Characterization of Simulation by Probabilistic Testing

Philipp Rümmer(✉) and Wang Yi

Department of Information Technology, Uppsala University, Uppsala, Sweden
philipp.ruemmer@it.uu.se

Abstract. Testing of systems naturally has a non-deterministic character: on the one hand, internal decisions of the system under test appear as non-determinism to an observer; on the other hand, the system under test inevitably receives inputs from the environment that are not controlled by the tester. To model both aspects, we investigate a probabilistic testing framework in which non-deterministic labelled transition systems are examined through execution of finite, probabilistic test-cases. We show that the simulation preorder on labelled transition systems can be tested probabilistically, elegantly recapturing the notion of conformance testing in this setting.

1 Introduction

For us, Frank is not only a great friend, but also a great scientist and a great leader. His contributions in many areas have been a source for inspiration in our work. His leadership has been a driving force in many large successful collaborating projects in Europe. Thank you Frank! Congratulations for the first successful 60 years; we look forward to working with you in the coming 60 years! The work presented in this paper was initiated many years ago when Frank was also a participant in a venue discussing research issues on concurrency and testing.

To study probabilistic phenomena such as randomisation and failure rates in distributed computing, significant research effort has been put into the extension of models and methods that have proven successful for non-probabilistic systems to the probabilistic setting. In the non-probabilistic setting, transition systems are well-established as a basic semantic model for sequential, concurrent, and distributed systems. This model has been extended in the literature to the probabilistic case by adding a mechanism for representing probabilistic choice.

In the work presented in this paper, we consider the specific combination of classical, *non-probabilistic* systems, examined with the help of *probabilistic* tests. More specifically, we consider tests as finite labelled transition systems that might contain both probabilistic and non-deterministic choice. As the main result, we show that the (non-probabilistic) simulation preorder can be tested by comparing the likelihood that probabilistic tests succeed. Probability, in this setting, is mainly used as a vehicle to examine the branching structure of processes, since probabilistic choice has the effect of copying and duplicating intermediate

E. Ábrahám et al. (Eds.): de Boer Festschrift, LNCS 9660, pp. 360–372, 2016.
DOI: 10.1007/978-3-319-30734-3_24

states of processes, in such a way that each copy can be examined separately. This concept has been exploited in a number of previous research results, including [1,2].

We outline how this theoretic result can practically be exploited in the context of *conformance testing,* where the relationship of a concrete implementation with an abstract behavioural specification is checked.

1.1 Related Work

Characterization of Bisimulation by Probabilistic Testing. Abramsky presented the first work in the 80s [1] to characterize bisimulation relations using probabilistic testing, which is the original motivation of this work. The essential idea of Abramsky is to utilize the "copying capability" in probability testing to characterize equivalence relations. In this work, we show that the copying feature can also be used to characterize simulation relation, which is a preorder. A relevant work along this line is [2], where we have shown that testing preorders can be characterized by simulation relations over probabilistic systems. The difference is that here we have probabilistic tests and the systems under test exhibit only non-deterministic behavior.

Statistical Model Checking. An area related to probabilistic testing is statistical model checking, which has been proposed as an alternative to exhaustive model checking for analyzing stochastic (e.g., timed or hybrid) systems [3,6]. In statistical model checking, the behavior of a system is *simulated,* thus obtaining a sample of possible system executions; afterwards, *hypothesis testing* is used to check whether the sample represents sufficient statistical evidence that some specification is satisfied or violated. In contrast to exhaustive methods, statistical model checking does not provide guarantees, but makes it possible to bound the likelihood of wrong answers. At the same time, runtime and memory consumption of statistical model checking can be drastically smaller than that of exhaustive techniques.

The results presented in this paper differ from statistical model checking methods in two important points: in our setting, it is the *tests* that are assumed to be probabilistic, whereas systems under test only exhibit non-deterministic behavior (in Sect. 3 and later); the situation in statistical model checking is the opposite. Second, we consider how testing is used to derive simulation relation between two systems, rather than checking that a system conforms to some independently defined property.

2 Preliminaries

We consider a model of probabilistic transition systems, containing probabilistic and non-deterministic choices as independent concepts. Processes, in most parts of the paper, are transition systems only containing non-deterministic choices, i.e., there is no probabilistic behaviour. In contrast, tests are defined as transition systems that can contain both non-deterministic and probabilistic behaviour, more precisely as finite trees in which certain states are “accepting.” As we will see, in this setting it is possible to give an exceptionally simple and elegant characterisation of *simulation* in terms of *tests.*

Most of the definitions follow the lines of [2].

2.1 Basic Concepts

A *weighting* on a set S is a function $\sigma : S \to \mathcal{R}_{\geq 0}$ from S to nonnegative real numbers. For a set S, we use $\sigma(S)$ to denote $\sum_{s \in S} \sigma(s)$. A *probability distribution* on a finite set S is a weighting σ on S such that $\sigma(S) = 1$. A *sub-distribution* on a finite set S is a weighting σ on S such that $\sigma(S) \leq 1$. We use $s \in \sigma$ to denote that $\sigma(s) > 0$. The *support* $\mathit{Supp}(\sigma)$ of a weighting σ is the set of elements s with $s \in \sigma$. A distribution whose support is a singleton set is called a *deterministic distribution.* Let $\mathit{Weight}(S)$ and $\mathit{Dist}(S)$ denote the sets of weightings and probability distributions on S, respectively. We will sometimes identify a single state s with the deterministic distribution that assigns probability 1 to s.

If σ is a weighting on S and ρ is a weighting on R, then $\sigma \times \rho$ is a weighting on $S \times R$, defined by $(\sigma \times \rho)(\langle s, r \rangle) = \sigma(s) * \rho(r)$. If σ is a weighting on S and $h : S \to R$ is a function from S to R, then $h(\sigma)$ is a weighting on R, defined by $h(\sigma)(r) = \sum_{h(s)=r} \sigma(s)$. If σ and ρ are weightings on S, then $\sigma \leq \rho$ denotes that $\sigma(s) \leq \rho(s)$ for all $s \in S$.

2.2 Probabilistic Transition Systems

We assume a finite set $\mathcal{A}ct$ of atomic actions, ranged over by a and b.

Definition 1. *A* (probabilistic) transition system *is a pair* $\langle S, \longrightarrow \rangle$*, where*

- S *is a non-empty finite set of* states*, and*
- $\longrightarrow \subseteq S \times \mathcal{A}ct \times \mathit{Dist}(S)$ *is a finite* transition relation*.*

We use $s \xrightarrow{a} \sigma$ *to denote that* $\langle s, a, \sigma \rangle \in \longrightarrow$*.*

A (probabilistic) process *is a tuple* $\langle \langle S, \longrightarrow \rangle, \sigma_0 \rangle$*, where* $\langle S, \longrightarrow \rangle$ *is a probabilistic transition system, and* $\sigma_0 \in \mathit{Dist}(S)$ *is an* initial probability distribution *on* S*.*

We write $s \xrightarrow{a}$ to denote that there is a σ such that $s \xrightarrow{a} \sigma$, and say that a state s is *terminal* (written $s \not\longrightarrow$) if there is no a and σ such that $s \xrightarrow{a} \sigma$. By slight abuse of notation, we write $s \xrightarrow{a} s'$ if $s \xrightarrow{a} \sigma$ such that $s' \in \sigma$. A *finite tree* is a process $\langle \langle S, \longrightarrow \rangle, \sigma_0 \rangle$ such that every state $s' \in S$ can be reached by exactly one path $s_0 \xrightarrow{a_1} s_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} s_n = s'$ with $s_0 \in \sigma_0$.

Each state of a probabilistic transition system has a potential for future dynamic behavior. When an action is performed, the system makes a probabilistic “choice” of next state. Thus, at each point in time, a snapshot of the system state will be a distribution over possible states.

2.3 Probabilistic Testing

To study testing, we define a synchronous parallel composition operator for probabilistic transition systems, in which two processes $\mathcal{P}$ and $\mathcal{Q}$ execute in parallel while synchronizing on all actions in $\mathcal{A}ct$.

Definition 2. *Let $\langle S, \longrightarrow\rangle$ and $\langle R, \longrightarrow\rangle$ be two transition systems. Their composition, denoted by the expression $\langle S, \longrightarrow\rangle \| \langle R, \longrightarrow\rangle$, is the transition system $\langle U, \longrightarrow\rangle$ where*

- *$U = S \times R$. A pair $(s, r) \in U$ is denoted $s\|r$.*
- *$\longrightarrow \subseteq U \times \mathcal{A}ct \times Dist(U)$ is defined by*

$$s\|r \xrightarrow{a} \sigma \times \rho \qquad \textit{iff} \qquad s \xrightarrow{a} \sigma \quad \textit{and} \quad r \xrightarrow{a} \rho$$

The composition of two processes $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ and $\mathcal{Q} = \langle\langle R, \longrightarrow\rangle, \rho_0\rangle$, denoted $\mathcal{P}\|\mathcal{Q}$, is the process $\langle\langle S, \longrightarrow\rangle \| \langle R, \longrightarrow\rangle, \sigma_0 \times \rho_0\rangle$.

Following Wang and Larsen [5], we define tests as finite trees with a certain subset of the terminal states being "accepting states."

Definition 3. *A* (probabilistic) test *is a tuple $\langle\langle\langle T, \longrightarrow\rangle, \tau_0\rangle, F\rangle$ in which the process $\langle\langle T, \longrightarrow\rangle, \tau_0\rangle$ is a finite tree, and $F \subseteq T$ is a set of* success states*, each of which is terminal.*

A test $\mathcal{T}$ is applied to a process $\mathcal{P}$ by putting the process $\mathcal{P}$ in parallel with the test $\mathcal{T}$ and measuring the likelihood of reaching a success state.

We define a testing system as the parallel composition of a process and a test.

Definition 4. *Let $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ be a process and $\mathcal{T} = \langle\langle\langle T, \longrightarrow\rangle, \tau_0\rangle, F\rangle$ be a test. The composition of $\mathcal{P}$ and $\mathcal{T}$, denoted $\mathcal{P}\|\mathcal{T}$, is called a* testing system*, defined as the process $\langle\langle S, \longrightarrow\rangle, \sigma_0\rangle \| \langle\langle T, \longrightarrow\rangle, \tau_0\rangle$ with success states $S \times F$.*

Our intention is that a testing system defines a probability of reaching a success state. However, since from each state there may be several outgoing transitions, such a probability is not uniquely defined. We will be interested in the maximal probabilities of success. These can be defined inductively on the structure of the testing system.

Definition 5. *Let $\mathcal{P}\|\mathcal{T}$ be a testing system, with a process $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ and test $\mathcal{T} = \langle\langle\langle T, \longrightarrow\rangle, \tau_0\rangle, F\rangle$. For each state $s\|t$ of $\mathcal{P}\|\mathcal{T}$ we define its* maximal probability of sucess*, denoted $t\lceil s\rceil$ inductively by*

- *If $s\|t$ is terminal, then $t\lceil s\rceil = 1$ if t is a success state, else $t\lceil s\rceil = 0$.*
- *If $s\|t$ is not terminal, then*

$$t\lceil s\rceil = \max_{s\|t \xrightarrow{a} \sigma \times \tau} \left(\sum_{s'\|t'} (\sigma \times \tau)(s'\|t') * t'\lceil s'\rceil \right)$$

For a distribution σ on S and a distribution τ on T, we define

$$\tau\lceil\sigma\rceil = \sum_{s\|t}(\sigma \times \tau)(s\|t) * t\lceil s\rceil$$

We define $\mathcal{T}\lceil\mathcal{P}\rceil = \sigma_0\lceil\tau_0\rceil$.

We note that, using the definition of $\tau\lceil\sigma\rceil$, we simplify the definition of $t\lceil s\rceil$ to

$$t\lceil s\rceil = \max_{s\|t \xrightarrow{a} \sigma \times \tau} \tau\lceil\sigma\rceil$$

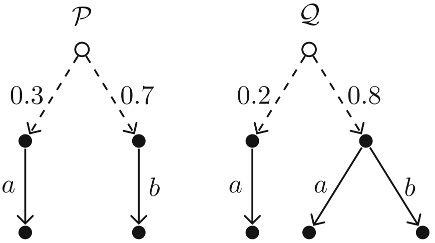
We now define a may-preorder of testing, which abstracts from the set of possible expected outcomes when testing a process $\mathcal{P}$ by a test $\mathcal{T}$: *may*-testing considers the highest possible expected outcome of $\mathcal{P}\|\mathcal{T}$.

Definition 6. *Given two processes $\mathcal{P}$ and $\mathcal{Q}$, define*

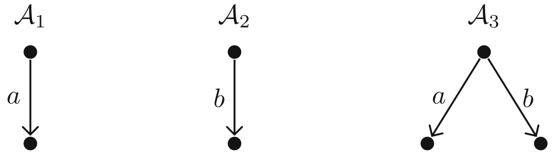
$$\mathcal{P} \sqsubseteq_t \mathcal{Q} \qquad \textit{iff} \qquad \forall\mathcal{T} : \mathcal{T}\lceil\mathcal{P}\rceil \leq \mathcal{T}\lceil\mathcal{Q}\rceil$$

The intention behind the definition of $\sqsubseteq_t$ is that intuitively, $\mathcal{P} \sqsubseteq_t \mathcal{Q}$ should mean that $\mathcal{P}$ refines $\mathcal{Q}$ with respect to "safety properties." The motivation is the following. We can regard the success states of a test as states defining when the tester has observed some "bad" or "unacceptable" behavior. A process then refines another one if it has a smaller potential for "bad behavior" with respect to any test. In the definition of $\mathcal{P} \sqsubseteq_t \mathcal{Q}$, this means that the maximal probability of observing bad behavior of $\mathcal{P}$ should not exceed the maximal probability of observing bad behavior of $\mathcal{Q}$.

Example 7. Consider the following processes $\mathcal{P}$ and $\mathcal{Q}$. The dashed arrows show the initial distribution of the processes, the straight arrows the (deterministic) transitions of the processes.



The probability that $\mathcal{P}$ may pass a test is always less or equal to the probability $\mathcal{Q}$ may pass the same test; therefore $\mathcal{P} \sqsubseteq_t \mathcal{Q}$. To see this, consider the sub-systems $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$:



Clearly, for any test $\mathcal{T}$ it is the case that $\mathcal{T}\lceil\mathcal{A}_1\rceil \leq \mathcal{T}\lceil\mathcal{A}_3\rceil$ and $\mathcal{T}\lceil\mathcal{A}_2\rceil \leq \mathcal{T}\lceil\mathcal{A}_3\rceil$. This implies that

$$\mathcal{T}\lceil\mathcal{P}\rceil = 0.3 \cdot \mathcal{T}\lceil\mathcal{A}_1\rceil + 0.7 \cdot \mathcal{T}\lceil\mathcal{A}_2\rceil \;\leq\; 0.2 \cdot \mathcal{T}\lceil\mathcal{A}_1\rceil + 0.8 \cdot \mathcal{T}\lceil\mathcal{A}_3\rceil = \mathcal{T}\lceil\mathcal{Q}\rceil.$$

3 Characterization of Simulation by Probabilistic Testing

In the following, we restrict our attention to non-probabilistic processes, but consider the analysis of such processes with the help of probabilistic tests. We call a process $\langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ *non-probabilistic* if σ_0 is a deterministic distribution, and if, likewise, σ is deterministic for every $\langle s, a, \sigma\rangle \in \longrightarrow$. The main result of this section is the relationship between the may-preorder for non-probabilistic processes, established through execution of probabilistic tests, and the classical notion of *simulation* [4]:

Definition 8 (Simulation). *Let $\langle S, \longrightarrow\rangle$ and $\langle R, \longrightarrow\rangle$ be two non-probabilistic transition systems. A* simulation relation *between $\langle S, \longrightarrow\rangle$ and $\langle R, \longrightarrow\rangle$ is a binary relation $W \subseteq S \times R$ such that, whenever $(s, r) \in W$ and $s \xrightarrow{a} s'$, there is a state $r' \in R$ with $r \xrightarrow{a} r'$ and $(s', r') \in W$. We say that a process $\langle\langle S, \longrightarrow\rangle, s_0\rangle$* simulates *a process $\langle\langle R, \longrightarrow\rangle, r_0\rangle$, denoted by $\langle\langle S, \longrightarrow\rangle, s_0\rangle \lhd \langle\langle R, \longrightarrow\rangle, r_0\rangle$, if there is a simulation relation W between $\langle S, \longrightarrow\rangle$ and $\langle R, \longrightarrow\rangle$ with $(s_0, r_0) \in W$.*

Lemma 9. *The relation $s \lhd r \;\equiv\; \langle\langle S, \longrightarrow\rangle, s\rangle \lhd \langle\langle R, \longrightarrow\rangle, r\rangle$ is the greatest simulation relation between the non-probabilistic transition systems $\langle S, \longrightarrow\rangle$ and $\langle R, \longrightarrow\rangle$.*

The simulation preorder is instrumental in various contexts, in particular (as discussed in the later sections of this paper) for checking the conformance of systems with behavioural specifications.

We are now able to give the main theorem of this section (and the paper), relating the may-preorder of testing with the classical simulation preorder. The result shows that the simulation preorder of non-probabilistic processes can be tested in a probabilistic setting, by considering tests possibly containing probabilistic choices.

Theorem 10 (Testability of simulation). *Suppose $\mathcal{P}, \mathcal{Q}$ are non-probabilistic processes. Then the following equivalence holds:*

$$\mathcal{P} \sqsubseteq_t \mathcal{Q} \qquad \textit{iff} \qquad \mathcal{P} \lhd \mathcal{Q}$$

For proving this theorem, we first need a number of intermediate results. We can first observe that every testing system gives rise to a finite set of *resolutions,* in which every state has an out-degree of at most one:

3.1 Linear Resolutions of Processes

Definition 11 (Linearity). *A finite tree $\langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ is called* linear *if $\sigma_0 = s_0$ is deterministic and every state has at most one outgoing transition:*

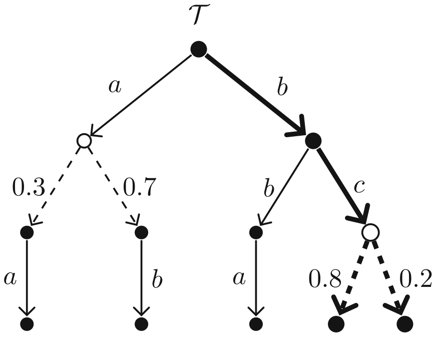
$$\textit{for each } s \in S: \quad s \xrightarrow{a} \sigma \textit{ and } s \xrightarrow{a'} \sigma' \quad \textit{imply} \quad a = a' \textit{ and } \sigma = \sigma'.$$

A linear resolution *of a finite tree $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ is a maximum linear subtree $\langle\langle S', \longrightarrow'\rangle, \sigma_0\rangle$ of $\mathcal{P}$, i.e., a linear tree consisting of maximum subsets of states $S' \subseteq S$ and transitions $\longrightarrow' \subseteq \longrightarrow$ of $\mathcal{P}$. The set of resolutions of a tree $\mathcal{P}$ is denoted by $Res(\mathcal{P})$.*

The notion of linear resolutions naturally extends to finite *acyclic* processes, i.e., to processes in which the length of paths $s_0 \xrightarrow{a_1} s_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} s_n$ is bounded. Note that, by definition of a tree, the resolution $\langle\langle S', \longrightarrow'\rangle, \sigma_0\rangle$ is closed under transitions: $Supp(\sigma_0) \subseteq S'$ and $Supp(\sigma) \subseteq S$ for each $\langle s, a, \sigma\rangle \in \longrightarrow'$. Maximality implies that a resolution does not have more terminal states than the original tree, i.e., $s \not\longrightarrow'$ implies $s \not\longrightarrow$ for any $s \in S'$.

Intuitively, if a state of a process has two outgoing transitions $s \xrightarrow{a} \sigma$ and $s \xrightarrow{a'} \sigma'$, any linear resolution of the process will contain at most one of the transitions, and remove the other one; if s is a state that is kept in the resolution, exactly one of the transitions will be kept. In the case of a finite non-probabilistic tree, resolutions correspond to maximum paths starting in the root of the tree.

Example 12. The following diagrams illustrates a linear resolution of a finite tree $\mathcal{T}$. The resolution is drawn bold:



Note that probabilistic choices are kept in a resolution, so that linear resolutions do not necessarily form simple chains of transitions.

The notion of a resolution leads to a more explicit characterisation of the maximum success probability of running a test:

Lemma 13. *Let $\mathcal{P} \| \mathcal{T}$ be a testing system, composed of process $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, \sigma_0\rangle$ and the test $\mathcal{T} = \langle\langle\langle T, \longrightarrow\rangle, \tau_0\rangle, F\rangle$. Then*

$$\mathcal{T}\lceil\mathcal{P}\rceil \;=\; \max_{R \in Res(\mathcal{P}\|\mathcal{T})} P(R)$$

where the success probability $P(R) = P_R(\sigma_0 \times \tau_0)$ of a test system resolution $R \in Res(\mathcal{P} \| \mathcal{T})$ is recursively defined by:

$$P_R(\sigma \times \tau) \;=\; \sum_{s\|t} (\sigma \times \tau)(s\|t) * P_R(s\|t)$$

$$P_R(s\|t) \;=\; \begin{cases} 1 & \textit{if } t \in F \textit{ is a success state} \\ 0 & \textit{if } s\|t \not\longrightarrow \textit{is a terminal state with } t \notin F \\ P_R(\sigma \times \tau) & \textit{if } s\|t \xrightarrow{a} \sigma \times \tau \textit{ (in the resolution } R\textit{)} \end{cases}$$

3.2 Necessary and Sufficient Conditions for the May-Preorder

It is unnecessary to consider the set of all tests for checking the may-preorder; rather, we can give necessary and sufficient conditions for the preorder by checking whether tests are guaranteed to succeed or not. These criteria will be helpful in proving the main Theorem 10 of the section:

Lemma 14. *For non-probabilistic processes $\mathcal{P}, \mathcal{Q}$:*

$$\mathcal{P} \sqsubseteq_t \mathcal{Q} \qquad \textit{iff} \qquad \forall \mathcal{T} : \big(\mathcal{T}\lceil\mathcal{P}\rceil = 1 \Longrightarrow \mathcal{T}\lceil\mathcal{Q}\rceil = 1\big)$$

Proof. "$\Longrightarrow$" By definition, $\mathcal{P} \sqsubseteq_t \mathcal{Q}$ means $\forall \mathcal{T} : \mathcal{T}\lceil\mathcal{P}\rceil \leq \mathcal{T}\lceil\mathcal{Q}\rceil$, which implies the right-hand side of the equivalence.

"$\Longleftarrow$" Proving by contradiction, we assume $\forall \mathcal{T} : \big(\mathcal{T}\lceil\mathcal{P}\rceil = 1 \Longrightarrow \mathcal{T}\lceil\mathcal{Q}\rceil = 1\big)$, but $\mathcal{P} \not\sqsubseteq_t \mathcal{Q}$, the latter of which implies that there is a test $\mathcal{T} = \langle\langle\langle T, \longrightarrow\rangle, \tau_0\rangle, F\rangle$ such that $\mathcal{T}\lceil\mathcal{P}\rceil > \mathcal{T}\lceil\mathcal{Q}\rceil$. According to Lemma 13, we can assume that $\mathcal{T}\lceil\mathcal{P}\rceil$ is realised by the resolution $R = \langle\langle S_R, \longrightarrow_R\rangle, \sigma_R\rangle \in Res(\mathcal{P}\|\mathcal{T})$, which means that the success probability of R is $P(R) = \mathcal{T}\lceil\mathcal{P}\rceil$.

We define a new test $\mathcal{T}' = \langle\langle\langle T', \longrightarrow'\rangle, \tau_0\rangle, F'\rangle$, in such a way that $\mathcal{T}'\lceil\mathcal{P}\rceil = 1$:

- $T' = \{t \in T \mid \exists s : s\|t \in S_R\}$ is the set of test states reachable in R;
- $\longrightarrow' = \{(t, a, \tau) \in \longrightarrow \mid \exists(s, a, \sigma) : (s\|t, a, \sigma \times \tau) \in \longrightarrow_R\}$ is the reduct of $\longrightarrow$ to transitions in R;
- $F' = \{t \in T' \mid \exists s : s\|t \not\longrightarrow_R\}$ are those test states that occur as final states in R.

To see that $\mathcal{T}'\lceil\mathcal{P}\rceil = 1$, observe that R also is a resolution of $\mathcal{P}\|\mathcal{T}'$; all terminal states of this resolution are success states.

Due to the assumption that $\forall \mathcal{T} : \big(\mathcal{T}\lceil\mathcal{P}\rceil = 1 \Longrightarrow \mathcal{T}\lceil\mathcal{Q}\rceil = 1\big)$, this implies $\mathcal{T}'\lceil\mathcal{Q}\rceil = 1$; in other words, also $\mathcal{Q}\|\mathcal{T}'$ has a resolution R' in which all terminal states are success states. This means, in particular, that all success states of $\mathcal{T}$ reached in R are also reached in R', because otherwise R' would contain paths not leading to success. But then also the test system $\mathcal{Q}\|\mathcal{T}$ has a resolution R'' containing at least all success states reached in R, which implies $P(R'') \geq P(R)$ and contradicts the assumption $\mathcal{T}\lceil\mathcal{P}\rceil > \mathcal{T}\lceil\mathcal{Q}\rceil$. □

Similarly, it would be sufficient to consider tests with success probability 0 to characterise the may-preorder:

Lemma 15. *For non-probabilistic processes $\mathcal{P}, \mathcal{Q}$:*

$$\mathcal{P} \sqsubseteq_t \mathcal{Q} \qquad \textit{iff} \qquad \forall \mathcal{T} : \big(\mathcal{T}\lceil\mathcal{Q}\rceil = 0 \Longrightarrow \mathcal{T}\lceil\mathcal{P}\rceil = 0\big)$$

3.3 The May-Preorder as Simulation

We prove the two directions of Theorem 10 separately. The more intricate proof concerns the observation that the may-preorder is a subset of the simulation preorder, which can be shown by induction over processes:

Lemma 16. *If $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, s_0\rangle$ and $\mathcal{Q} = \langle\langle R, \longrightarrow\rangle, r_0\rangle$ are non-probabilistic processes, then:*

$$\mathcal{P} \sqsubseteq_t \mathcal{Q} \qquad \textit{implies} \qquad \mathcal{P} \lhd \mathcal{Q}$$

Proof. We conduct a proof by contradiction, showing that $\mathcal{P} \ntrianglelefteq \mathcal{Q}$ implies $\mathcal{P} \not\sqsubseteq_t \mathcal{Q}$. Since $\ntriangleleft$ can be defined as a least fixed-point, we can prove the implication by means of induction over processes $\mathcal{P}, \mathcal{Q}$ not in simulation relation.

Assume $\mathcal{P} \ntriangleleft \mathcal{Q}$. Since $\lhd$ is the greatest simulation relation, this means that there is a transition $s_0 \xrightarrow{a} s'$, but for all a-transitions $r_0 \xrightarrow{a} r_1, \ldots, r_0 \xrightarrow{a} r_n$ of $\mathcal{Q}$ we have $\mathcal{P}' = \langle\langle S, \longrightarrow\rangle, s'\rangle \ntriangleleft \langle\langle R, \longrightarrow\rangle, r_i\rangle = \mathcal{Q}_i$ (for $i \in \{1, \ldots, n\}$). Together with the induction hypothesis and Lemma 14, this implies that there are tests $\mathcal{T}_1, \ldots, \mathcal{T}_n$ such that $\mathcal{T}_i\lceil\mathcal{P}'\rceil = 1$, but $\mathcal{T}_i\lceil\mathcal{Q}_i\rceil < 1$ for all $i \in \{1, \ldots, n\}$.

We construct a new test $\mathcal{T}$, in such a way that $\mathcal{T}\lceil\mathcal{P}\rceil = 1$, but $\mathcal{T}\lceil\mathcal{Q}\rceil < 1$. By Lemma 14, this implies $\mathcal{P} \not\sqsubseteq_t \mathcal{Q}$.

We assume $\mathcal{T}_i = \langle\langle\langle T_i, \longrightarrow_i\rangle, \tau_i\rangle, F_i\rangle$, and, without loss of generality, that the sets $(T_i)_{i=1}^n$ are pairwise disjoint. The test $\mathcal{T} = \langle\langle\langle T, \longrightarrow\rangle, t_0\rangle, F\rangle$ is defined by:

- $T = \{t_0\} \cup \bigcup_{i=1}^n T_i$, where t_0 is a fresh state not occurring in any of the sets T_i;
- $\longrightarrow \,=\, \{(t_0, a, \tau_a)\} \cup \left(\bigcup_{i=1}^n \longrightarrow_i\right)$, with τ_a being the distribution

$$\tau_a(t) \;=\; \begin{cases} \tau_i(t)/n & \text{if } t = t_i \\ 0 & \text{otherwise;} \end{cases}$$

- $F = \bigcup_{i=1}^n F_i$.

We then have $\mathcal{T}\lceil\mathcal{P}\rceil = 1$, since $\mathcal{T}_i\lceil\mathcal{P}'\rceil = 1$ for all $i \in \{1, \ldots, n\}$:

$$\begin{aligned} \mathcal{T}\lceil\mathcal{P}\rceil \;&=\; t_0\lceil s_0\rceil \;=\; \max_{s_0 \| t_0 \xrightarrow{b} \sigma \times \tau} \tau\lceil\sigma\rceil \\ &\geq\; \tau_a\lceil s'\rceil \;=\; \sum_{i=1}^n \frac{\tau_i\lceil s'\rceil}{n} \;=\; \sum_{i=1}^n \frac{\mathcal{T}_i\lceil\mathcal{P}'\rceil}{n} \;=\; \sum_{i=1}^n \frac{1}{n} \;=\; 1 \end{aligned}$$

Similarly, we can observe that $\mathcal{T}\lceil\mathcal{Q}\rceil < 1$:

$$\begin{aligned} \mathcal{T}\lceil\mathcal{Q}\rceil \;&=\; t_0\lceil r_0\rceil \;=\; \max_{r_0 \| t_0 \xrightarrow{b} \sigma \times \tau} \tau\lceil\sigma\rceil \\ &=\; \max_{i \in \{1, \ldots, n\}} \tau_a\lceil r_i\rceil \;=\; \max_{i \in \{1, \ldots, n\}} \sum_{j=1}^n \frac{\tau_j\lceil r_i\rceil}{n} \;\overset{(*)}{<}\; 1 \end{aligned}$$

At $(*)$, we make use of the fact that $\tau_j\lceil r_i\rceil \leq 1$ for all $i, j \in \{1, \ldots, n\}$, but in particular $\tau_i\lceil r_i\rceil = \mathcal{T}_i\lceil\mathcal{Q}_i\rceil < 1$ for $i \in \{1, \ldots, n\}$. □

The proof for the other direction of Theorem 10 proceeds by induction over tests:

Lemma 17. *If $\mathcal{P} = \langle\langle S, \longrightarrow\rangle, s_0\rangle$ and $\mathcal{Q} = \langle\langle R, \longrightarrow\rangle, r_0\rangle$ are non-probabilistic processes, then:*

$$\forall \mathcal{T} : \big(\mathcal{P} \lhd \mathcal{Q} \textit{ implies } \mathcal{T}\lceil\mathcal{P}\rceil \le \mathcal{T}\lceil\mathcal{Q}\rceil\big)$$

Proof. We prove the lemma by induction over tests $\mathcal{T} = \langle\langle\langle T, \longrightarrow\rangle, \tau_0\rangle, F\rangle$. Suppose $t_i \xrightarrow{a_i} \tau_i$ are all transitions outgoing from initial states $t_i \in \tau_0$, for $i \in \{1, \ldots, n\}$. For each $t \in T \setminus \mathit{Supp}(\tau_0)$, we can identify a sub-test $\mathcal{T}_t$ of $\mathcal{T}$ that has t as root.

Assuming $\mathcal{P} \lhd \mathcal{Q}$, the transitions outgoing from s_0 are $s_0 \xrightarrow{b_j} s_j$ (for $j \in \{1, \ldots, m\}$), and the transitions outgoing from r_0 are $r_0 \xrightarrow{c_l} r_l$ (for $l \in \{1, \ldots, k\}$). Due to $\mathcal{P} \lhd \mathcal{Q}$, we know that for every $j \in \{1, \ldots, m\}$ there is a $l_j \in \{1, \ldots, k\}$ such that $b_j = c_{l_j}$ and $\langle\langle S, \longrightarrow\rangle, s_j\rangle \lhd \langle\langle R, \longrightarrow\rangle, r_{l_j}\rangle$. By the induction hypothesis, it follows that $t\lceil s_j\rceil \le t\lceil r_{l_j}\rceil$ for all $t \in T \setminus \mathit{Supp}(\tau_0)$. From this we can derive $\mathcal{T}\lceil\mathcal{P}\rceil \le \mathcal{T}\lceil\mathcal{Q}\rceil$:

$$\mathcal{T}\lceil\mathcal{P}\rceil \;=\; \tau_0\lceil s_0\rceil \;=\; \sum_{t_0} \tau_0(t_0) \times t_0\lceil s_0\rceil \;\overset{(*)}{\le}\; \sum_{t_0} \tau_0(t_0) \times t_0\lceil r_0\rceil \;=\; \mathcal{T}\lceil\mathcal{Q}\rceil$$

At $(*)$, we use the following sub-derivation, for a state $t_0 \in \tau_0$:

$$\begin{aligned} t_0\lceil s_0\rceil \;&=\; \max_{s_0 \| t_0 \xrightarrow{a} \sigma \times \tau} \tau\lceil\sigma\rceil \;=\; \max_{\substack{i,j\\ t_i = t_0\\ a_i = b_j}} \tau_i\lceil s_j\rceil \;=\; \max_{\substack{i,j\\ t_i = t_0\\ a_i = b_j}} \sum_t \tau_i(t) * t\lceil s_j\rceil \\ &\le\; \max_{\substack{i,j\\ t_i = t_0\\ a_i = b_j}} \sum_t \tau_i(t) * t\lceil r_{l_j}\rceil \;=\; \max_{\substack{i,j\\ t_i = t_0\\ a_i = b_j}} \tau_i\lceil r_{l_j}\rceil \;\le\; \max_{\substack{i,l\\ t_i = t_0\\ a_i = c_l}} \tau_i\lceil r_l\rceil \;=\; t_0\lceil r_0\rceil \end{aligned}$$

This concludes the proof. □

Lemmas 16 and 17 together imply Theorem 10.

3.4 Linear Tests

Up to this point, we have considered tests as arbitrary finite trees that can, in particular, exhibit non-deterministic behaviour (transitions $t \xrightarrow{a} t_1$ and $t \xrightarrow{a} t_2$) or have states in which multiple actions are offered to the system under test (transitions $t \xrightarrow{a_1} t_1$ and $t \xrightarrow{a_2} t_2$). From the perspective of practical testing, both properties are somewhat unusual and can be difficult to implement. We show in this section that such a rich language of tests is in fact unnecessary, our main results (in particular Theorem 10) also hold if only *linear tests* (following Definition 11) are considered.

Definition 18. *Given two processes $\mathcal{P}$ and $\mathcal{Q}$, we define the* linear may-preorder *by:*

$$\mathcal{P} \sqsubseteq_t^l \mathcal{Q} \qquad \textit{iff} \qquad \forall \mathcal{T} : \Big(\mathcal{T} \textit{ is linear} \Longrightarrow \mathcal{T}\lceil\mathcal{P}\rceil \leq \mathcal{T}\lceil\mathcal{Q}\rceil \Big)$$

Lemma 19. *For non-probabilistic processes $\mathcal{P}, \mathcal{Q}$, the linear may-preorder coincides with the may-preorder:*

$$\mathcal{P} \sqsubseteq_t \mathcal{Q} \qquad \textit{iff} \qquad \mathcal{P} \sqsubseteq_t^l \mathcal{Q}$$

Proof. "$\Longrightarrow$" Holds since every linear test is a test.

"$\Longleftarrow$" There are different ways to prove the implication; importantly, it can be observed that the proof of Lemma 16 only requires linear tests to be constructed, from which the implication follows.

We give an independent proof by contradiction as well. Assume $\mathcal{P} \sqsubseteq_t^l \mathcal{Q}$, but $\mathcal{P} \not\sqsubseteq_t \mathcal{Q}$. By Lemma 14, this means that there is a test $\mathcal{T}$ such that $\mathcal{T}\lceil\mathcal{P}\rceil = 1$, but $\mathcal{T}\lceil\mathcal{Q}\rceil < 1$. Since $\mathcal{T}\lceil\mathcal{P}\rceil = 1$, by Lemma 13 there is a resolution $R \in \mathit{Res}(\mathcal{P}\|\mathcal{T})$ with $P(R) = 1$. In the same way as in the proof of Lemma 14, it is possible to derive a new, linear test $\mathcal{T}'$ from R with $\mathcal{T}'\lceil\mathcal{P}\rceil = 1$; in fact, $\mathcal{T}'$ is a linear resolution of $\mathcal{T}$.

From the assumption $\mathcal{P} \sqsubseteq_t^l \mathcal{Q}$, it follows that $\mathcal{T}'\lceil\mathcal{Q}\rceil = 1$. However, $\mathit{Res}(\mathcal{Q}\|\mathcal{T}') \subseteq \mathit{Res}(\mathcal{Q}\|\mathcal{T})$, which (by Lemma 13) implies $\mathcal{T}\lceil\mathcal{Q}\rceil = 1$, contradicting the assumption $\mathcal{P} \not\sqsubseteq_t \mathcal{Q}$. □

Using Lemma 19 and Theorem 10, we can derive a stronger form of our main theorem:

Theorem 20 (Linear testability of simulation). *For non-probabilistic processes $\mathcal{P}, \mathcal{Q}$, the following equivalence holds:*

$$\mathcal{P} \sqsubseteq_t^l \mathcal{Q} \qquad \textit{iff} \qquad \mathcal{P} \lhd \mathcal{Q}$$

4 Probabilistic Conformance Testing

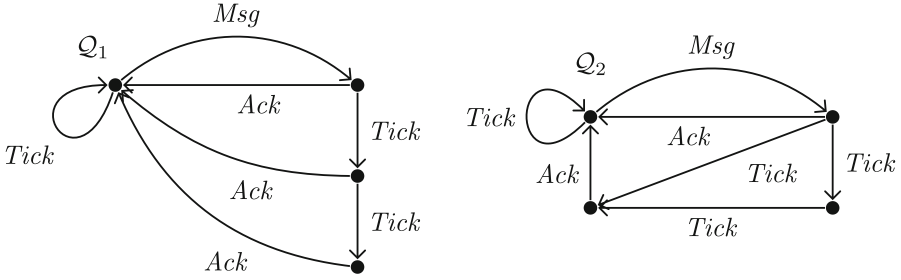
Conformance testing is concerned with checking that a system (or a piece of software) behaves correctly with respect to a given specification or standard. Many well-known applications of testing, for instance the verification of partial functional properties, can be considered as a part of conformance testing. Since conformance can pertain to safety- and security-critical aspects, as well as to contractual commitments, it is of great practical importance when developing systems.

A common setup for conformance testing is that of *black-box testing,* which means that implementation details of the system under test (SUT) are not taken into account during the testing process. In this scenario, the SUT is executed for a finite (but large) set of concrete test inputs, observing the responses of the system, in order to answer (with high confidence) the question whether the SUT conforms with a given specification.

We discuss how probabilistic testing of simulation relations, developed in the last sections, can be used to formally capture this kind of testing. There are typically a number of sources of non-determinism that have to be considered:

– the SUT might appear to behave non-deterministically, due to internal mechanisms (like a scheduler) that are not visible to the environment;
– the specification can be non-deterministic, in order to describe a whole set of scenarios of system execution, and in order to allow some degree of freedom in the behavior of the SUT;
– the set of considered concrete tests can be generated randomly, according to some chosen distributions, and depending on the responses given by the SUT.

Example 21. We consider the following, simplistic model $\mathcal{Q}_1$ of a server communicating with its environment using the messages *Msg* (sent to the server) and *Ack* (sent by the server). We adopt a discrete model of time and assume the presence of a further action *Tick*, expressing that one unit of time has passed. In the initial state $\mathcal{Q}_1$, the server is expected to remain silent until it has received *Msg*; then, after at most two *Tick*s, the server is supposed to respond with an *Ack*, returning to the state $\mathcal{Q}_1$:



As a specification of an actual implementation $\mathcal{P}$ of such a server, it could be required that $\mathcal{P}$ *simulates* the model $\mathcal{Q}_1$, i.e., $\mathcal{P} \lhd \mathcal{Q}_1$. Note that this kind of specification is able to capture very intricate behavioral properties related to the branching structure of a system. For instance, the model $\mathcal{Q}_2$ mostly coincides with $\mathcal{Q}_1$, but is stronger since it requires the server to decide about the delay before sending *Ack* at an earlier point ($\mathcal{Q}_2 \lhd \mathcal{Q}_1$, but $\mathcal{Q}_1 \not\lhd \mathcal{Q}_2$). Also, note that we disregard probabilistic aspects both of the implementation and the specification; while either might behave non-deterministically, we do not specify or check the distribution of behavior. □

4.1 Random Testing of Simulation Relations

A methodology for testing whether a SUT simulates a process (given as specification) can be as follows:

1. A number of linear, non-probabilistic tests is generated, and for each of the tests it is checked whether the SUT $\mathcal{P}$ passes the test (considering

the unique terminal state of the test as success state). This yields a multiset $O \subseteq \mathcal{A}ct^* \times \mathbb{B}$, recording both the sequences of input/output actions, and the test outcomes. The number of tests in O with positive outcome determines the overall success rate $s_{\mathcal{P}}$ of the SUT.

2. The set O is summarized as a single linear test $\mathcal{T}$, using the distribution of tests in O to synthesize probabilities.
3. The measured success rate $s_{\mathcal{P}}$ is compared with the maximum success probability $\mathcal{T}\lceil\mathcal{Q}\rceil$ predicted by the specification. Since $s_{\mathcal{P}}$ can be considered as a lower bound of the precise maximum success probability $\mathcal{T}\lceil\mathcal{P}\rceil$ (for a sufficiently large number of tests), a result $s_{\mathcal{P}} > \mathcal{T}\lceil\mathcal{Q}\rceil$ is an indication for $\mathcal{T}\lceil\mathcal{P}\rceil > \mathcal{T}\lceil\mathcal{Q}\rceil$, and by Theorem 10 for $\mathcal{P} \ntrianglelefteq \mathcal{Q}$.

5 Conclusions

We have shown that the simulation relation between non-probabilistic processes can be characterised through probabilistic testing, and outlined how this result might be useful for the purpose of conformance testing on non-deterministic processes. It is planned to study this latter application on more detail, and evaluate how tools for property-based random testing can be used to implement the conformance testing approach in practice.

References

1. Abramsky, S.: Observation equivalence as a testing equivalence. Theor. Comput. Sci. **53**, 225–241 (1987)
2. Jonsson, B., Yi, W.: Testing preorders for probabilistic processes can be characterized by simulations. Theor. Comput. Sci. **282**(1), 33–51 (2002)
3. Legay, A., Delahaye, B., Bensalem, S.: Statistical model checking: an overview. In: Barringer, H., Falcone, Y., Finkbeiner, B., Havelund, K., Lee, I., Pace, G., Roşu, G., Sokolsky, O., Tillmann, N. (eds.) RV 2010. LNCS, vol. 6418, pp. 122–135. Springer, Heidelberg (2010)
4. Milner, R.: Communication and Concurrency. Prentice-Hall Inc., Upper Saddle River (1989)
5. Yi, W., Larsen, K.G.: Testing probabilistic and nondeterministic processes. In: Proceedings of the IFIP TC6/WG6.1 Twelth International Symposium on Protocol Specification, Testing and Verification XII, pp. 47–61. North-Holland Publishing Co., Amsterdam (1992)
6. Younes, H.L.S., Simmons, R.G.: Statistical probabilistic model checking with a focus on time-bounded properties. Inf. Comput. **204**(9), 1368–1409 (2006)