Metrics for Action-labelled Quantitative Transition Systems [1]

Yuxin Deng [a,2] Tom Chothia [b] Catuscia Palamidessi [c]
Jun Pang [c]

[a] *INRIA Sophia-Antipolis and Université Paris 7, France*
deng.yuxin@pps.jussieu.fr

[b] *LIX, École Polytechnique, France*
tomc@lix.polytechnique.fr

[c] *INRIA Futurs and LIX, École Polytechnique, France*
{catuscia,pangjun}@lix.polytechnique.fr

Abstract

This paper defines *action-labelled quantitative transition systems* as a general framework for combining qualitative and quantitative analysis. We define *state-metrics* as a natural extension of bisimulation from non-quantitative systems to quantitative ones. We then prove that any single state-metric corresponds to a bisimulation and that the greatest state-metric corresponds to bisimilarity. Furthermore, we provide two extended examples which show that our results apply to both probabilistic and weighted automata as special cases of action-labelled quantitative transition systems.

Key words: Transition systems, quantitative, metrics, bisimulations, processes.

1 Introduction

Bisimulation, the widely used notion of equivalence for process calculi [16], provides a definition of equality that can capture similarities between processes without forcing them to be syntactically the same. The idea is to match any step in one process with a step, labelled by the same action, in the other process.

[1] Partially supported by the projet Rossignol of the ACI Sécurité Informatique (Ministère de la recherche et nouvelles technologies).

[2] The work of Yuxin Deng was supported by the EU project PROFUNDIS.

This is a preliminary version. The final version will be published in Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

Recently, there has been growing interest in systems that model quantitative processes. In these systems steps are associated with a given quantity, such as the probability that the step will happen [15,20,6] or the resources (e.g. time or cost) needed to perform that step [28,2,17]. The standard notion of bisimulation can be adapted to these systems by treating the quantities as labels (see for example [15,20,19,6]), but this does not provide a robust relation, since quantities are matched only when they are identical. Processes that differ for a very small probability, for instance, would be considered just as different as processes that perform completely different actions. This is particularly relevant to security systems where specifications can be given as perfect, but impractical processes and other, practical processes are considered safe if they only differ from the specification with a negligible probability.

To find a more flexible way to differentiate processes, researchers in this area have borrowed from pure mathematics the notion of metric [3]. A metric is defined as a function that associates a set distance with a pair of elements. Whereas topologists use metrics as a tool to study continuity and convergence, we will use them to provide a measure of the difference between two processes that are not quite bisimilar.

The first proposal based on metrics was by Giacalone *et al.* [10] for deterministic probabilistic processes. Later, Desharnais *et al.* [7,8] and van Breugel and Worrell [25,23] investigated the notion of metric for more general probabilistic systems, using much more sophisticated techniques to deal with the combination of probabilistic distribution, nondeterminism and recursion. In particular, they used the notion of Hutchinson metric [12] on distributions; this metric is also known under many different names including Kantorovich metric [13] and Vaserstein metric [27]. In [7,8], Desharnais *et al.* treated the case of labelled Markov chains and labelled concurrent Markov chains respectively, and defined the intended metric as the greatest fixed point of a monotonous function. In contrast, the authors of [25,23] used a construction based on the (unique) fixed point of a contractive transformation. They considered similar classes of automata, namely fully probabilistic systems and reactive models.

In this paper, we extend the approach of [7,8] to a more general framework that we call *Action-labelled Quantitative Transition Systems* (AQTS). Our framework subsumes some other well-known quantitative systems such as probabilistic automata, simple probabilistic automata [20], fully probabilistic models [1], reactive models, generative models [26] (see [21] for other related models and the relationship among them), as well as (a simplified version of) weighted automata [9,17].

The main contributions of this work are the following:

- We define a notion of metrics, which we call *state-metrics*, for AQTSs. These are based on the Hutchinson distance and coincide with the metrics in [7,8]

[3] For simplicity, in this paper we use the term metric to denote both metric and pseudometric. All the results of the paper are based on pseudometrics.

in the case of probabilistic transition systems.

- We show that each state-metric corresponds to a bisimulation and that as a consequence the greatest state-metric corresponds to bisimilarity. (In previous works, only the latter result was shown.) We show that the greatest state-metric can be characterised as the greatest fixed point of a monotonous function on state-metrics, which is closely analogous to Milner's characterization of bisimilarity as the greatest fixed point of a monotonous function on bisimulations [16].
- We consider two process calculi whose operational semantics are based on probabilistic and weighted automata respectively. We show that prefixing, choices and parallel composition constructors are non-expansive, which means that when different processes are placed in the same context they become more similar. This is a natural extension of the notion of congruence and matches our intuition that the larger the parts of two processes that are identical and the closer their behaviour the smaller the distance between them should be. Our result for probabilistic automata is similar to those in [7,8] for simple probabilistic automata.

The rest of the paper is structured as follows. In Section 2, after giving the formal definition of AQTS, we define state-metrics and the greatest state-metric. In Section 2.2 we show that state-metrics concur with bisimulations. Sections 3 and 4 both give specific examples of action-labelled quantitative transition systems. The first corresponds to Segala and Lynch's probabilistic automata and the second to weighted automata. In each case we define a process calculus and we show the non-expansiveness of some constructors with respect to the metric. In Section 5 we discuss related work. Section 6 concludes.

2 Action-labelled quantitative transition system

In this section we introduce the concept of AQTS. Then we define a pseudometric suitable for finite-state AQTSs, and we relate it to bisimulation.

Definition 2.1 A *action-labelled quantitative transition system* is defined as a tuple $(S, A, s_0, c, \rightarrow)$, where

(i) S is a set of states, and s_0 is the start state;

(ii) A is a finite set of action labels;

(iii) c is a positive real number;

(iv) Let $\mathcal{D}$ be all the functions $\eta : A \times S \mapsto [0, c]$ s.t. $\sum_{(a,s)\in(A\times S)} \eta(a, s) \leq c$. $\rightarrow \subseteq S \times \mathcal{D}$ represents a transition.

We shall use the more suggestive notation $s \rightarrow \eta$ instead of $(s, \eta) \in \rightarrow$. Note that AQTSs subsume various other models which have appeared in the literature. We illustrate some examples. Let $\mathcal{S} = (S, A, s_0, c, \rightarrow)$; if $c = 1$ then

$\mathcal{S}$ is a probabilistic automaton. By adding further constraints, we can obtain other models such as a simple probabilistic automaton, the reactive model and the generative model. On the other hand, if for each transition $s \to \eta$ there exists a unique pair $(a,t) \in A \times S$ s.t. $\eta(a,t) \neq 0$, then $\mathcal{S}$ is a *weighted automaton*, which is similar to an ordinary automaton in which a transition from s to t is labelled by a pair (a,w) where $w = \eta(a,t)$.

2.1 State-metrics

We fix an AQTS with the set of states S finite, and consider pseudometrics on S. A pseudometric is a function that yields a non-negative real number for each pair of states and satisfies the following: $m(s,s) = 0$; $m(s,t) = m(t,s)$; and $m(s,t) \leq m(s,u) + m(u,t)$. We say a pseudometric m is c-bounded if $\forall s,t : m(s,t) \leq c$, where c is a positive real number.

Definition 2.2 $\mathcal{M}_c$ is the class of c-bounded pseudometrics on states with the ordering

$$m_1 \preceq m_2 \text{ if } \forall s,t : m_1(s,t) \geq m_2(s,t).$$

Here we reverse the ordering with the purpose of characterizing bisimilarity as the *greatest* fixed point (cf: Corollary 2.14).

Lemma 2.3 $(\mathcal{M}_c, \preceq)$ *is a complete lattice.*

Proof. The top element is given by $\forall s,t : \top(s,t) = 0$; the bottom element is given by $\bot(s,t) = c$ if $s \neq t$, 0 otherwise. Greatest lower bounds are given by $(\bigsqcap X)(s,t) = \sup\{m(s,t) \mid m \in X\}$ for any $X \subseteq \mathcal{M}_c$. Finally, least upper bounds are given by $\bigsqcup X = \bigsqcap \{m \in \mathcal{M}_c \mid \forall m' \in X : m' \preceq m\}$. □

In order to define the notion of state-metrics (which will correspond to bisimulations) and the monotonous transformation on metrics, we need to associate a metric with $\mathcal{D}$. We give a definition based on the Hutchinson metric [12] on probability measures, which has been used by van Breugel and Worrell for defining metrics on fully probabilistic systems [23] and reactive probabilistic systems [24]; and by Desharnais *et al.* for labelled Markov chains [7] and labelled concurrent Markov chains [8], respectively.

In the following, for $\eta \in \mathcal{D}$, we will call the *total mass* of η the number $\sum_{(a,s)\in A\times S} \eta(a,s)$.

Definition 2.4 For each $m \in \mathcal{M}_c$, we lift it to be a metric on distributions. Given $\eta, \eta' \in \mathcal{D}$, we define $m(\eta,\eta')$ (with a slight abuse of the notation m) as follows:

(i) if the total mass of η is not less than the total mass of η', then $m(\eta,\eta')$ is given by the solution to the following linear program:

$$
\begin{array}{lll}
\text{maximize} & \frac{1}{c}\cdot\sum_{(a_i,s_i)\in A\times S}(\eta(a_i,s_i)-\eta'(a_i,s_i))x_i & (*)\\
\text{subject to} & \text{- } \forall i: 0\leq x_i\leq c & \\
& \text{- } \forall i,j: x_i - x_j \leq \hat{m}((a_i,s_i),(a_j,s_j)) & \\
& \text{where } \hat{m}((a_i,s_i),(a_j,s_j)) = \begin{cases} c & \text{if } a_i\neq a_j \\ m(s_i,s_j) & \text{otherwise}\end{cases} &
\end{array}
$$

(ii) if the total mass of η is less than the total mass of η', then $m(\eta,\eta')$ is defined to be $m(\eta',\eta)$.

Note that since i and j range over all indexes it is unnecessary to require $|x_i - x_j|$ to be less than or equal to $\hat{m}((a_i,s_i),(a_j,s_j))$ in the second constraint. It can be shown that m defined in this way is a pseudometric on $\mathcal{D}$.

An alternative definition would be to scale the above $m(\eta,\eta')$ by a factor $e\in(0,1]$, see van Breugel and Worrell [25] for more discussions. Here we simply let $e=1$ because all the main results obtained in this paper are independent from e.

Definition 2.5 $m\in\mathcal{M}_c$ is a *state-metric* if, for all $\epsilon\in[0,c)$, $m(s,t)\leq\epsilon$ implies:

- if $s\to\eta$ then there exists some η' such that $t\to\eta'$ and $m(\eta,\eta')\leq\epsilon$.

Note that if m is a state-metric then it is also a metric. By $m(s,t)\leq\epsilon$ we have $m(t,s)\leq\epsilon$, which implies

- if $t\to\eta'$ then there exists some η such that $s\to\eta$ and $m(\eta',\eta)\leq\epsilon$.

In the above definition, we prohibit ϵ to be c because throughout this paper c represents the distance between any two states including the case where one state may perform a transition and the other may not.

The greatest state-metric is defined as

$$m_{max} = \bigsqcup\{m\in\mathcal{M}_c \mid m \text{ is a state-metric}\}.$$

When compared with the labelled transition system in CCS [16], it turns out that state-metrics correspond to bisimulations and the greatest state-metric corresponds to bisimilarity. To make the analogy closer, in what follows we will characterize m_{max} as a fixed point of a suitable monotonous function on $\mathcal{M}_c$. First we recall the definition of Hausdorff distance.

Definition 2.6 Given a c-bounded metric d on Z, the *Hausdorff distance* between two subsets X,Y of Z is defined as follows:

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

where $\inf\emptyset = c$ and $\sup\emptyset = 0$.

Next we define a function F on $\mathcal{M}_c$ by using the Hausdorff distance.

Definition 2.7 Let $tr(s) = \{\eta \mid s \rightarrow \eta\}$. $F(m)$ is a pseudometric given by:

$$F(m)(s,t) = H_m(tr(s), tr(t)).$$

Thus we have the following property.

Lemma 2.8 *For all $\epsilon \in [0, c)$, $F(m)(s,t) \leq \epsilon$ if and only if:*

- *if $s \rightarrow \eta$ then there exists some η' such that $t \rightarrow \eta'$ and $m(\eta, \eta') \leq \epsilon$;*
- *if $t \rightarrow \eta'$ then there exists some η such that $s \rightarrow \eta$ and $m(\eta', \eta) \leq \epsilon$.*

The above lemma can be proved by directly checking the definition of F, as can the next lemma.

Lemma 2.9 *m is a state-metric iff $m \preceq F(m)$.*

Consequently, we have the following characterization:

$$m_{max} = \bigsqcup \{m \in \mathcal{M}_c \mid m \preceq F(m)\}.$$

Lemma 2.10 *F is monotonous on $\mathcal{M}_c$.*

Because of Lemma 2.3 and 2.10, we can apply Tarski's fixed point theorem [22], which tells us that m_{max} is the greatest fixed point of F. Furthermore, by Lemma 2.9 we know that m_{max} is indeed a state-metric, and it is the greatest state-metric.

In addition, the finite-stateness of AQTSs ensures that the closure ordinal of F is ω (cf:[7], Lemma 3.10). Therefore one can proceed in a standard way to show that

$$m_{max} = \bigsqcap \{F^i(\top) \mid i \in \mathbb{N}\}$$

where $\top$ is the top metric in $\mathcal{M}_c$ and $F^0(\top) = \top$.

2.2 Bisimulations

Let $\eta \in \mathcal{D}$, $a \in A$ and $V \subseteq S$, we write $\eta(a, V)$ for $\sum_{t \in V} \eta(a, t)$. We lift an equivalence relation on S to a relation on $\mathcal{D}$ in the following way:

Definition 2.11 Let $\eta, \eta' \in \mathcal{D}$, we say they are equivalent w.r.t. an equivalence relation $\mathcal{R}$ on S, written $\eta \equiv_{\mathcal{R}} \eta'$, if

$$\forall a \in A, \forall V \in S/\mathcal{R} : \eta(a, V) = \eta'(a, V).$$

We now show the correspondence between our state-metrics and bisimulations. More precisely, the correspondence is with the extension of Larsen and Skou's probabilistic bisimulation [15] to AQTSs.

Definition 2.12 An equivalence relation $\mathcal{R} \subseteq S \times S$ is a *(strong) bisimulation* if $s\mathcal{R}t$ implies:

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

Two states s, t are *bisimilar*, written $s \sim t$, if there exists a bisimulation $\mathcal{R}$ s.t. $s\mathcal{R}t$.

In the above definition, the "vice-versa" case is covered by the fact that $\mathcal{R}$ is an equivalence relation.

A bisimulation is related to a state-metric by the following theorem:

Theorem 2.13 *Given a binary relation $\mathcal{R}$ and a pseudometric $m \in \mathcal{M}_c$ such that*

$$m(s,t) = \begin{cases} 0 & \text{if } s\mathcal{R}t \\ c & \text{otherwise.} \end{cases}$$

Then $\mathcal{R}$ is a bisimulation iff m is a state-metric.

Proof. Given two distributions η, η', let us consider how to compute $m(\eta, \eta')$ if $\mathcal{R}$ is an equivalence relation. Since S is finite, we may assume that $V_1, ..., V_n \in S/\mathcal{R}$ are all the equivalence classes of S under $\mathcal{R}$. We rewrite the linear program (*) in the form:

$$\frac{1}{c} \cdot \sum_{a_i \in A} \sum_{s_j \in S} (\eta(a_i, s_j) - \eta'(a_i, s_j))x_{ij} \tag{1}$$

If $s_{j_1}, s_{j_2} \in V_j$ for some $j \in 1..n$, then $\hat{m}((a_i, s_{j_1}), (a_i, s_{j_2})) = m(s_{j_1}, s_{j_2}) = 0$, which implies $x_{ij_1} = x_{ij_2}$ by the second constraint of (*). Thus, some summands of (1) can be grouped together and we have the following linear program:

$$\frac{1}{c} \cdot \sum_{a_i \in A} \sum_{j \in 1..n} (\eta(a_i, V_j) - \eta'(a_i, V_j))x_{ij} \tag{2}$$

with the constraint $x_{ij} - x_{i'j'} \leq c$ for any two variables x_{ij} and $x_{i'j'}$. Therefore if $\mathcal{R}$ is an equivalence relation then $m(\eta, \eta')$ is obtained by maximizing the linear program (2).

($\Rightarrow$) Suppose $\mathcal{R}$ is a bisimulation and $m(s,t) = 0$. Clearly $\mathcal{R}$ is an equivalence relation. By the definition of m we have $s\mathcal{R}t$. If $s \rightarrow \eta$ then $t \rightarrow \eta'$ for some η' such that $\eta \equiv_{\mathcal{R}} \eta'$. To show that m is a state-metric it suffices to prove $m(\eta, \eta') = 0$. We know from $\eta \equiv_{\mathcal{R}} \eta'$ that $\eta(a_i, V_j) = \eta'(a_i, V_j)$, for each $j \in 1..n$. It follows that (2) is maximized to be 0, thus $m(\eta, \eta') = 0$.

($\Leftarrow$) Suppose m is as defined in the hypothesis. It is clear that $\mathcal{R}$ is an equivalence relation. We show that it is a bisimulation. Suppose $s\mathcal{R}t$, which means $m(s,t) = 0$. If $s \rightarrow \eta$ then $t \rightarrow \eta'$ for some η' such that $m(\eta, \eta') = 0$. Without loss of generality we assume that the total mass of η is not less than the total mass of η'.

$$\sum_{a_i \in A} \eta(a_i, S) \geq \sum_{a_i \in A} \eta'(a_i, S) \tag{3}$$

To ensure that $m(\eta, \eta') = 0$, in (2) the following two conditions must be satisfied.

(i) No coefficient is positive. Otherwise, if $\eta(a_i, V_j) - \eta'(a_i, V_j) > 0$ then (2) would be maximized to a value not less than $(\eta(a_i, V_j) - \eta'(a_i, V_j))$, which is greater than 0.

(ii) It is not the case that at least one coefficient is negative and the other coefficients are either negative or 0. Otherwise, by summing all the coefficients, we would get

$$\sum_{a_i \in A} (\eta(a_i, S) - \eta'(a_i, S)) < 0$$

which contradicts (3).

Therefore, the only possibility is that all coefficients in (2) are 0, i.e., $\eta(a_i, V_j) = \eta'(a_i, V_j)$ for any action a_i and equivalence class $V_j \in S/\mathcal{R}$. It follows that $\eta \equiv_{\mathcal{R}} \eta'$. So we have shown that $\mathcal{R}$ is indeed a bisimulation. □

Corollary 2.14 *$s \sim t$ iff $m_{max}(s, t) = 0$.*

Proof. ($\Rightarrow$) If $s \sim t$ then there exists a bisimulation $\mathcal{R}$ such that $s\mathcal{R}t$. By Theorem 2.13 there exists some state-metric m such that $m(s, t) = 0$. By the definition of m_{max} we have $m \preceq m_{max}$. Therefore $m_{max}(s, t) \leq m(s, t) = 0$.

($\Leftarrow$) From m_{max} we construct a pseudometric m as follows.

$$m(s, t) = \begin{cases} 0 & \text{if } m_{max}(s, t) = 0 \\ c & \text{otherwise.} \end{cases}$$

Since m_{max} is a state-metric, it is easy to see that m is also a state-metric. Now we construct a binary relation $\mathcal{R}$ such that $\forall s, s' : s\mathcal{R}s'$ iff $m(s, s') = 0$. If follows from Theorem 2.13 that $\mathcal{R}$ is a bisimulation. If $m_{max}(s, t) = 0$, then $m(s, t) = 0$ and thus $s\mathcal{R}t$. Therefore we have the required result $s \sim t$ because $\sim$ is the largest bisimulation. □

3 Example: probabilistic finite behaviours

In this section we consider a simple process calculus whose semantics is given by Segala and Lynch's general probabilistic automata, which admit both probability and nondeterminism. We define a parallel composition constructor for the process calculus and show that, like probabilistic and nondeterministic choices, it is non-expansive in the sense of [7].

First we give some preliminary notations. Let U be a set. A function $\eta : U \mapsto [0, 1]$ is called a *discrete probability distribution*, or *distribution* for short, on U if the *support* of η, defined as $spt(\eta) = \{x \in U \mid \eta(x) > 0\}$, is finite or countably infinite and $\sum_{x \in U} \eta(x) \leq 1$. Given two distributions η and η', we can define their sum $\eta \uplus \eta'$, if the function given by $(\eta \uplus \eta')(x) = \eta(x) + \eta'(x)$ is still a distribution. If η is a distribution with finite support and $V \subseteq spt(\eta)$ we use the set $\{(s_i : \eta(s_i))\}_{s_i \in V}$ to enumerate the probability associated with each element of V. For convenience of presentation, we may consider distributions as either functions or sets, depending on their contexts.

Processes are defined by the following syntax:

$$P, Q ::= \bigoplus_{i \in 1..n} p_i a_i . P_i \;\Big|\; \bar{a}.P \;\Big|\; \sum_{i \in 1..m} P_i \;\Big|\; P|Q$$

conm	$\bar{a}.P \to \{(\bar{a}, P : 1)\}$	psum	$\bigoplus_{i\in 1..n} p_i a_i.P_i \to \biguplus_{i\in 1..n}\{(a_i, P_i : p_i)\}$
par	$\dfrac{P \to \{(a_i, P_i : p_i)\}_i}{P\|Q \to \{(a_i, P_i\|Q : p_i)\}_i}$	nsum	$\dfrac{P_j \to \eta}{\sum_{i\in 1..m} P_i \to \eta}$ for some $j \in 1..m$
com	$\dfrac{P \to \{(\bar{a}, P' : 1)\} \qquad Q \to \{(a, Q_i : p_i)\}_{i\in I} \uplus \{(a_j, Q_j : q_j)\}_{j\in J}}{P\|Q \to \{(\tau, P'\|Q_i : p_i)\}_{i\in I} \uplus \{(a_j, P\|Q_j : q_j)\}_{j\in J}}$		

Table 1
Transitions of probabilistic automata

Here $\bigoplus_{i\in 1..n} p_i a_i.P_i$ stands for a *probabilistic choice* constructor, where the p_i's represent positive probabilities, i.e., they satisfy $p_i \in (0,1]$ and $\sum_{i\in 1..n} p_i = 1$. When $n = 0$ we abbreviate the probabilistic choice as $\mathbf{0}$. Sometimes we are interested in certain branches of the probabilistic choice; in this case we write $\bigoplus_{i\in 1..n} p_i a_i.E_i$ as $p_1 a_1.E_1 \oplus \cdots \oplus p_n a_n.E_n$. The third construction $\sum_{i\in 1..m} P_i$ stands for *nondeterministic choice*, and occasionally we may write it as $P_1 + ... + P_m$. The above syntax defines only finite processes.

The operational semantics of a process P is defined as a probabilistic automaton whose states are the processes reachable from P and the transition relation is defined by the axioms and inference rules in Table 1, where $P \to \eta$ describes a transition that leaves from P and leads to a distribution η over $A \times S$. In rule com we require the condition that $\forall j \in J : a_j \neq a$. The symmetric rules of par and com are omitted. Here parallel composition is defined to be asynchronous as in [11]. This way of treating parallel composition is itself of interest: admitting parallelism in probabilistic automata is often considered as a hard problem [20], while our treatment here is quite simple.

We show that prefixing, choices and parallel composition are non-expansive.

Proposition 3.1 *If* $m_{max}(P,\ Q) \leq \epsilon$ *then*

(i) $m_{max}(\bar{a}.P,\ \bar{a}.Q) \leq \epsilon$

(ii) $m_{max}(p_1 a_1.R_1 \oplus ... \oplus p_n a_n.R_n \oplus pa.P,\ p_1 a_1.R_1 \oplus ... \oplus p_n a_n.R_n \oplus pa.Q) \leq \epsilon$

(iii) $m_{max}(R + P,\ R + Q) \leq \epsilon$

(iv) $m_{max}(R|P,\ R|Q) \leq \epsilon$.

Proof. The first three clauses are straightforward. The last clause is proved by induction on the size of the process $R|P + R|Q$, as only finite processes are involved. When the size is 0 the result is immediate. For the inductive step, there are four cases, among which we consider the hardest one. Since m_{max} is a fixed point of F, we only need to show that if $R|P \to \eta$ there is $R|Q \to \eta'$ such that $m_{max}(\eta, \eta') \leq \epsilon$. Suppose $P \to \theta = \{(a, P_i : p_i)\}_{i\in I_1} \uplus \{(a_i, P_i : p_i)\}_{i\in I_2}$ with $a_i \neq a$ for all $i \in I_2$. Since $m_{max}(P, Q) \leq \epsilon$, there exists $Q \to \theta' = \{(a, P_i : p_i)\}_{i\in I_1'} \uplus \{(a_i, P_i : p_i)\}_{i\in I_2'}$ such that $a_i \neq a$ for all $i \in I_2'$ and $m_{max}(\theta, \theta') \leq \epsilon$.

That is, the linear program

$$\sum_{i\in I_1\cup I_2} p_i x_i - \sum_{i\in I_1'\cup I_2'} p_i x_i \tag{4}$$

subject to

- $x_i - x_j \le m_{max}(P_i, P_j)$ for all $i, j \in I_1 \cup I_1'$
- $x_i - x_j \le \hat{m}_{max}((a_i, P_i), (a_j, P_j))$ for all $i, j \in I_2 \cup I_2'$
- $x_i - x_j \le 1$ for all $i \in I_1 \cup I_1', j \in I_2 \cup I_2'$ or $i \in I_2 \cup I_2', j \in I_1 \cup I_1'$

is maximized to a value not greater than ϵ. Let $R \to \{(\bar{a}, R' : 1)\}$. We have $R|P \to \eta = \{(\tau, R'|P_i : p_i)\}_{i\in I_1} \uplus \{(a_i, R|P_i : p_i)\}_{i\in I_2}$ and $Q|R \to \eta' = \{(\tau, R'|Q_i : p_i)\}_{i\in I_1'} \uplus \{(a_i, R|Q_i : p_i)\}_{i\in I_2'}$. Then $m_{max}(\eta, \eta')$ is the maximum value of the linear program (4) subject to

- $x_i - x_j \le m_{max}(R'|P_i, R'|P_j)$ for all $i, j \in I_1 \cup I_1'$
- $x_i - x_j \le \hat{m}_{max}((a_i, R|P_i), (a_j, R|P_j))$ for all $i, j \in I_2 \cup I_2'$
- $x_i - x_j \le \hat{m}_{max}((\tau, R'|P_i), (a_j, R|P_j)$ for all $i \in I_1 \cup I_1', j \in I_2 \cup I_2'$
- $x_i - x_j \le \hat{m}_{max}((a_i, R|P_i), (\tau, R'|P_j)$ for all $i \in I_2 \cup I_2', j \in I_1 \cup I_1'$

By induction hypothesis, we have

- $m_{max}(R'|P_i, R'|P_j) \le m_{max}(P_i, P_j)$ for all $i, j \in I_1 \cup I_1'$
- $\hat{m}_{max}((a_i, R|P_i), (a_j, R|P_j)) \le \hat{m}_{max}((a_i, P_i), (a_j, P_j))$ for all $i, j \in I_2 \cup I_2'$

and clearly

- $\hat{m}_{max}((\tau, R'|P_i), (a_j, R|P_j) \le 1$ for all $i \in I_1 \cup I_1', j \in I_2 \cup I_2'$
- $\hat{m}_{max}((a_i, R|P_i), (\tau, R'|P_j) \le 1$ for all $i \in I_2 \cup I_2', j \in I_1 \cup I_1'$.

It follows that $m_{max}(\eta, \eta') \le m_{max}(\theta, \theta') \le \epsilon$. □

Here we do not consider recursion $\mu_X E$ because this constructor is not non-expansive. For example, let $E = \frac{1}{2}a.\mathbf{0} \oplus \frac{1}{2}a.X$ and $F = \frac{1}{3}a.\mathbf{0} \oplus \frac{2}{3}a.X$. Suppose the distance between $\mathbf{0}$ and X is 1, then $m_{max}(E, F) = \frac{1}{6}$. On the other hand we have $m_{max}(\mu_X E, \mu_X F) = \frac{1}{3} > \frac{1}{6}$.

4 Example: weighted automata

4.1 Bisimulations on weighted automata

In this section we consider weighted automata, which are degenerate AQTSs in that for each transition $s \to \eta$ there is a unique pair $(a, t) \in A \times S$ s.t. $\eta(a, t) \neq 0$. For simplicity we write this transition as $s \xrightarrow{a[w]} t$, where $w = \eta(a, t)$. In the literature, the weight w is interpreted as *cost* in some places (e.g. [17]) and as *time* in some other places (e.g. [2]), and strong bisimulation is often defined as follows.

Definition 4.1 A binary relation $\mathcal{R} \subseteq S \times S$ is a labelled bisimulation if $s\mathcal{R}t$ implies:

- if $s \xrightarrow{a[w]} s'$ there exists t' such that $t \xrightarrow{a[w]} t'$ and $s'\mathcal{R}t'$;
- if $t \xrightarrow{a[w]} t'$ there exists s' such that $s \xrightarrow{a[w]} s'$ and $s'\mathcal{R}t'$.

Two states s, t are labelled bisimilar, written $s \sim' t$, if there exists a labelled bisimulation $\mathcal{R}$ s.t. $s\mathcal{R}t$.

Unlike bisimulation (cf: Definition 2.12) a labelled bisimulation is not necessarily an equivalence relation. However, as far as weighted automata are concerned, $\sim$ coincides with $\sim'$.

Lemma 4.2 *In weighted automata, $s \sim t$ iff $s \sim' t$.*

Proof. ($\Rightarrow$) It is easy to see that if $\mathcal{R}$ is a bisimulation then $\mathcal{R}$ is a labelled bisimulation.

($\Leftarrow$) It can be shown that if $\mathcal{R}$ is a labelled bisimulation then $\mathcal{R}^*$ is a bisimulation, where $\mathcal{R}^*$ is the equivalence (reflexive, symmetric and transitive) closure of $\mathcal{R}$. □

Corollary 4.3 *In weighted automata, $s \sim' t$ iff $m_{max}(s, t) = 0$.*

Proof. By Corollary 2.14 and the preceding lemma. □

4.2 Non-expansiveness of some constructors

We now give a process calculus with weighted automata as its underlying operational semantics. Processes are defined by the following syntax:

$$P, Q ::= \mathbf{0} \;\Big|\; \alpha.P \;\Big|\; P + Q \;\Big|\; P|Q$$

$$\alpha ::= a[w] \;\Big|\; \bar{a}[w] \;\Big|\; \tau[w]$$

The operational behaviour of a process P is described as a weighted automaton whose states are the derivatives of P and the transition relation is defined by the rules in Table 2. The symmetric rules of sum, par and com are omitted. The intuition for the semantics is as follows. In a transition $P \xrightarrow{a[w]} P'$, process P performs action a, with the maximal cost or maximal delay of time w, before evolving into P'. Therefore in rule com the weight w should be the smaller one between w_1 and w_2.

In this process calculus, the constructions prefixing, choice and parallel composition are non-expansive.

Proposition 4.4 *If $m_{max}(P,\ Q) \leq \epsilon$ then*

(i) $m_{max}(\alpha.P,\ \alpha.Q) \leq \epsilon$

(ii) $m_{max}(R + P,\ R + Q) \leq \epsilon$

pre	$\dfrac{}{\alpha.P \xrightarrow{\alpha} P}$	sum	$\dfrac{P \xrightarrow{\alpha} P'}{P+Q \xrightarrow{\alpha} P'}$
par	$\dfrac{P \xrightarrow{\alpha} P'}{P\vert Q \xrightarrow{\alpha} P'\vert Q}$	com	$\dfrac{P \xrightarrow{a[w_1]} P' \quad Q \xrightarrow{\bar{a}[w_2]} Q' \quad w = \min\{w_1, w_2\}}{P\vert Q \xrightarrow{\tau[w]} P'\vert Q'}$

Table 2
Transitions of weighted automata

(iii) $m_{max}(R|P,\ R|Q) \le \epsilon$

Proof. The first two clause are easy to show. The last clause is proved by induction on the size of the process $R|P + R|Q$. When the size is 0 the result is immediate. For the inductive step we consider one case. Suppose $R|P \to \eta$ with $\eta = \{(\tau, R'|P' : w)\}$, i.e., $R|P \xrightarrow{\tau[w]} R'|P'$. We consider the situation that the transition comes from $R \xrightarrow{\bar{a}[w_1]} R'$ and $P \xrightarrow{a[w_2]} P'$, with $w = \min\{w_1, w_2\}$. We need to show that there exists some η' such that $R|Q \to \eta'$ and $m_{max}(\eta, \eta') \le \epsilon$. Since $m_{max}(P, Q) \le \epsilon$, it can be shown that there exists a transition $Q \xrightarrow{b[w_2']} Q'$ such that

(5) $$|w_2 - w_2'| + \frac{\min\{w_2, w_2'\}}{c} \cdot \hat{m}_{max}((a, P'), (b, Q')) \le \epsilon.$$

Now there are two possibilities.

- If $a \neq b$ then $\hat{m}_{max}((a, P'), (b, Q')) = c$, thus it follows from (5) that $\max\{w_2, w_2'\} \le \epsilon$. Now we have the transition $R|Q \xrightarrow{b[w_2']} R|Q'$ and

$$\begin{aligned}
&|w - w_2'| + \tfrac{\min\{w, w_2'\}}{c} \cdot \hat{m}_{max}((\tau, R'|P'), (b, R|Q')) \\
&\le |w - w_2'| + \min\{w, w_2'\} \\
&= \max\{w, w_2'\} \\
&\le \max\{w_2, w_2'\} \\
&\le \epsilon.
\end{aligned}$$

  In other words, we have the required condition that $R|Q \to \eta'$ and $m_{max}(\eta, \eta') \le \epsilon$, where $\eta' = \{(b, R|Q' : w_2')\}$.

- If $a = b$ then it follows from (5) that

  (6) $$|w_2 - w_2'| + \frac{\min\{w_2, w_2'\}}{c} \cdot m_{max}(P', Q') \le \epsilon.$$

  Observe that we have the transition $R|Q \xrightarrow{\tau[w']} R'|Q'$ where $w' = \min\{w_1, w_2'\}$. That is, $R|Q \to \eta'$, where $\eta' = \{(\tau, R'|Q' : w')\}$.

Now we assert that

$$|w - w'| + \frac{\min\{w, w'\}}{c} \cdot m_{max}(P', Q') \le \epsilon, \tag{7}$$

which can be proved by analyzing the relation between w_1 and w_2, w_2'. There are four cases:

(i) $w_1 \le w_2$ and $w_1 \le w_2'$;
(ii) $w_2 \le w_1$ and $w_2' \le w_1$;
(iii) $w_2' \le w_1 \le w_2$;
(iv) $w_2 \le w_1 \le w_2'$.

As an example we only consider the last case; the first three cases are similar. If $w_2 \le w_1 \le w_2'$ then $w = w_2$ and $w' = w_1$. Thus we have the following:

$$\begin{array}{ll} & |w - w'| + \frac{\min\{w,w'\}}{c} \cdot m_{max}(P', Q') \\ = & |w_2 - w_1| + \frac{\min\{w_2,w_1\}}{c} \cdot m_{max}(P', Q') \\ \le & |w_2 - w_2'| + \frac{\min\{w_2,w_2'\}}{c} \cdot m_{max}(P', Q') \\ \le & \epsilon \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{by (6).} \end{array}$$

Therefore (7) holds.

At last, we can derive that

$$\begin{array}{lll} & |w - w'| + \frac{\min\{w,w'\}}{c} \cdot \hat{m}_{max}((\tau, R'|P'), (\tau, R'|Q')) & \\ = & |w - w'| + \frac{\min\{w,w'\}}{c} \cdot m_{max}(R'|P', R'|Q') & \\ \le & |w - w'| + \frac{\min\{w,w'\}}{c} \cdot m_{max}(P', Q') & \text{by induction hypothesis} \\ \le & \epsilon & \text{by (7).} \end{array}$$

It follows that $m_{max}(\eta, \eta') \le \epsilon$.

□

One can also define restriction and relabelling constructors in the process calculus in the style of CCS, and then show their non-expansiveness. However, as in Section 3 recursion is still not non-expansive.

5 Related work

Giacalone *et al.* [10] were the first to suggest a metric between probabilistic transition systems to formalize the notion of distance between processes. Metrics were used also in [14,18] to give denotational semantics for reactive models. De Vink and Rutten [4] showed that discrete probabilistic transition systems can be viewed as coalgebras. They considered the category of complete ultrametric spaces. Similar ultrametric spaces are considered by den Hartog in [5]. De Alfaro *et al.* [3] presented a quantitative transition system by interpreting propositions as numbers between 0 and 1, without considering action labels on the transitions. This system is quite different from the usual labelled transition systems and it is hard to precisely compare their metrics with the metrics in all other models (see [21] for an overview) and our transition systems.

The works most related to ours are [8,7,25,23]. Desharnais *et al.* [8] studied a logical pseudometric for labelled Markov chains, which is a reactive model of probabilistic systems. The metric has the property that two processes have distance of 0 if and only if they are probabilistic bisimilar. They also introduced a probabilistic process calculus and showed that some of the process constructors are non-expansive. A similar pseudometric was defined by van Breugel and Worrell [24] via the terminal coalgebra of a functor based on a metric on the space of Borel probability measures. Interestingly, van Breugel and Worrell [23] also presented a polynomial-time algorithm to approximate their coalgebraic distances. In [7] Desharnais *et al.* dealt with labelled concurrent Markov chains (this model can be captured by the simple probabilistic automata of [20]). They showed that the greatest fixed point of a monotonous function on pseudometrics corresponds to the weak probabilistic bisimilarity of [19]. They also showed that some process constructors of a probabilistic process calculus are non-expansive.

In comparison with the works [8,7,25,23] discussed above, we note that: (i) our results on state-metrics hold for a strictly more general framework; (ii) besides characterizing bisimilarity, we have the more refined property of using state-metrics to characterize every bisimulation relation; (iii) our results on non-expansiveness of some constructors in Sections 3 and 4 hold for probabilistic and weighted automata respectively, while the non-expansiveness results of [8] and [7] are for a kind of reactive models and simple probabilistic automata respectively; (iv) the metric of [8,25] works for continuous probabilistic transition systems, while in this work we concentrate on discrete systems.

Another interesting work is [29], in which Ying proposed the notion of bisimulation index for the usual labelled transition systems, by using ultrametrics on actions instead of using pseudometrics on states. He applied bisimulation indexes on timed CCS and real time ACP. But the deeper connection between [29] and our work worths some further studies.

6 Concluding remarks

We have presented the notion of action-labelled quantitative transition systems, a class of quantitative automata which subsume various traditional models used in quantitative verification, such as probabilistic automata and weighted automata. We have investigated a metric semantics on the new transition systems and we have related it to the classical bisimulation-based semantics. More precisely, we have shown that state-metrics correspond to bisimulations and the greatest state-metric corresponds to bisimilarity. Additionally, we have shown the non-expansiveness of some constructors for probabilistic and weighted automata.

In this paper we have considered only strong bisimulations. The extension of our results to weak bisimulations is not difficult if weak transitions are appropriately defined. For example, as far as probabilistic automata are concerned, we could define weak transition $\Rightarrow$ in the way as in [6], thus weak state-metric and weak bisimulation would be as follows.

$m \in \mathcal{M}_c$ is a *weak state-metric* if, for all $\epsilon \in [0, c)$, $m(s,t) \leq \epsilon$ implies:
- if $s \Rightarrow \eta$ then there exists some η' such that $t \Rightarrow \eta'$ and $m(\eta, \eta') \leq \epsilon$.

An equivalence relation $\mathcal{R} \subseteq S \times S$ is a *weak bisimulation* if $s\mathcal{R}t$ implies:
- whenever $s \Rightarrow \eta$, there exists η' such that $t \Rightarrow \eta'$ and $\eta \equiv_{\mathcal{R}} \eta'$.

We write $s \approx t$ if there exists a weak bisimulation $\mathcal{R}$ s.t. $s\mathcal{R}t$.

By arguments similar to those in Section 2, one could verify the property that weak state-metrics correspond to weak bisimulations and the greatest weak state-metric characterizes weak bisimilarity $\approx$.

As to the future work, it might be interesting to see what kind of non-expansiveness is enjoyed by recursion construct. It is also worth studying logical characterization of state-metrics and developing efficient algorithms to compute the distance between any two finite-state AQTSs.

Acknowledgement

We thank the anonymous referees for their valuable comments on a preliminary version of the paper.

References

[1] C. Baier and H. Hermanns. Weak bisimulation for fully probabilistic processes. In *Proceedings of the 9th International Conference on Computer Aided Verification*, volume 1254 of *Lecture Notes in Computer Science*, pages 119–130. Springer-Verlag, 1997.

[2] L. Chen. Timed processes: Models, axioms and decidability. Technical Report ECS-LFCS-93-271, LFCS, University of Edinburgh, 1993.

[3] L. de Alfaro, M. Faella, and M. I. A. Stoelinga. Linear and branching metrics for quantitative transition systems. In *Proceedings of the 31st International Colloquium on Automata, Languages and Programming*, volume 3142 of *Lecture Notes in Computer Science*, pages 97–109. Springer–Verlag, 2004.

[4] E. P. de Vink and J. J. M. M. Rutten. Bisimulation for probabilistic transition systems: A coalgebraic approach. *Theoretical Computer Science*, 221(1/2):271–293, 1999.

[5] J. I. den Hartog. *Probabilistic Extensions of Semantical Models*. PhD thesis, Free University Amsterdam, 2002.

[6] Y. Deng and C. Palamidessi. Axiomatizations for probabilistic finite-state behaviors. In *Proceedings of the 8th International Conference on Foundations of Software Science and Computation Structures*, volume 3441 of *Lecture Notes in Computer Science*, pages 110–124. Springer, 2005.

[7] J. Desharnais, R. Jagadeesan, V. Gupta, and P. Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science*, pages 413–422. IEEE Computer Society, 2002.

[8] J. Desharnais, R. Jagadeesan, V. Gupta, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004. An early version of this paper appeared in the proceedings of CONCUR'99.

[9] S. Eilenberg. *Automata, Languages, and Machines.* Academic Press, 1974.

[10] A. Giacalone, C.-C. Jou, and S. A. Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Proceedings of IFIP WG 2.2/2.3 Working Conference on Programming Concepts and Methods*, pages 453–459, 1990.

[11] O. M. Herescu and C. Palamidessi. Probabilistic asynchronous π-calculus. In *Proceedings of the 3rd International Conference Foundations of Software Science and Computation Structures*, volume 1784 of *Lecture Notes in Computer Science*, pages 146–160. Springer-Verlag, 2000.

[12] J. E. Hutchinson. Fractals and self-similarity. *Indiana University Mathematics Journal*, 30:713–747, 1981.

[13] L. Kantorovich. On the transfer of masses (in Russian). *Doklady Akademii Nauk*, 5(1):1–4, 1942. Translated in *Management Science*, 5(1):1–4, 1958.

[14] M. Z. Kwiatkowska and G. J. Norman. Probabilistic metric semantics for a simple language with recursion. In *Proceedings of 21st International Symposium on Mathematical Foundations of Computer Science*, volume 1113 of *Lecture Notes in Computer Science*, pages 419–430. Springer-Verlag, 1996.

[15] K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.

[16] R. Milner. *Communication and Concurrency.* Prentice-Hall, 1989.

[17] M. Mohri. Edit-distance of weighted automata: General definitions and algorithms. *International Journal of Foundations of Computer Science*, 14(6):957–982, 2003.

[18] G. J. Norman. *Metric Semantics for Reactive Probabilistic Systems.* PhD thesis, University of Birmingham, 1997.

[19] A. Philippou, I. Lee, and O. Sokolsky. Weak bisimulation for probabilistic systems. In *Proceedings of the 11th International Conference on Concurrency Theory*, volume 1877 of *Lecture Notes in Computer Science*, pages 334–349. Springer-Verlag, 2000.

[20] R. Segala. Modeling and verification of randomized distributed real-time systems. Technical Report MIT/LCS/TR-676, PhD thesis, Massachusetts Institute of Technology, 1995.

[21] A. Sokolova and E. de Vink. Probabilistic automata: system types, parallel composition and comparison. In *Validation of Stochastic Systems: A Guide to Current Research*, volume 2925 of *Lecture Notes in Computer Science*, pages 1–43. Springer-Verlag, 2004.

[22] A. Tarski. A lattice-theoretical fixed point theorem and its applications. *Pacific Journal of Mathematics*, 5:285–309, 1955.

[23] F. van Breugel and J. Worrell. An algorithm for quantitative verification of probabilistic transition systems. In *Proceedings of the 12th International Conference on Concurrency Theory*, volume 2154 of *Lecture Notes in Computer Science*, pages 336–350. Springer-Verlag, 2001.

[24] F. van Breugel and J. Worrell. Towards quantitative verification of probabilistic transition systems. In *Proceedings of the 28th International Colloquium Automata, Languages and Programming*, volume 2076 of *Lecture Notes in Computer Science*, pages 421–432. Springer-Verlag, 2001.

[25] F. van Breugel and J. Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 2004. In press.

[26] R. J. van Glabbeek, S. A. Smolka, and B. Steffen. Reactive, generative, and stratified models of probabilistic processes. *Information and Computation*, 121(1):59–80, 1995.

[27] L. Vaserstein. Markov processes over denumerable products of spaces, describing large systems of automata. *Problems in Information Transmission*, 5(3):47–52, 1969.

[28] Y. Wang. CCS + time = an interleaving model for real time systems. In *Proceedings of the 18th Colloquium on Automata, Languages and Programming*, volume 510 of *Lecture Notes in Computer Science*, pages 217–228. Springer-Verlag, 1991.

[29] M. Ying. Bisimulation indexes and their applications. *Theoretical Computer Science*, 275(1/2):1–68, 2002.