Computers and Mathematics with Applications 57 (2009) 991–999

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

On metrics for probabilistic systems: Definitions and algorithms

Taolue Chen [a,*], Tingting Han [b,c], Jian Lu [d]

[a] CWI, Department of Software Engineering, PO Box 94079, 1090 GB Amsterdam, The Netherlands

[b] RWTH Aachen, MOVES, Ahornstraße 55, D-52056 Aachen, Germany

[c] University of Twente, Faculty of EEMCS, Formal Methods and Tools, PO Box 217, 7500 AE Enschede, The Netherlands

[d] Nanjing University, State Key Laboratory of Novel Software Technology, Nanjing, Jiangsu, 210093, PR China

ARTICLE INFO

Keywords:
Probabilistic systems
Simple probabilistic automata
Behavioral equivalence
Metric
Algorithm

ABSTRACT

In this paper, we consider the behavioral pseudometrics for probabilistic systems, which are a quantitative analogue of probabilistic bisimilarity in the sense that the distance zero captures the probabilistic bisimilarity. 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 $\frac{1}{\varepsilon}$. One of the key ingredients 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 the binary search to approximate the behavioral distance.

1. Introduction

Probability, like nondeterminism, is an abstraction mechanism used to hide inessential or unknown details. Statistical mechanics, originated by Boltzmann, Gibbs, Maxwell et al., is a celebrated successful example of using the probabilistic abstraction. The probability abstraction can be exploited, among others, to model a certain chance of error or other stochastic behavior occurring in various real world applications, thus has been extensively applied in many areas, in particular, natural computation, fuzzy systems and knowledge discovery. This motivates the investigation of *probabilistic systems*, where system dynamics encodes the probability of making a transition between states rather than just the existence of such transitions. To support the specification and analysis of probabilistic systems, numerous models have been proposed in the literature and are study subjects of a rapidly growing research community, for which we refer the readers to [2] for a comprehensive exposition.

In [12] van Glabbeek et al. classified probabilistic models into *reactive*, *generative* and *stratified*. Albeit the diversity of these models, Segala [17] argued convincingly that none of them captures real *nondeterminism*, an essential notion for modeling, say, scheduling freedom, implementation freedom, the external environment and incomplete information. To remedy this, he then introduced a new family of models, namely the *simple probabilistic automata* and *probabilistic automata* (SPA and PA in short, respectively), where both *probability* and *nondeterminism* are taken into account. In a nutshell, (S)PA

* Corresponding author.
E-mail addresses: chen@cwi.nl (T. Chen), tingting.han@cs.rwth-aachen.de (T. Han), lj@nju.edu.cn (J. Lu).

0898-1221/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2008.10.041

constitutes a very expressive framework for the specification and analysis of probabilistic systems. They are based on state transition systems and make a clear distinction between probabilistic and nondeterministic choices. Technically speaking, SPAs relate a state and an action with a distribution over target states while PAs relate a state with a distribution over actions and target states. In contrast, probabilistic models without nondeterminism are sometimes referred to as *purely probabilistic*, for example, the discrete-time Markov chains (DTMCs) and a more "traditional" probabilistic automaton due to Rabin [16]. We refer the reader to [18] for a leisure exposition and comparison.

In this paper, we concentrate on SPAs. Admittedly, SPA is a special case of PA. However, problems arise when defining the parallel composition operators for PAs, which prevents us from potential modular (compositional) analysis. Since compositionality is an extremely desired property, we prefer SPA. Moreover, we mention that these two models are equivalent in some sense and there is no real obstacle to extend our result from SPA to PA. For more discussions, we refer the readers to the relevant part of [2].

In system theory, usually one models the systems and analyzes their properties. For any model, one of the fundamental research questions is the notions of *equivalence* and *congruence*: when can two systems be deemed the same and when can they be inter-substituted for each other? In the classical investigations in concurrency theory, *bisimulation*, introduced by Park and Milner [14], is a ubiquitous notion of equivalence which has become one of the primary tools in the analysis of systems: when two systems are bisimilar, known properties are readily transferred from one system to the other. In probabilistic systems, the standard notion of bisimulation has to be adapted, usually by treating the probability as labels. This line of research, to the best of our knowledge, dates back to Larsen and Skou's work on pure probabilistic systems [13] and now is very fruitful, see, among others, [13,17,15].

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 (e.g. see [10]). To see this, first let us note that the probabilities appearing in models cannot be viewed as exact numbers, instead they should be read as numbers with some error estimate. Hence reasoning principles based on the exact value of numbers are of dubious practical value and thus it is unfortunately problematic if the notion of equivalence of probabilistic systems is sensitive to the exact probabilities of transitions, since a slight change in the transition probabilities will result in that two equivalent states which are deemed no longer equivalent. This instability is due to the quantitative nature of probabilistic systems. In summary, in a situation where the system behavior has a quantitative aspect it should come up with a more robust (or "fuzzy") approach to equivalence.

To find a more flexible way to distinguish system states, researchers in this area 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. A couple of papers have addressed this problem and to make our presentation smooth, we defer the discussion of related works to Section 5.

Having a nice pseudometric definition for systems at hand, the next natural question is: how to *compute* it? This raises some algorithmic challenges. It turns out that this problem is relevant to the technical definitions of pseudometrics. Desharnasi et al. [10] introduced a family of behavioral pseudometric for probabilistic transition systems. Each pseudometric is parameterized by a discount factor γ, a real number in $(0, 1]$. The smaller the discount factor, the more (behavioral differences in) the future is discounted. If the discount factor is 1, then the future is *not* discounted at all. In this case, the differences in behavior, whether in the near of far future, contribute alike to the distance. For systems that (in principle) run forever, e.g. the reactive systems, we may be interested in all these differences and, hence, in the pseudometric that does *not* discount the future [4]. In [10], Desharnais et al. have presented an algorithm to approximate the behavioral distances in the *discounted* setting. Van Breugel and Worrell also proposed an approximation algorithm in the same setting. However, there is a fundamental difference between pseudometric that discounts the future and the one that does not. For example, from the topology-theoretic point of view: all pseudometrics that discount the future give rise to the same topology, whereas the pseudometric that does *not* discount the future gives rise to a different topology [10]. As a consequence, although there are a couple of (approximation) algorithms to compute the behavioral distance when $\gamma < 1$, none of them can be adapted (at least not in an obvious way) to handle the case that $\gamma = 1$.

Our contributions. The main contribution of this paper is two-fold: first, we instantiate the (abstract) pseudometric definition given in [8] for a general quantitative model in the setting of simple probabilistic automata and provide a concrete account; second, we present an approximation algorithm to *compute* the behavioral distance. Since the pseudometric we consider here does not discount the future, some novel approach other than the ones presented in [10] or [5] has to be exploited. Our main technique is the characterization of distance as semi-algebraic quantities. 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 very natural and provide an algorithm and thus a complexity upper bound on the problem of approximating the distance between two states in a very general model of probabilistic systems.

[1] In this paper, as in [8], 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.

Structure of the paper. This paper is set up as follows. In Section 2, we present some background knowledge, including the probabilistic automata with traditional bisimulation definition and brief introduction on real closed fields. In Section 3, we provide the pseudometric definition and discuss its properties. In Section 4, an approximation algorithm is proposed and we show its correctness. We conclude our work in Section 5 where some related works, as well as the future works are also discussed in brief. An extended abstract of the current paper has appeared as [7].

2. Preliminaries

Before starting our exposition, let us first fix some general notations. Throughout the paper, we assume a fixed $\mathbb{F}$ of some real closed ordered field. An ordered field $\mathbb{F}$ is *real closed* if no proper algebraic extension of $\mathbb{F}$ is ordered. Examples of real closed fields include 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 $\delta : X \mapsto [0, 1] \cap \mathbb{F}$ such that $\sum_{x\in X} \delta(x) = 1$. We denote the set of probability distributions on X by $\mathcal{D}(X)$. For a probability distribution $\delta \in \mathcal{D}(X)$ we define $\|\delta\|$, the *support* of δ, as $\|\delta\| = \{x \in X \mid \delta(x) > 0\}$. Note here we impose a restriction to each distribution η such that $\eta(x) \in \mathbb{F}$. Namely, for instance, if one sets $\mathbb{F}$ as the set of real algebraic numbers, then the probability can not be, say, $\frac{\pi}{4}$, which is transcendental.

2.1. Simple probabilistic automata

Definition 1 (*Simple Probabilistic Automata, [17]*). A *simple probabilistic automaton* is a tuple $\mathcal{P} = (S, A, \rightarrow)$ where

- S is a *finite* set of states;
- A is a *finite* set of actions; and
- $\rightarrow \subseteq S \times A \times \mathcal{D}(S)$ is the transition relation.

We shall write $s \xrightarrow{a} \eta$ as a more suggestive notion instead of $(s, a, \eta) \in \rightarrow$.

Note that it is possible that from state s, there are more than one transition labeled by a which lead to different distributions (in terms of concurrency, SPAs feature internal nondeterminism). Henceforth for each state s and action a, we write $\mathit{l}(s, a)$ for $\{\eta \in \mathcal{D}(S) \mid s \xrightarrow{a} \eta\}$. The *size* of SPA is the sum of the number of its states and transitions.

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 $\eta(V)$ for $\sum_{s\in V} \eta(s)$. We first lift an equivalence relation on S to an equivalence relation between *distributions over* S in the following way:

Definition 2. Let $\eta, \eta' \in \mathcal{D}(S)$, we say that they are equivalent w.r.t. an equivalence $\mathcal{R}$ on S, written $\eta \equiv_{\mathcal{R}} \eta'$ if

$$\forall U \in S/\mathcal{R}.\ \eta(U) = \eta'(U).$$

Definition 3 (*Probabilistic Bisimulation*). An equivalence relation $\mathcal{R} \subseteq S \times S$ is a (strong) *probabilistic bisimulation* if $s\mathcal{R}t$ implies:

whenever $s \xrightarrow{a} \eta$, there exists η' such that $t \xrightarrow{a} \eta'$ and $\eta \equiv_{\mathcal{R}} \eta'$.

Two states s, t are *probabilistic bisimilar*, denoted by $s \leftrightarrow t$, if there exists some bisimulation $\mathcal{R}$ s.t. $s\mathcal{R}t$.

2.2. Theory of real closed fields

Our main technique is to represent the value as an elementary formula in the theory of real closed fields, and uses a decision procedure for the theory of reals with addition and multiplication dating back to [19]. To facilitate the understanding of the algorithm in the sequel, here we include a brief introduction on this aspect.

We denote by $\mathbf{F}$ the real closed field $(\mathbb{F}, +, \cdot, 0, 1, \leq)$ with addition and multiplication. An *atomic formula* a is an expression of the form $p > 0$ or $p = 0$ where p is a (possibly) multi-variate polynomial with coefficients in $\mathbb{F}$. An *elementary formula* is constructed from atomic formulae by the grammar:

$$\phi ::= a \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \exists x.\phi \mid \quad \forall x.\phi,$$

where a is an atomic formula. The semantics of an elementary formula is given in a standard way. A variable x is *free* in the formula ϕ if it is not in the scope of a quantifier $\exists x$ or $\forall x$. An *elementary sentence* is a formula without free variables.

It is well-known that the elementary (i.e., first-order) theory of real closed fields in the language of ordered fields is decidable, namely, we have:

Theorem 4 (*[19]*). *The theory of real closed fields in the language of ordered fields is decidable.*

Basically, Tarski's theorem [19] tells us that the theory of real closed fields, including the predicate symbol $<$, admits elimination of quantifiers, which in turn entails it is a complete and decidable theory. Furthermore the results of [1] show that quantifier elimination in the theory of reals 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.

3. Behavioral metrics

In this section, we define pseudometric as the greatest fixpoint of a certain functional. Note that most of material of this section is adapted from [8]. Let us fix a simple probabilistic automata $\mathcal{P} = (S, A, \rightarrow)$ and consider pseudometrics on its set of states S. We note that this metric suffices even if one wants to compute the distance between the states in two different SPAs, say $\mathcal{P}$ and $\mathcal{P}'$, since we can simply take the disjoint union of the state space $S \uplus S'$ and view them as a single automaton.

Definition 5. A *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 6. 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 that as in [3], we make the reverse direction of $\sqsubseteq$ and $\geq$ in the above definition on purpose. This is used to make d the *greatest* fixpoint, in analogy with the characterization of bisimilarity, rather than the *least* fixpoint.

Lemma 7. *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.*

Proof. Standard. It suffices to note that meets (g.l.b.) are suprema and joins (l.u.b.) are infima. Namely for any $X \subseteq \mathcal{M}$, $(\bigsqcap X)(s, t) = \sup\{d(s, t) \mid d \in X\}$. Dually, $(\bigsqcup X)(s, t) = \inf\{d(s, t) \mid d \in X\}$. □

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 simple probabilistic automata, the transitions are generally from state to distribution.

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

Definition 8 (*Hutchinson Metric*). Given a metric space (S, d), 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:

$$\begin{array}{ll} \text{maximize} & \sum_{s \in S} \big(\eta(s) - \eta'(s)\big) \cdot x_s \\ \text{s.t.} & \text{for any } s \in S,\ 0 \leq x_s \leq 1 \\ & \text{for any } s, t \in S,\ x_s - x_t \leq d(s, t). \end{array}$$

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 $\hat{d}(\eta, \eta')$ by a factor $\gamma \in (0, 1)$, can be found in [10].

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 9. *Let* $d, d_1, d_2 \in \mathcal{M}$ *and* $d_1 \sqsubseteq d_2$*. It holds that:*

- *For any* $\eta_1, \eta_2, \eta_3 \in \mathcal{D}(S)$, $\hat{d}(\eta_1, \eta_3) \leq \hat{d}(\eta_1, \eta_2) + \hat{d}(\eta_2, \eta_3)$*;*
- *For any* $\eta, \eta' \in \mathcal{D}(S)$, $\hat{d}_1(\eta, \eta') \geq \hat{d}_2(\eta, \eta')$.

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

Definition 10 (*Hausdorff Distance*). Given a 1-bounded pseudometric on Z, the *Hausdorff distance* between two subsets $X, Y \subseteq Z$ is given as follows:

$$H_d(X, Y) = \max\left\{\sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(y, x)\right\}$$

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

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

Definition 11 (*Functional* Δ). Let d be a 1-bounded pseudometric on S. The distance function $\Delta(d) : S \times S \to [0, 1]$ is defined by:

$$\Delta(d)(s, t) = \max_{a \in A} \left\{ H_{\hat{d}} \left(\mathcal{U}(s, a), \mathcal{U}(t, a) \right) \right\}.$$

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

Lemma 12. *Δ is monotonic on $\mathcal{M}$.*

According to the remarkable Knaster–Tarski theorem [20], the fixpoints of a monotonic functional on a complete lattice form a complete lattice and hence, the functional Δ has a greatest and least fixpoint. In the following, we denote the *greatest fixpoint* of Δ by $\mathsf{gfp}(\Delta)$.

Definition 13. We define d_S as the greatest fixpoint of Δ, formally:

$$d_S \stackrel{\text{def}}{=} \mathsf{gfp}(\Delta).$$

To justify the soundness of the pseudometic 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 [8], we omit it here.

Theorem 14. *Given a simple probabilistic automaton $\mathcal{P} = (S, A, \to)$. For any two states $s, t \in S$, $s \leftrightarrow t$ if and only if $d_S(s, t) = 0$.*

4. Approximation algorithm

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

Definition 15. d is a post-fixpoint of Δ if $d \sqsubseteq \Delta(d)$.

We give an explicit characterization of post-fixpoint.

Lemma 16. *d is a post-fixpoint of Δ if and only if for any action $a \in A$:*

- *If $s \xrightarrow{a} \eta$, then there exists some η' such that $t \xrightarrow{a} \eta'$ and $\hat{d}(\eta, \eta') \leq d(s, t)$;*
- *If $t \xrightarrow{a} \eta'$, then there exists some η such that $s \xrightarrow{a} \eta$ and $\hat{d}(\eta, \eta') \leq d(s, t)$;*

Proof. By definition. □

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 a set.

Lemma 17.

$$\mathsf{gfp}(\Delta) = \bigsqcup \{d \in \mathcal{M} \mid d \sqsubseteq \Delta(d)\}.$$

Having done some technical preparations, now we devote ourselves to expressing the fact "d is pseudometric on state space S according to Definition 13" in the first-order (elementary) theory of real closed fields $\mathbb{F}$. For this purpose, we first introduce some meta notations to make the predicates accessible.

4.1. Meta notations

- Given any *finite* set $X = \{x_1, \cdots, x_n\}$, we denote:
$$\bigwedge_{x \in X} \varphi(x) = \varphi(x_1) \wedge \cdots \wedge \varphi(x_n).$$
The same notation applies to disjunction $\bigvee$.
- Assuming a *finite* metric space (X, d), when we write d, we mean implicitly a vector $\vec{d_{s,t}}$ where s, t range over X (note that s and t are *not* necessarily different). This also applies to the case of probabilistic distribution which has a *finite* support. Namely, given η with finite $\|\eta\|$, we write η for $\vec{\eta}_s$ where $s \in \|\eta\|$.
- For any *finite* set $X = \{x_1, \cdots, x_n\}$, we write $\forall \vec{X}.\varphi$ for $\forall x_1 \cdots \forall x_n.\varphi$. The same notation applies to existential quantification $\exists$.

4.2. Predicates

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:

$$\begin{aligned}&\mathsf{pseudo}(d) \equiv\\ &\bigwedge_{s,t\in S} 0 \le d_{s,t} \le 1 \ \wedge \ \bigwedge_{s\in S} d_{s,s} = 0 \ \wedge\\ &\bigwedge_{s,t\in S} d_{s,t} = d_{t,s} \ \wedge \ \bigwedge_{s,t,u\in S} d_{s,u} \le d_{s,t} + d_{t,u}.\end{aligned}$$

- Given two probabilistic distribution $\eta, \eta' \in \mathcal{D}(S)$, where S is finite, we define the predicate $\hbar d(y, d, \eta, \eta')$ stating the fact that y is the Hutchinson metric (c.f. Definition 8) of η and η' w.r.t. the metric d on S, formally $y = \hat{d}(\eta, \eta')$ in the following way:

  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 let us set $X = \{x_s \mid s \in S\}$.

$$\begin{aligned}&\ell p(y, d, \eta, \eta') \equiv\\ &\exists \vec{X}.(y = \sum_{s\in S}(\eta(s) - \eta'(s)) \cdot x_s)\\ &\wedge \bigwedge_{s\in S} 0 \le x_s \le 1 \wedge \bigwedge_{s,t\in S} x_s - x_t \le d(s, t).\end{aligned}$$

  We remark that since S is finite, this formula is a first-order sentence.

  It follows the definition of $\hbar d(y, d, \eta, \eta')$:

$$\begin{aligned}&\hbar d(y, d, \eta, \eta') \equiv\\ &\ell p(y, d, \eta, \eta') \wedge \forall z.\left(\ell p(z, d, \eta, \eta') \Longrightarrow y \ge z\right).\end{aligned}$$

- We proceed to define the predicate regarding the Hausdorff distance (c.f. Definition 10). Given a pseudometric d on S, two states $s, t \in S$, an action $a \in A$ and a distribution η such that $s \xrightarrow{a} \eta$, we define, under the condition that $\mathcal{A}(t, a) \neq \emptyset$, that:

$$\begin{aligned}&\mathbf{inf}(y, d, a, s, t, \eta) \equiv\\ &\bigvee_{\eta'\in\mathcal{A}(t,a)} \hbar d(y, d, \eta, \eta') \ \wedge \ \forall z.\left(\bigvee_{\eta'\in\mathcal{A}(t,a)} \hbar d(z, d, \eta, \eta') \Longrightarrow y \le z\right).\end{aligned}$$

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

$$\begin{aligned}&\sup\inf(y, d, a, s, t) \equiv\\ &\bigvee_{\eta\in\mathcal{A}(s,a)} \inf(y, d, a, s, t, \eta) \ \wedge \ \forall z.\left(\bigvee_{\eta\in\mathcal{A}(s,a)} \inf(z, d, a, s, t, \eta) \Longrightarrow y \ge z\right).\end{aligned}$$

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

$$\begin{aligned}&\mathsf{haus}(y, d, a, s, t) \equiv\\ &\sup\inf(y, d, a, s, t) \ \wedge \ \sup\inf(y, d, a, t, s)\\ &\wedge \forall z.\,(\sup\inf(z, d, a, s, t) \wedge \sup\inf(z, d, a, t, s) \Longrightarrow y \ge z)\,.\end{aligned}$$

- In view of Lemma 16, to define d is a post-fixpoint w.r.t. states s and t, we have to distinguish three cases:
  1. For any $a \in A$, $\mathcal{A}(s, a) \neq \emptyset \Leftrightarrow \mathcal{A}(t, a) \neq \emptyset$ and there exists some a, $\mathcal{A}(s, a) \neq \emptyset$.

$$\begin{aligned}&\mathsf{postfixpoint}_1(d, s, t) \equiv\\ &\forall a \in A.\ \mathcal{A}(s, a) = \emptyset \Leftrightarrow \mathcal{A}(t, a) = \emptyset\\ &\wedge \exists y. \bigvee_{\{a\in A \mid \mathcal{A}(s,a)\neq\emptyset\}} \mathsf{haus}(y, a, d, s, t)\\ &\wedge \forall z. \bigvee_{\{a\in A \mid \mathcal{A}(s,a)\neq\emptyset\}} \mathsf{haus}(z, a, d, s, t)\\ &\qquad\qquad \Longrightarrow y \ge z.\end{aligned}$$

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

$$\begin{aligned}&\mathsf{postfixpoint}_2(d, s, t) \equiv\\ &\forall a \in A.\ \mathcal{A}(s, a) = \emptyset \ \wedge \ \mathcal{A}(t, a) = \emptyset \ \wedge \ d_{s,t} = 0.\end{aligned}$$

3. There exists some a such that $\mathbb{1}(s, a) = \emptyset \not\Leftrightarrow \mathbb{1}(t, a) \neq \emptyset$.

$$\begin{aligned}&\mathsf{postfixpoint}_3(d, s, t) \equiv\\ &\quad \exists a \in A.\ \neg\,(\mathbb{1}(s, a) = \emptyset \Leftrightarrow \mathbb{1}(t, a) = \emptyset)\ \wedge\ d_{s,t} = 1.\end{aligned}$$

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

$$\begin{aligned}&\mathsf{postfixpoint}(d, s, t) \equiv\\ &\quad \mathsf{postfixpoint}_1(d, s, t) \vee \mathsf{postfixpoint}_2(d, s, t) \vee \mathsf{postfixpoint}_3(d, s, t).\end{aligned}$$

Here we note that the predicate concerning $\mathbb{1}(s, a)$ and $\mathbb{1}(t, a)$ can be instantiated to *true* or *false* when the concrete SPA is considered.

- It follows that

$$\overline{\mathsf{postfixpoint}}(d) \equiv \bigwedge_{s,t \in S} \mathsf{postfixpoint}(d, s, t).$$

According to Lemma 16, it is not difficult to see that the following theorem holds:

Theorem 18. *Assume any simple probabilistic automaton $\mathcal{P} = (S, A, \mathcal{P})$, $\overline{\mathsf{postfixpoint}}(d)$ holds iff d is a post-fixpoint of Δ given in Definition 11.*

4.3. Algorithm

Let us fix a simple probabilistic automata $\mathcal{P} = (S, A, \rightarrow)$, two states $s, t \in S$ and ε as the desired accuracy. Recall that our goal is to find an interval $[\ell, u] \subseteq [0, 1]$ such that $u - \ell \leq \varepsilon$ and $d_S(s, t) \in [\ell, u]$. An algorithm that approximates the distance within a tolerance of ε is obtained by a binary search, see Algorithm 1.

Algorithm 1 Approximating the distance

Require: A simple probabilistic automata $\mathcal{P} = (S, A, \rightarrow)$, and a rational value ε as tolerance, two states s, t in S.
Ensure: An interval $[\ell, u]$ such that $u - l \leq 2\varepsilon$ and $d(s, t) \in [\ell, u]$

```
l := 0, u := 1, m := 1/2;
for ⌈log(1/ε)⌉ steps do
  if ∃d.postfixpoint(d) ∧ pseudo(d) ∧ d(s, t) ≤ m then
    u := m, m := (l+u)/2;
  else
    l := m, m := (l+u)/2;
  end if
end for
return [ℓ, u];
```

We now sketch the correctness of the algorithm.

Proof. The termination of the algorithm is obvious. And clearly $d(s, t) \in [0, 1]$. Note that after each loop, the size of interval $[\ell, u]$ will decrease into half. Therefore, after $\lceil \log(\frac{1}{\varepsilon}) \rceil$ times of the loop, $u - \ell \leq \varepsilon$. Therefore, to see the correctness, it is enough to notice the following two cases, according to the entry condition of the **for** loop.

- $u - \ell \geq \varepsilon$ and $\exists d.\overline{\mathsf{postfixpoint}}(d) \wedge \mathsf{pseudo}(d) \wedge d_{s,t} \leq m$. Then there exists some pseudometric d which is a post-fixpoint of Δ and $d(s, t) \leq m$. According to Lemma 17, $d \sqsubseteq d_S$, namely, $d_S(s, t) \leq d(s, t) \leq m$. Hence $d_S(s, t) \in [\ell, m]$.
- $u - \ell \geq \varepsilon$ and $\neg\exists d.\overline{\mathsf{postfixpoint}}(d)\ \wedge\ \mathsf{pseudo}(d) \wedge d_{s,t} \leq m$. Then for any 1-bounded pseudometric d which is a post-fixpoint of Δ, we have $d(s, t) > m$. Clearly, d_S is a post-fixpoint of Δ. It follows that $d_S(s, t) > m$. Hence $d_S(s, t) \in [m, u]$.

The correctness of the algorithm follows. □

4.4. Complexity

It is easy to see that the length of the formula $\exists d.\mathsf{postfixpoint}(d) \wedge \mathsf{pseudo}(d)$ is quadratic in the size of a given SPA. In addition, the number of quantifier alternations is a *constant* in this formula. As we mentioned before, 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 EXPTIME upper complexity bound. Formally, we can obtain:

Corollary 19. *Given a simple probabilistic automaton $\mathcal{P} = (S, A, \rightarrow)$ and two states $s, t \in S$, the pseudometric distance can be approximated up to any $\varepsilon > 0$ in time exponential in the size of $\mathcal{P}$ and logarithmic in $\frac{1}{\varepsilon}$.*

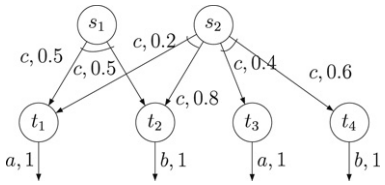

Fig. 1. The simple probabilistic automaton.

4.5. Example

We present a simple example to illustrate our algorithm. The SPA is depicted in Fig. 1, where s_1 has a c transition leading to a distribution μ_1 over t_1, t_2 with $\mu_1(t_1) = \mu_1(t_2) = 0.5$ and s_2 has two transitions both labeled with c which lead to μ_2 and μ_3 respectively, where $\mu_2(t_1) = 0.2$ and $\mu_2(t_2) = 0.8$ and $\mu_3(t_3) = 0.4$ and $\mu_3(t_4) = 0.6$. For t_1, t_2, t_3, t_4, each has a transition labeled either by a or b resulting in a Dirac distribution. Assume the tolerance $\varepsilon = 0.1$, and we intend to compute the distance of s_1 and s_2. Note the main part of the algorithm is the **for** loop (for a binary search) and the decision of $\exists d.\overline{\mathsf{postfixpoint}}(d) \wedge \mathsf{pseudo}(d) \wedge d(s,t) \leq m$. Clearly, the for loop will run $\lceil \log(\frac{1}{0.1}) \rceil = 4$ times.

(1) $m = 0.5$. Then **if** condition $\exists d.\overline{\mathsf{postfixpoint}}(d) \wedge \mathsf{pseudo}(d) \wedge d(s,t) \leq 0.5$ returns TRUE. To see this, we note that the following metric d with typically $d(s_1, s_2) = 0.1, d(t_1, t_3) = 0, d(t_2, t_4) = 0, d(t_1, t_2) = 1, d(t_2, t_3) = 1$ (the distance of other pairs of states are 1) suffices as the witness of this first-order formula. It follows that $l = 0, u = 0.5$ and $m = 0.25$ when loop body is entered in the next time;
(2) $m = 0.25$. As before, $\exists d.\overline{\mathsf{postfixpoint}}(d) \wedge \mathsf{pseudo}(d) \wedge d(s,t) \leq 0.25$ returns TRUE. Hence, $l = 0, u = 0.25$ and $m = 0.125$;
(3) $m = 0.125$. This time $\exists d.\overline{\mathsf{postfixpoint}}(d) \wedge \mathsf{pseudo}(d) \wedge d(s,t) \leq 0.125$ returns FALSE. Hence, $l = 0.125, u = 0.25$ and $m = 0.0625$;
(4) $m = 0.0625$. This time $\exists d.\overline{\mathsf{postfixpoint}}(d) \wedge \mathsf{pseudo}(d) \wedge d(s,t) \leq 0.0625$ returns FALSE. We can conclude the distance of s_1, s_2 falls into the scope $[0.0625, 0.125]$.

In conclusion, the distance is approximated within the tolerance 0.1. If more precise approximation is expected, one can narrow ε.

5. Conclusion

In this paper, we have considered the behavioral pseudometrics for simple probabilistic automata, which are very expressive models including both the nondeterministic and probabilistic choices. The pseudometric is a quantitative analogue of probabilistic bisimilarity and is characterized as the greatest fixpoint of a monotonic functional on the complete lattice of state metrics. We focused on the case that pseudometrics do not discount the future. We have provided an approximation algorithm, which can compute the distance of any two states up to any desired degree of precision ε in time exponential in the size of the model and logarithmic in $\frac{1}{\varepsilon}$. A key point 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 pseudometric.

5.1. Related work

Giacalone et al. [11] were the first to suggest a metric between probabilistic transition systems to formalize the notion of distance between processes. Subsequently, [10] studied a logical pseudometric for labeled Markov chains, which is a reactive model of probabilistic systems. The metric has the property that two processes have distance 0 if and only if they are probabilistic bisimilar. A similar pseudometric was defined by van Breugel and Worrell [6] via the terminal coalgebra of a functor based on a metric on the space of Borel probability measures. In [9] Desharnais et al. dealt with labeled concurrent Markov chains (this model can be captured by our model). They showed that the greatest fixpoint of a monotonous functional on pseudometrics corresponds to the weak probabilistic bisimilarity of [15]. [8] considered a more general framework, called *action-labeled quantitative transition systems* (AQTS). They define a pseudometric which is an adaption of the one in [9]. The definition of pseudometric studied in this paper does not deviate very far from this line of research in the sense that it can be viewed as an instantiation the (abstract) pseudometric definition given in [8] in the setting of simple probabilistic automata. We claim that our algorithm can be extended to AQTS without any difficulty. However, to work on SPA can avoid unnecessary technical details which could obscure the essential points of our algorithm.

For the algorithmic aspect, as we have mentioned, [5] and [10] both provided algorithms when the metric does discount the future. However, they can not be applied for the metric defined in the current paper. Furthermore, it is worth mentioning that after the conference version of current paper [7] was finished, we are notified that [4] independently proposed an algorithm when the future is not discounted. The authors use basically the same technique to deal with this problem and

thus obtain a very similar algorithm. However, they only considered the fully probabilistic model (i.e. DTMC) while the model considered in this paper (i.e. the simple probabilistic automata) is much more general since it takes the nondeterminism into account.

5.2. Future work

Many open questions remain. First we do not know whether the distance can be computed *exactly* (note our algorithm is an approximation algorithm). Moreover, it is interesting to establish the lower bound of the complexity (here we only give an EXPTIME upper bound). Furthermore, it is interesting to see whether we can extend the algorithm to more general (not restricted to probabilistic) systems, say metric transition systems studied in [3] where the underlying state space is not necessarily finite and other closely related models, say fuzzy automata.

Acknowledgement

We are grateful to Jun Pang who drew our attention on the algorithmic aspect of metric for probabilistic systems. We are also in debt to two anonymous reviewers for their constructive comments.

The first author is partially supported by the Dutch Bsik project BRICKS (Basic Research in Informatics for Creating the Knowledge Society). The second author is partially supported by the Dutch NWO project QUPES (Verification of Quantitative Properties of Embedded Software). The third author is partially supported by the Chinese national 863 program (2007AA01Z178), NSFC (60736015) and JSNSF (BK2006712).

References

[1] S. Basu, New results on quantifier elimination over real closed fields and applications to constraint databases, Journal of ACM 46 (4) (1999) 537–555.
[2] C. Baier, B.R. Haverkort, H. Hermanns, J.-P. Katoen, M. Siegle, Validation of Stochastic Systems — A Guide to Current Research, in: LNCS, vol. 2925, Springer, 2004.
[3] F. van Breugel, A behavioural pseudometric for metric labelled transition systems, in: Proc. of CONCUR'05, in: LNCS, vol. 3653, Springer, 2005, pp. 141–155.
[4] F. van Breugel, B. Sharma, J. Worrell, Approximating a behavioural pseudometric without discount, in: Proc. of FoSSaCS'07, in: LNCS, vol. 4423, Springer, 2007, pp. 123–137.
[5] F. van Breugel, J. Worrell, An algorithm for quantitative verification of probabilistic transition systems, in: Proc. of CONCUR'01, in: LNCS, vol. 2154, Springer, 2001, pp. 336–350.
[6] F. van Breugel, J. Worrell, A behavioural pseudometric for probabilistic transition systems, Theoretical Computer Science 331 (1) (2005) 115–142.
[7] T. Chen, T. Han, J. Lu, On behavioral metric for probabilistic systems: Definition and approximation Algorithm, in: Proc. of FSKD'07, IEEE Computer Society, 2007, pp. 21–25.
[8] Y. Deng, T. Chothia, C. Palamidessi, J. Pang, Metrics for action-labelled quantitative transition systems, ENTCS 153 (2) (2006) 79–96.
[9] J. Desharnais, R. Jagadeesan, V. Gupta, P. Panangaden, The metric analogue of weak bisimulation for probabilistic processes, in: Proc. of LICS'02, IEEE Computer Society, 2002, pp. 413–422.
[10] J. Desharnais, R. Jagadeesan, V. Gupta, P. Panangaden, Metrics for labelled Markov processes, Theoretical Computer Science 318 (3) (2004) 323–354.
[11] A. Giacalone, C.-C. Jou, S.A. Smolka, Algebraic reasoning for probabilistic concurrent systems, in: Proc. of IFIP WG 2.2/2.3 PCM'90, 1990, pp. 453–459.
[12] R.J. van Glabbeek, S.A. Smolka, B. Steffen, Reactive, generative, and stratified models of probabilistic processes, Information and Computation 121 (1) (1995) 59–80.
[13] K.G. Larsen, A. Skou, Bisimulation through probabilistic testing, Information and Computation 94 (1) (1991) 1–28.
[14] R. Milner, Communication and Concurrency, Prentice-Hall, 1989.
[15] A. Philippou, I. Lee, O. Sokolsky, Weak bisimulation for probabilistic systems, in: Proc. of CONCUR'00, in: LNCS, vol. 1877, Springer, 2000, pp. 334–349.
[16] M.O. Rabin, Probabilistic automata, Information and Control 6 (1963) 230–245.
[17] R. Segala, Modeling and verification of randomized distributed real-time systems, Technical Report MIT/LCS/TR-676, PhD thesis, Massachusetts Institute of Technology, 1995.
[18] M. Stoelinga, An introduction to probabilistic automata, Bulletin of the EATCS 78 (2002) 176–198.
[19] A. Traski, A Decision Method for Elementary Algebra and Geometry, Univ. of California Press, Berkeley, 1951.
[20] A. Tarski, A lattice-theoretical fixpoint theorem and its applications, Pacific Journal of Mathematics 5 (1955) 285–309.