On Behavioral Metric for Probabilistic Systems: Definition and Approximation Algorithm

Taolue Chen¹,³ Tingting Han², Jian Lu³
¹ CWI, Department of Software Engineering, P.O. Box 94079, 1090GB Amsterdam, The Netherlands
² RWTH Aachen, MOVES, Ahornstraße 55, D-52056 Aachen, Germany
³ State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing, Jiangsu, P.R.China, 210093

Abstract

In this paper, we consider the behavioral pseudometrics for probabilistic systems. The model we are interested in is probabilistic automata, which are based on state transition systems and make a clear distinction between probabilistic and nondeterministic choices. The pseudometrics are defined as the greatest fixpoint of a monotonic functional on the complete lattice of state metrics. A distinguished characteristic of this pseudometric lies in that it does not discount the future, which addresses some algorithmic challenges to compute the distance of two states in the model. We solve this problem by providing an approximation algorithm: up to any desired degree of accuracy ε, the distance can be approximated to within ε in time exponential in the size of the model and logarithmic in 1/ε. A key ingredient of our algorithm is to express a pseudometric being a post-fixpoint as the elementary sentence over real closed fields, which allows us to exploit Tarski's decision procedure, together with binary search to approximate the behavioral distance.

1 Introduction

Probabilistic systems, where system dynamics encodes the probability of making a transition between states rather than just the existence of such transitions, have received considerable attention from a rapidly growing research community. In order to capture nondeterminism, Segala introduced a new family of models, namely the simple probabilistic automata (SPA in short), where both probability and nondeterminism are taken into account. In a nutshell, SPAs are based on state transition systems and make a clear distinction between probabilistic and nondeterministic choices and thus constitute a very expressive framework for the specification and analysis of probabilistic systems.

In the system theory, for a given model, one of the fundamental research questions is the notion of equivalence. In the classical investigations in concurrency, bisimulation is a ubiquitous notion of system equivalence that has become one of the primary tools in the analysis of systems. In probabilistic systems, the standard notion of bisimulation has to be adapted, usually by treating the probability as labels. This line of research, dates back to Larsen and Skou’s work on pure probabilistic systems [9] and now is very fruitful.

However, it is now widely recognized that traditional equivalence is not a robust concept in the presence of quantitative (i.e. numerical) information in the model, in particular, for probabilistic models [8]. To find a more flexible way to differentiate system states, researchers have borrowed from pure mathematics the notion of metric.¹ A metric is often defined as a function that associates some distance with a pair of elements. Here, it is exploited to provide a measure of the difference between two states that are not exactly bisimilar. Having a nice pseudometric definition for systems at hand, the next natural question is: How to compute it? This raises some algorithmic challenges.

In this paper we deal with both of these questions and now summarize the main contributions as follows: First, we instantiate the (abstract) pseudometric def-

¹In this paper, the term metric is used to denote both metric and pseudometric. It turns out that in the probabilistic system, pseudometric is a more natural notion.
inition given in [6] in the setting of probabilistic automata and provide a concrete account; We emphasize that the pseudometric defined here, in contrast to the one in [4][8], does not discount the future; Second, we present an approximation algorithm to compute the behavioral distance.

Our main technique exploits a decision procedure for the theory of reals with addition and multiplication dating from [11]. Namely, we show how to express a pseudometric being a post-fixpoint in first order theory of real closed fields which is quadratic in the size of the system and has a constant number of quantifier alternations. It is known that the (first-order) theory of real closed fields is decidable in time exponential in the size of the formula and doubly exponential in the quantifier alternation depth [1]. This, together with binary search on the range of values gives rise to an exponential algorithm to approximate the value to any given ε. Our techniques are simple and combine known results to provide an algorithm and complexity bound on the general problem of approximating the distance between two states in a very general model of probabilistic systems. Due to space restriction, all of the proofs, the explicit presentation of the algorithm and more comprehensive references are omitted in the current form. We refer the interested readers to our technical report [5] for details.

2 Preliminaries

Before starting our exposition, let us first fix some general notations. Throughout the paper, we assume a fixed \( F \) of some real closed ordered field. An ordered field \( F \) is real closed if no proper algebraic extension of \( F \) is ordered. Examples of real closed fields are the real algebraic numbers, the computable numbers, the real numbers, superreal numbers, hyperreal numbers, etc. For a countable set \( X \), a probability distribution on \( X \) is a function \( δ : X \rightarrow [0,1] \cap \mathbb{F} \) such that \( \sum_{x \in X} δ(x) = 1 \). We denote the set of probability distributions on \( X \) by \( D(X) \). For a probability distribution \( δ \in D(X) \) we define \( ||δ|| \), the support of \( δ \), as \( ||δ|| = \{ x \in X \mid δ(x) > 0 \} \). Note here we exert a restriction to each distribution \( η \) such that \( η(x) \in \mathbb{F} \). Namely, for instance, if one sets \( \mathbb{F} \) as the set of real algebraic numbers, then the probability can not be, say, \( \frac{1}{e} \), which is transcendental.

2.1 Probabilistic Automata

Definition 1 (Probabilistic Automata) A probabilistic automaton is a tuple \( P = (S, A, →) \) where \( S \) is a finite set of states; \( A \) is a finite set of actions; and \( → \subseteq S \times A \times D(S) \) is the transition relation. We shall write \( s \xrightarrow{a, η} \eta \) as a more suggestive notion instead of \( (s, a, η) \subseteq → \). For each state \( s \) and action \( a \), we write \( \mathcal{I}(s, a) = \{ η \in D(S) \mid s \xrightarrow{a} η \} \).

Now, we provide a classical notion of equivalence between states, namely, the (strong) bisimulation. Assuming \( η \) is a distribution on \( S \) and \( V \subseteq S \), we write \( η(V) = \sum_{s \in V} η(s) \). We first lift an equivalence relation on \( S \) to an equivalence relation between distributions over \( S \) in the following way:

Definition 2 Let \( η, η' \in D(S) \), we say that they are equivalent w.r.t. an equivalence \( R \) on \( S \), written \( η \equiv_R η' \) if

\[ \forall U \subseteq S/R. \ η(U) = η'(U) \]

Definition 3 (Probabilistic Bisimulation) An equivalence relation \( R \subseteq S \times S \) is a (strong) bisimulation if \( sRt \) implies:

- whenever \( s \xrightarrow{a} η \), there exists \( η' \) such that \( t \xrightarrow{a} η' \) and \( η \equiv_R η' \).

Two states \( s, t \) are bisimilar, denoted by \( s \equiv_R t \), if there exists some bisimulation \( R \) s.t. \( sRt \).

3 Behavioral Metrics

In this section, we define pseudometric as a greatest fixpoint of a certain functional. Let us fix a simple probabilistic automata \( P = (S, A, →) \) and consider pseudometrics on its set of states. We note that this metric suffices even if one wants to compute the distance between the states in two different SPAs, say \( P \) and \( P' \), since we can simply take the disjoint union of the state space \( S \sqcup S' \) and view them as a single automaton.

Definition 4 An 1-bounded pseudometric space is a pair \((X, d_X)\) consisting of a set \( X \) and a distance function \( d_X : X \times X \rightarrow [0,1] \) s.t. (1) for all \( x \in X \), \( d_X(x,x) = 0 \); (2) for all \( x, y \in X \), \( d_X(x,y) = d_X(y,x) \); and (3) for all \( x, y, z \in X \), \( d_X(x,z) \leq d_X(x,y) + d_X(y,z) \). As a convention, we often write \( X \) instead of \((X, d_X)\) and we denote the distance function of a metric space \( X \) by \( d_X \).

In this paper, we focus on the behavioral pseudometric which does not discount the future. We characterize the pseudometric as the greatest fixpoint of a functional from a complete lattice to itself. This characterization can be viewed as a quantitative analogue of the greatest fixpoint characterization of bisimilarity.

Definition 5 Let \( \mathcal{M} \) be the class of 1-bounded pseudometric on state set \( S \). The order \( \sqsubseteq \) on \( \mathcal{M} \) is defined by \( d_1 \sqsubseteq d_2 \) if for all \( s, t \in S \), \( d_1(s,t) \geq d_2(s,t) \).
Remark 1 Note the reverse direction of $\sqsubseteq$ and $\geq$ in the above definition. This is to make $d$ the greatest fixpoint, in analogy with the characterization of bisimilarity, rather than a least fixpoint.

Lemma 6 The set of 1-bounded pseudometric over $S$ endowed with the order $\sqsubseteq$ forms a complete lattice. Formally, $(\mathcal{M}, \sqsubseteq)$ is a complete lattice.

Our goal is to introduce a functional from the complete lattice $(\mathcal{M}, \sqsubseteq)$ to itself of which the behavioral pseudometric $d_S$ is the greatest fixpoint. For this purpose, first we have to lift each metric to be a metric on distributions, namely, we need to endow a metric to the distribution $\mathcal{D}$ on sets of states, since in probabilistic automata, the transitions are generally from state to distribution.

It turns out that the classical Hutchinson metric on probabilistic distributions suffices.

Definition 7 (Hutchinson Metric) Given a metric space $(S, d_S)$, we lift it to be a metric over $\mathcal{D}(S)$. Assuming $\eta, \eta' \in \mathcal{D}(S)$, we define $\hat{d}(\eta, \eta')$ as the solution of the following linear program:

\[
\begin{align*}
\text{maximize} & \quad \sum_{s \in S}(\eta(s) - \eta'(s)) \cdot x_s \\
\text{s.t.} & \quad \text{for any } s \in S, 0 \leq x_s \leq 1 \\
& \quad \text{for any } s, t \in S, x_s - x_t \leq d(s, t)
\end{align*}
\]

Remark 2 As mentioned in Section 1, here, we follow the nondiscounted version of pseudometric definition. An alternative one, i.e. the discounted version, which scales the above $d(\eta, \eta')$ by a factor $\gamma \in (0, 1)$, can be found in [8].

The following lemma shows that this extension to distributions satisfies the triangle inequality and is consistent with the ordering on pseudometrics. From the first conclusion, it is not difficult to show that $\hat{d}$ is indeed a pseudometric on $\mathcal{D}(S)$.

Lemma 8 • Let $d \in \mathcal{M}$. Then for any $\eta_1, \eta_2, \eta_3 \in \mathcal{D}(S)$, $\hat{d}(\eta_1, \eta_3) \leq \hat{d}(\eta_1, \eta_2) + \hat{d}(\eta_2, \eta_3)$;
• Let $d_1 \sqsubseteq d_2$. Then for any $\eta, \eta' \in \mathcal{D}(S)$, $d_1(\eta, \eta') \geq d_2(\eta, \eta')$.

We now are in a position to define a monotonic transformation (i.e. functional) on $\mathcal{M}$. First let us recall the definition of Hausdorff distance.

Definition 9 (Hausdorff Distance) Given a 1-bounded pseudometric on $Z$, the Hausdorff distance between two subsets $X, Y \subseteq Z$ is given as follows:

\[
H_d(X, Y) = \max\{\sup_{x \in X} d(x, y), \sup_{y \in Y} d(y, x)\}
\]

where we set $\inf\emptyset = 1$ and $\sup\emptyset = 0$.

As the next step, we define a functional $\Delta$ on $\mathcal{M}$ based on the Hausdorff distance. Recall that $I(s, a) = \{\eta \mid s \xrightarrow{a} \eta\}$.

Definition 10 (Functional $\Delta$) Let $d$ be a 1-bounded pseudometric on $S$. The distance function $\Delta(d): S \times S \rightarrow [0, 1]$ is defined by

\[
\Delta(d)(s, t) = \max_{a \in A}\{H_d(I(s, a), I(t, a))\}
\]

It is not difficult to see that $\Delta$ is well-defined. To ensure the existence of the greatest fixpoint, it suffices to show that $\Delta$ is monotonic. The proof is straightforward and thus omitted here.

Lemma 11 $\Delta$ is monotonic on $\mathcal{M}$.

According to the remarkable Knaster-Tarski theorem, the fixpoints of a monotonic functional on a complete lattice form a complete lattice and hence, the functional has a greatest and least fixpoint. In the following, we denote the greatest fixpoint of $\Delta$ by $\text{gfp}(\Delta)$.

Definition 12 We define $d_S$ as the greatest fixpoint of $\Delta$, formally

\[
d_S \overset{\text{def}}{=} \text{gfp}(\Delta)
\]

To justify the soundness of the pseudometric definition, we have to establish the correspondence between the behavioral pseudometrics and probabilistic bisimulation (c.f. Definition 3): the distance zero captures probabilistic bisimilarity, which is stated by the following theorem formally. Since the proof is similar to the one in [6], we omit it here.

Theorem 13 $s \xleftrightarrow{a} t$ if and only if $d(s, t) = 0$.

4 Approximation Algorithm

In this section, we present the main ingredients of our approximation algorithm. First, we have to provide some technical definitions.

Definition 14 $d$ is a post-fixpoint of $\Delta$ if $d \sqsubseteq \Delta(d)$.

We give an explicit characterization of post-fixpoint.

Lemma 15 $d$ is a post-fixpoint of $\Delta$ if and only if for any action $a \in A$:

• if $s \xrightarrow{a} \eta$, then there exists some $\eta'$ such that $t \xrightarrow{a} \eta'$ and $d(\eta, \eta') \leq d(s, t)$;
• if $t \xrightarrow{a} \eta'$, then there exists some $\eta$ such that $s \xrightarrow{a} \eta$ and $d(\eta, \eta') \leq d(s, t)$;
Clearly, a fixpoint is also a post-fixpoint. Consequently, we have the following characterization. Note here $\sqcup$ returns the greatest lower bound (a.k.a. infimum, meet) of the subset.

**Lemma 16**

$$\text{gfp}(\Delta) = \bigcup \{d \in M \mid d \sqsubseteq \Delta(d)\}$$

Having completed some technical preparations, now we devote ourselves to expressing the fact “$d$ is pseudometric on state space $S$ according to Definition 12” in the first-order (elementary) theory of real closed fields $\mathbb{F}$. Most of shortcuts are standard. We note that when we write $d$, we mean implicitly a vector $d_{s,t}$, where $s, t$ range over $X$. This also applies to the case of distributions with finite support.

In order to make our technical developments streamline, we introduce a series of predicates, which lead to the encoding of a pseudometric being a post-fixpoint as the elementary sentence over real closed fields.

- The fact that $d$ is a 1-bounded pseudometric can be captured as follows:

  $$\text{pseudo}(d) \equiv \\land_{s,t \in S} 0 \leq d_{s,t} \leq 1 \land \land_{s \in S} d_{s,s} = 0 \land \land_{s,t \in S} d_{s,t} = d_{t,s} \land \land_{s,t,u \in S} d_{s,u} \leq d_{s,t} + d_{t,u}$$

- Given two probabilistic distributions $\eta, \eta' \in \mathcal{D}(S)$, where $S$ is finite, we define the predicate $\text{hd}(y, d, \eta, \eta')$ stating the fact that $y$ is the Hutchinson metric (c.f. Definition 7) of $\eta$ and $\eta'$ w.r.t. the metric $d$ on $S$, formally $y = d(\eta, \eta')$ as follows.

  As an auxiliary predicate, we first propose the following predicate $\ell p(y, d, \eta, \eta')$ which encodes the constraints in the linear programming. Note here we let set $X = \{x_s \mid s \in S\}$.

  $$\ell p(y, d, \eta, \eta') \equiv \\exists \hat{X}. (y = \sum_{s \in S} (\eta(s) - \eta'(s)) \cdot x_s) \land \land_{s \in S} 0 \leq x_s \leq 1 \land \land_{s,t \in S} x_s - x_t \leq d(s,t)$$

  It follows the definition of $\text{hd}(y, d, \eta, \eta')$:

  $$\text{hd}(y, d, \eta, \eta') \equiv \ell p(y, d, \eta, \eta') \land \forall z. (\ell p(z, d, \eta, \eta') \implies y \geq z)$$

- We proceed to define the predicate regarding the Hausdorff distance (c.f. Definition 9). Given a pseudometric $d$ on $S$, two states $s, t \in S$, an action $a \in A$ and a distribution $\eta$ such that $s \xrightarrow{a} \eta$, we define, under the condition that $\mathcal{I}(t,a) \neq \emptyset$, that:

  $$\text{inf}(y, d, a, s, t, \eta) \equiv \bigvee_{\eta' \in \mathcal{I}(t,a)} \text{hd}(y, d, \eta, \eta') \land \forall z. (\forall \eta' \in \mathcal{I}(t,a) \text{hd}(z, d, \eta, \eta') \implies y \leq z)$$

  It follows that we define, under the condition that $\mathcal{I}(s,a) \neq \emptyset$ that:

  $$\sup \inf(y, d, a, s, t, \eta) \equiv \bigvee_{\eta \in \mathcal{I}(s,a)} \text{inf}(y, d, a, s, t, \eta) \land \forall z. (\forall \eta \in \mathcal{I}(s,a) \text{inf}(z, d, a, s, t, \eta) \implies y \geq z)$$

- The fact that $y$ is the distance w.r.t. a 1-bounded pseudometric on distributions, under the constraint that $\mathcal{I}(s,a) \neq \emptyset$ and $\mathcal{I}(t,a) \neq \emptyset$ can be captured as follows:

  $$\text{haus}(y, d, a, s, t) \equiv \sup \inf(y, d, a, s, t) \land \sup \inf(y, d, a, t, s) \land \forall z. (\sup \inf(z, d, a, s, t) \land \sup \inf(z, d, a, t, s) \implies y \geq z)$$

- In view of Lemma 15, to define $d$ is a post-fixpoint w.r.t. states $s$ and $t$, we have to distinguish three cases:

  (1) For any $a \in A$, $\mathcal{I}(s,a) \neq \emptyset \iff \mathcal{I}(t,a) \neq \emptyset$ and there exists some $a$, $\mathcal{I}(s,a) \neq \emptyset$.

    $$\text{postfixpoint}_1(d, s, t) \equiv \forall a \in A. \mathcal{I}(s,a) = \emptyset \iff \mathcal{I}(t,a) = \emptyset$$

    $$\land \exists y. \bigvee_{a \in A} \mathcal{I}(s,a) \neq \emptyset \text{haus}(y, a, d, s, t) \land \forall z. \bigvee_{a \in A} \mathcal{I}(s,a) \neq \emptyset \text{haus}(z, a, d, s, t) \implies y \geq z$$

  (2) For any $a \in A$, $\mathcal{I}(s,a) \neq \emptyset \iff \mathcal{I}(t,a) \neq \emptyset$ and for all $a \in A$, $\mathcal{I}(s,a) = \emptyset$.

    $$\text{postfixpoint}_2(d, s, t) \equiv \forall a \in A. \mathcal{I}(s,a) = \emptyset \land \mathcal{I}(t,a) = \emptyset \land d_{s,t} = 0$$

  (3) There exists some $a$ such that $\mathcal{I}(s,a) = \emptyset \land \mathcal{I}(t,a) \neq \emptyset$.

    $$\text{postfixpoint}_3(d, s, t) \equiv \exists a \in A. \neg(\mathcal{I}(s,a) = \emptyset \land \mathcal{I}(t,a) = \emptyset) \land d_{s,t} = 1$$

  We note the above three cases (1)(2)(3) are mutual exclusive. To combine them together, we obtain:

  $$\text{postfixpoint}(d, s, t) \equiv \text{postfixpoint}_1(d, s, t) \lor \text{postfixpoint}_2(d, s, t) \lor \text{postfixpoint}_3(d, s, t)$$

  Here we note that the predicate concerning $\mathcal{I}(s,a)$ and $\mathcal{I}(t,a)$ can be instantiated to $\text{true}$ or $\text{false}$ when the concrete SPA is considered.

- It follows that

  $$\text{postfixpoint}(d) \equiv \land_{s,t \in S} \text{postfixpoint}_1(d, s, t) \lor \text{postfixpoint}_2(d, s, t) \lor \text{postfixpoint}_3(d, s, t)$$

  This completes our constructions.
According to Lemma 15, it is not difficult to see that the following lemma holds:

**Lemma 17** Assume any probabilistic automaton \( \mathcal{P} \). Then postfixpoint \((d)\) holds iff \( d \) is a post-fizpoint of \( \Delta \) given in Definition 10.

Let us fix a probabilistic automata \( \mathcal{P} = (S, A, \rightarrow) \), two states \( s, t \in S \) and \( \varepsilon \) as the desired accuracy. Recall that our goal is to find an interval \([\ell, u]\) such that \( u - \ell \leq \varepsilon \) and \( d_S(s, t) \in [\ell, u] \). With the predicate postfixpoint \((d)\) at hand, the algorithm that approximates the distance within a tolerance of \( \varepsilon \) can be obtained by a binary search. Due to space constraint, we omit the detailed presentation. The crucial point is to check whether a given \( d \) is a pseudometric on state space according to \( \Delta \) given in Definition 12, which is encoded by the formula \( \exists! \text{postfixpoint}(d) \wedge \text{pseudo}(d) \). In addition, we note that the length of the formula is quadratic in the size of a given SPA. In addition, the number of quantifier alternations is a constant in this formula. The results of [1] shows that quantifier elimination in the theory of real closed fields over addition and multiplication can be achieved in time exponential in the size of the formula and double exponential in the number of quantifier alternations. Thus we obtain the \text{EXPTIME} upper complexity bound. To conclude, we obtain:

**Theorem 18** Given a simple probabilistic automaton \( \mathcal{P} \), two states \( s, t \), the pseudometric distance can be approximated up to any \( \varepsilon > 0 \) in time exponential in the size of \( \mathcal{P} \) and logarithmic in \( \frac{1}{\varepsilon} \).

## 5 Related Works

In this section, we only can mention the most relevant works in brief. [8] studied a logical pseudometric for labelled Markov chains. A similar pseudometric was defined by van Breugel and Worrell [4] via the terminal coalgebra of a functor based on a metric on the space of Borel probability measures. In [7] Desharnais et al. dealt with labelled concurrent Markov chains (this model can be captured by our model). [6] considered a more general framework, called \textit{action-labeled quantitative transition systems} (AQTS). The definition of pseudometric studied in this paper does not deviate very far from this line of research in the sense it can be viewed as an instantiate the (abstract) pseudometric definition given in [6] in the setting of simple probabilistic automata. For the algorithmic aspect, [3] and [8] both provided algorithms when the metric discounts the future, which simply do not work in our setting. More recently, [2] independently proposed an algorithm when the future is not discounted. Their algorithm is very similar to us. However, they only considered the fully probabilistic model while the model considered in this paper is much more general.

## References


