;=\; 1 - \sup_{\eta} Pr^{\eta}(\bar{E})$$

Thus, model checking MDPs boils down to solving successive reachability optimization problems. Note that direct application of this method to IMCs with open interval is not possible since no scheduler exists which may yields the value $\inf_{\eta} Pr^{\eta}_{D_{\mathcal{M}}}(s \models g)$.

In the rest of the section we use the above mentioned model checking mechanism to show that model checking IMCs with open interval in IMDP semantics, reduces to model checking its closure.

Theorem 2. *Given a PCTL formula f and an IMC $\mathcal{M}$,*

$$\mathcal{M}, s \models f \quad \textit{iff} \quad \bar{\mathcal{M}}, s \models f$$

Proof. We assume that $\mathcal{M}$ has open intervals. We proceed by induction on the structure of the formula f. We have the following cases:

1. Let $f := a$. The labeling function of s in $\mathcal{M}$ and $\bar{\mathcal{M}}$ are identical. Thus, $\mathcal{M}, s \models f$ iff $\bar{\mathcal{M}}, s \models f$.
2. Let $f := {\sim} f'$. From the induction hypothesis, $\mathcal{M}, s \not\models f'$ iff $\bar{\mathcal{M}}, s \not\models f'$. Thus, $\mathcal{M}, s \models f$ iff $\bar{\mathcal{M}}, s \models f$.
3. Let $f := f_1 \wedge f_2$. From the induction hypothesis, $\mathcal{M}, s \models f_1$ iff $\bar{\mathcal{M}}, s \models f_1$ and $\mathcal{M}, s \models f_2$ iff $\bar{\mathcal{M}}, s \models f_2$. Thus, $\mathcal{M}, s \models f$ iff $\bar{\mathcal{M}}, s \models f$.
4. Let $f := [\mathsf{X} f']_{\bowtie c}$. Consider the case $\bowtie \in \{\geq, >\}$. Suppose $\overset{*}{\eta}$ be the optimal scheduler of $\bar{\mathcal{M}}$ such that $\mathrm{Pr}^{\overset{*}{\eta}}_{\bar{\mathcal{M}}}(\mathsf{X} f') = \min_{\eta} \mathrm{Pr}^{\eta}_{\bar{\mathcal{M}}}(\mathsf{X} f')$.
   We show that for every ε we can construct a scheduler $\hat{\eta}$ of $\mathcal{M}$ such that

   $$Pr^{\hat{\eta}}_{\mathcal{M}}(\mathsf{X} f') - Pr^{\overset{*}{\eta}}_{\bar{\mathcal{M}}}(\mathsf{X} f') \leq \varepsilon.$$

   Observe that, any scheduler of $\mathcal{M}$ is also a scheduler of $\bar{\mathcal{M}}$, since for any states $s, t \in S$ $\delta(s,t) \subseteq \bar{\delta}(s,t)$. Thus, Corollary 1. is applicable. Let $Q_s \triangleq \{t \mid \overset{*}{\eta}(s,t) > 0, \overset{*}{\eta}(s,t) \notin \delta(s)\}$ and $R_s \triangleq \{t \mid \overset{*}{\eta}(s,t) > 0, \overset{*}{\eta}(s,t) \in \delta(s,t)\}$. We assume that Q_s, R_s are not empty and there are no point intervals. Let $\hat{\eta}(s) = \boldsymbol{d}$, where $\boldsymbol{d}$ is defined as follows:

– Let $t \in Q_s$. This implies $\overset{*}{\eta}(s,t) = \delta(s,t)\uparrow$ or $\overset{*}{\eta}(s,t) = \delta(s,t)\downarrow$. If $\overset{*}{\eta}(s,t) = \delta(s,t)\uparrow$ then $\delta(s,t)$ is open from above and $\boldsymbol{d}_t = \overset{*}{\eta}(s,t) - \frac{\varepsilon\rho}{|S|}$, where ρ is the minimum of the length of the non-zero interval in $\mathcal{M}$ and the $\overset{*}{\eta}(s,t)$ for $t \in R_s$. Similarly, if $\overset{*}{\eta}(s,t) = \delta(s,t)\downarrow$ then $\delta(s,t)$ is open from below and $\boldsymbol{d}_t = \overset{*}{\eta}(s,t) + \frac{\varepsilon\rho}{|S|}$.
– Let $t \in R_s$ and $\alpha \triangleq 1 - \sum_{t \in Q_s} \boldsymbol{d}_t - \sum_{t \in R_s} \overset{*}{\eta}(s,t)$. We have $\boldsymbol{d}_t = \overset{*}{\eta}(s,t) + \frac{\alpha}{|R_s|}$.

It follows that $\boldsymbol{d}$ is a distribution on the states of $\mathcal{M}$ and is a solution to $\delta(s)$. Let $E \triangleq \{w \mid \overset{*}{\eta}(w_0, w_1) > 0 \text{ and } \bar{\mathcal{M}}, w_1 \models f'\}$ and $E' \triangleq \{w \mid \hat{\eta}(w_0, w_1) > 0 \text{ and } \mathcal{M}, w_1 \models f'\}$.

$$|\Pr_{\bar{\mathcal{M}}}^{\overset{*}{\eta}}(E) - \Pr_{\mathcal{M}}^{\hat{\eta}}(E')| \leq \sum_{t \in support(\hat{\eta}(s))} \frac{\varepsilon\rho}{|S|} \quad \leq \quad \varepsilon$$

Thus we can conclude that $\inf_\eta \Pr^\eta_{\mathcal{M}}(\mathsf{X} f') = \min_\eta \Pr^\eta_{\bar{\mathcal{M}}}(\mathsf{X} f')$. By similar argument:

$$\sup_\eta Pr^\eta_{\mathcal{M}}(\mathsf{X} f') \;=\; \max_\eta Pr^\eta_{\bar{\mathcal{M}}}(\mathsf{X} f').$$

$\mathcal{M}, s \models [\mathsf{X} f']_{\bowtie c}$ iff $\bar{\mathcal{M}}, s \models [\mathsf{X} f']_{\bowtie c}$, where $\bowtie \in \{\leq, <\}$.

5. Let $f := [f_1 \mathsf{U} f_2]_{\bowtie c}$. Suppose $\bowtie \in \{\geq, >\}$. By induction hypothesis, for every s, $\mathcal{M}, s \models f_1$ iff $\bar{\mathcal{M}}, s \models f_1$ and $\mathcal{M}, s \models f_2$ iff $\bar{\mathcal{M}}, s \models f_2$. Let $S_1 \triangleq \{s \mid s, \mathcal{M} \models f_1\}$ and $T \triangleq \{s \mid s, \mathcal{M} \models f_2\}$. The IMC $\mathcal{M}'$ is obtained from $\mathcal{M}$ by omitting states not present in the set $S_1 \cup T$. It is easy to see that, if E is the event of reaching T in $\mathcal{M}'$, then $\inf_\eta \Pr^\eta_{\mathcal{M}'}(E) = \inf_\eta \Pr^\eta_{\mathcal{M}}(f)$.
From Corollary 1 it follows that for any $0 < \varepsilon \leq 1$ we can find $\hat{\eta}$ such that $\Pr^{\hat{\eta}}_{\mathcal{M}'}(E) - \min_\eta \Pr^\eta_{\bar{\mathcal{M}}'}(E) \leq \varepsilon$, where E is the event of reaching T in $\mathcal{M}'$. Thus $\inf_\eta \Pr^\eta_{\mathcal{M}}(f) = \min_\eta \Pr^\eta_{\bar{\mathcal{M}}}(f)$. Similar argument holds for $\bowtie \in \{<, \leq\}$.

This concludes the proof.

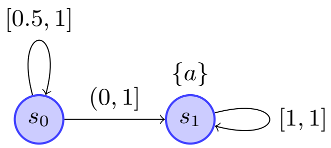


Fig. 3. A interval Markov chain

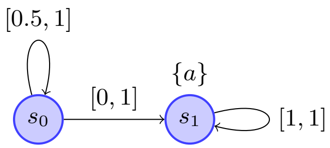


Fig. 4. A closed interval Markov chain

Example 4. Consider PCTL model checking of IMCs in UMC semantics. This involves existentially quantifying the transition probabilities and creating a formula in closed real field [4]. This captures a strict set of DTMC as compared to IMDP semantics, i.e, $[\mathcal{M}]_u \subsetneq [\mathcal{M}]_d$. For example, DTMC where the transition probability between two states s, t change over time cannot be represented in UMC semantics. This is exemplified by the IMC $\mathcal{M}$ in Figure 3. The probability of satisfying the path formula $g = \mathsf{G}\ (\sim a \wedge [\mathsf{X}a]_{>0})$ in the UMC semantics is 0. But we can find schedulers which can make the probability of satisfying g arbitrarily close to 1. The scheduler has the freedom to define an infinite Markov chain by assigning monotonically increasing probabilities for the transition $s_0 \to s_0$).

The model checking of the open IMC $\mathcal{M}$ is done by closing it (Figure 4). This gives us the closed IMC $\bar{\mathcal{M}}$, shown below: The maximum probability of satisfying g in $\bar{\mathcal{M}}$ is 1. Which implies, for every $0 < \varepsilon \leq 1$, there exists a scheduler $\hat{\eta}$, for which the probability of staying in a state that satisfies $\sim a \wedge [\mathsf{X}a]_{>0}\ (s_0)$ is greater than $1 - \varepsilon$, by Theorem 2.

6 Conclusion

We presented the problem of model checking Interval Markov chains with open intervals. We proved that as far as model checking (and reachability) is concerned open intervals do not cause any problem in interval Markov decision process semantics and thus can be safely ignored. Interval Markov chains are but special cases of more complex Markovian models, called *constraint Markov chains* (CMC) [3]. Transition probabilities in these models are defined as a solution to linear equations. Let F_V be the set of linear in-equations on variables V. A constraint Markov chain is a tuple $\mathcal{M} \triangleq (S, L, \delta)$, where the transition function $\delta : S \to 2^{F_V}$, maps each state to a set of linear in-equations. Thus IMCs are a strict sub-class of convex Markov decision process. The behaviour of a CMC can be defined in the UMC and IMDP semantics. We say, a system of in-equation are closed if they have non-strict inequalities, otherwise they are open. A CMC is called open if the transition function maps to an open system of linear equations. Model checking open CMCs have the same kinds of problems as described for IMCs. Theorem 2. can be extended to CMCs as well. We can define *basic feasible solutions* for a system of linear in-equations as well. Let s be a state of a CMC $\mathcal{M}$ and $\delta(s)$ be a system of linear in-equations on variables $\{x_1, \ldots, x_k\}$ such that x_i denotes the probability of moving from state s to s_i. The BFSs of $\delta(s)$ are the vertices of the convex hull defined by the set of in-equations $\delta(s) \cup \{x_1 + \ldots + x_k = 1\}$. The same argument as in the proof of Theorem 2 shows that, model checking of PCTL formulas on CMCs can be done by first closing the system of in-equations, this is done by replacing the strict inequalities $(<, >)$ with non-strict inequalities $(\leq, \geq)$, and then model checking on the closed model.

Acknowledgement. The authors thank Hongfei Fu for discussions on the topic of this paper.

References

1. Baier, C., Katoen, J.-P.: Principles of Model Checking (Representation and Mind Series). The MIT Press (2008)
2. Bianco, A., De Alfaro, L.: Model checking of probabilistic and nondeterministic systems. In: Thiagarajan, P.S. (ed.) FSTTCS 1995. LNCS, vol. 1026, pp. 499–513. Springer, Heidelberg (1995)
3. Caillaud, B., Delahaye, B., Larsen, K.G., Legay, A., Pedersen, M.L., Wasowski, A.: Constraint Markov chains. Theoretical Computer Science **412**(34), 4373–4404 (2011)
4. Chatterjee, K., Henzinger, T., Sen, K.: Model-checking omega-regular properties of interval Markov chains. In: Amadio, R.M. (ed) Foundations of Software Science and Computation Structure (FoSSaCS) 2008, pp. 302–317, March 2008
5. Clarke Jr., E.M., Grumberg, O., Peled, D.A.: Model Checking. MIT Press, Cambridge (1999)
6. Courcoubetis, C., Yannakakis, M.: The complexity of probabilistic verification. J. ACM **42**(4), 857–907 (1995)
7. Allen Emerson, E.: Temporal and modal logic. In: Handbook of Theoretical Computer Science (vol. b), pp. 995–1072. MIT Press, Cambridge (1990)
8. Hansson, H., Jonsson, B.: A logic for reasoning about time and reliability. Formal Aspects of Computing **6**(5), 512–535 (1994)
9. Jonsson, B., Larsen, K.G.: Specification and refinement of probabilistic processes. In: LICS, pp. 266–277. IEEE Computer Society (1991)
10. Renegar, J.: On the computational complexity and geometry of the first-order theory of the reals. part i: Introduction. preliminaries. the geometry of semi-algebraic sets. the decision problem for the existential theory of the reals. Journal of Symbolic Computation **13**(3), 255–299 (1992)
11. Sen, K., Viswanathan, M., Agha, G.: Model-checking Markov Chains in the presence of uncertainties. In: Hermanns, H., Palsberg, J. (eds.) TACAS 2006. LNCS, vol. 3920, pp. 394–410. Springer, Heidelberg (2006)
12. Walley, P.: Measure of uncertainty in expert systems. Artificial Intelligence **83**(1), 1–58 (1996)
13. Škulj, D.: Discrete time Markov Chains with interval probabilities. International Journal of Approximate Reasoning **50**(8), 1314–1329 (2009). Special Section on Interval/Probabilistic Uncertainty