Symbolic protocol verification with dice: process equivalences in the presence of probabilities

Vincent Cheval
Inria Paris

Raphaëlle Crubillé
Aix-Marseille Université
LIS, CNRS

Steve Kremer
Université de Lorraine
Inria Nancy Grand-Est & LORIA

***Abstract*—Symbolic protocol verification generally abstracts probabilities away, considering computations that succeed only with negligible probability, such as guessing random numbers or breaking an encryption scheme, as impossible. This abstraction, sometimes referred to as the perfect cryptography assumption, has shown very useful as it simplifies automation of the analysis. However, probabilities may also appear in the control flow where they are generally not negligible. In this paper we consider a framework for symbolic protocol analysis with a probabilistic choice operator: the probabilistic applied pi calculus. We define and explore the relationships between several behavioral equivalences. In particular we show the need for randomized schedulers and exhibit a counter-example to a result in a previous work that relied on non-randomized ones. As in other frameworks that mix both non-deterministic and probabilistic choices, schedulers may sometimes be unrealistically powerful. We therefore consider two subclasses of processes that avoid this problem. In particular, when considering purely non-deterministic protocols, as is done in classical symbolic verification, we show that a probabilistic adversary has—maybe surprisingly—a strictly superior distinguishing power for may testing, which, when the number of sessions is bounded, we show to coincide with purely possibilistic similarity.**

***Index Terms*—Security protocols, symbolic verification, probabilistic process equivalences.**

1. Introduction

Automated symbolic protocol verification, based on the seminal work of Dolev and Yao [1], has nowadays reached a level of maturity enabling successful use on complex real-world security protocols, including TLS [2], [3], Signal [4], authentication protocols of the 5G standard [5], or EMV's secure payment protocols [6] to name only a few. In the symbolic model, a non-deterministic, computationally unbounded attacker is assumed to have complete control of the network, being able to intercept any messages, and forge new ones. As a counterpart, cryptography is *idealized* and the attacker can only use predefined rules to manipulate messages that are represented by terms, *e.g.*, expressed by an equation $dec(enc(m, k), k) = m$ stating that a message m encrypted with k can be decrypted with the same key. This treatment of cryptography is in opposition to computational models where we assume a probabilistic polynomial time attacker, messages are represented by bitstrings and assumptions that an arbitrary such attacker has at most *negligible* probability of breaking a cryptographic primitive. Similarly, in the symbolic model, random values, such as keys or nonces, are chosen freshly from an infinite domain, rather than chosen randomly from a sufficiently large domain. These symbolic abstractions of cryptography and randomness have even been shown sound [7] (under rather strong assumptions) and significantly ease the automation of proofs. Hence, symbolic modeling of messages is arguably useful for formally analyzing cryptographic protocols.

However, the above-described abstractions of randomness only apply to the *messages*, and not to the *control flow*. Typical examples which crucially rely on randomized control flow are mechanisms for providing anonymity, such as the dining cryptographers protocol [8], mix-nets [9] or Crowds [10]. In this paper, we will investigate indistinguishability properties, expressed as equivalences in a cryptographic process calculus, the applied π-calculus [11], extended with a probabilistic choice operator. Typically, the testing equivalence expresses that two processes are equivalent if they exhibit the same behaviour when put in parallel with an arbitrary attacker process. Our work presents foundations for a model that *(i)* extends the scope of symbolic protocol analysis to probabilistic protocols, and *(ii)* allows to consider a probabilistic attacker (even on non-probabilistic protocols). In particular, when we consider purely concurrent processes–without probabilistic behavior–the equivalence we obtain is strictly stronger than the standard testing equivalence on such purely concurrent processes; in other terms, probabilistic adversaries are–for good reasons, as we will argue–more powerful in order to distinguish such processes than the purely concurrent adversaries considered in existing works and tools.

Our contributions. Our contributions can be split into three parts.

In a first part we introduce a probabilistic applied π-calculus and its semantics, which has similarities to [12], with two major differences. *(i)* We express our semantics in terms of general non-deterministic probabilistic transition

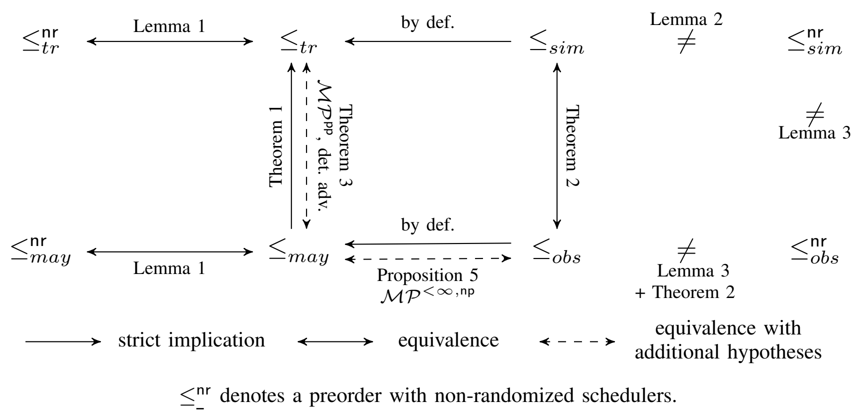


Figure 1: Summary of the relationship between preorders.

systems (NPLTS)–also called probabilistic automata in the literature–which allows us to benefit from a large body of existing results on these systems [13], [14], [15], [16], [17], [18], [19]. *(ii)* More importantly, we differ in the way non-determinism is resolved: unlike [12] we allow for *randomized* schedulers—rather than choosing one particular non-deterministic choice, we allow the scheduler to choose an arbitrary distribution on the available non-deterministic choices.

Second, we define several notions of preorders and equivalences and study their relations. The main results are also summarized in Figure 1, focusing on preorders (with similar relations between corresponding equivalences). We show, in particular, that

- unlike in the purely non-deterministic case, the may-testing preorder ($\leq_{may}$) is strictly stronger than the trace equivalence preorder ($\leq_{tr}$) (Theorem 1);
- simulation ($\leq_{sim}$) and observational pre-order ($\leq_{obs}$), respectively bisimilarity and observational equivalence, coincide for randomized schedulers (Theorem 2);
- for non-randomized schedulers, these equivalences ($\leq^{\mathsf{nr}}_{sim}$ and $\leq^{\mathsf{nr}}_{obs}$) do *not* coincide (Lemma 3), which provides a counter-example to one of the main results in [12].

Third, a well-known phenomenon [20], [21] in process calculi that are both probabilistic and non-deterministic is the existence of some *nonrealistic schedulers* that are able to use the internal probabilistic choices done by an agent in order to schedule another agent's non-deterministic choices, *i.e.*, the scheduler leaks the probabilistic choices. Therefore, we study two important *subclasses* of processes that avoid this phenomenon.

We first consider the classical class of *non-probabilistic processes* (denoted $\mathcal{MP}^{\mathsf{np}}$), as in the original applied π-calculus, but in the presence of probabilistic adversaries. We show that, if we additionally bound the number of sessions (denoted $\mathcal{MP}^{<\infty,\mathsf{np}}$),

- may-testing with probabilistic adversaries coincides with the classical, purely possibilistic notion of similarity (Proposition 5 and Theorem 2). This also provides a contextual characterization of the notion of similarity which is reminiscent of [17] in the setting of CSP;
- verification of testing equivalence with probabilistic adversaries is co-NEXPTIME complete for a large class of cryptographic primitives, relying on results from [22].

We next consider a class of *purely probabilistic* processes with a bounded number of sessions (denoted $\mathcal{MP}^{\mathsf{pp}}$), which is reminiscent of a probabilistic version of simple processes in [23], [24], and a slight generalization of the processes in [25]. We show that trace equivalence as considered in [25], is

- weaker than may-testing, but
- coincides with a version of may-testing with *determinate attackers*: attacker processes are restricted disallowing replication, parallel, and non-deterministic choice, but allowing probabilistic choices (Theorem 3).

Finally, we briefly discuss how the algorithm for deciding trace equivalence in the DeepSec verifier [22] could be adapted to this fully probabilistic case, providing a more general setting than Bauer *et al.* [26] who additionally bound the size of input messages.

A full version with detailed proofs is available at [27].

2. Probabilistic Applied π-calculus

In this section we introduce the probabilistic applied π-calculus, a probabilistic variant of the applied π-calculus introduced by Goubault-Larrecq *et al.* [12].

2.1. Message as terms

Atomic values such as keys and nonces are modelled by *names*. We assume an infinite set of such names $\mathcal{N} = \{a, b, \ldots, \}$ and partition it into two disjoint infinite sets $\mathcal{N}_{pub}$ and $\mathcal{N}_{priv}$. The set of *private* names $\mathcal{N}_{priv}$ is a priori unknown to the attacker and models, *e.g.*, honest keys in a protocol. The set of *public* names $\mathcal{N}_{pub}$ models public values, known to the attacker. The distinction between public and private names is analogous to the distinction between free and bound names in the original applied pi calculus. We also define an infinite set of variables $\mathcal{X}$. Finally, we consider a finite set of *function symbols* each equipped with their arity $\mathcal{F} = \{f/n, g/m, \ldots\}$. Function symbols model cryptographic operations, *e.g.*, $enc/2$ is a binary symbol that could be used to model encryption. *Terms* are defined as names, variables, and function symbols applied to other terms. For instance, given two names $a, k \in \mathcal{N}$, $enc(a, k)$ represents the encryption of a with the key k. For any $\mathsf{F} \subseteq \mathcal{F}$, $\mathsf{N} \subseteq \mathcal{N}$ and $\mathsf{V} \subseteq \mathcal{X}$, the set of terms built from N and V by applying function symbols in F is denoted by $\mathcal{T}(\mathsf{F}, \mathsf{N} \cup \mathsf{V})$.

We also suppose that terms are equipped with a binary relation $\doteq$ that expresses that two terms evaluate to the same result, and a predicate $Msg(\cdot)$ that is intended to hold when evaluation succeeds. How $\doteq$ and $Msg(\cdot)$ are precisely defined is not relevant for the results of this paper and we wish to capture several formalisms. $\doteq$ can for instance be defined by an equational theory, as in the applied π-calculus [11] (where $Msg(\cdot)$ would evaluate to true on any term), by a constructor-destructor rewrite system, allowing evaluation to fail when a destructor application does not reduce, as in the DeepSec tool [22], or a combination of these as in the ProVerif tool [28].

Formally, we require that $\doteq$ is symmetric, transitive, and closed under substitution of names and variables by other terms and application of function symbols. Moreover, for all $a, b \in \mathcal{N}$, $a = b$ if and only if $a \doteq b$. $Msg(\cdot)$ is supposed to hold on any names, be closed under renamings and $t_1 \doteq t_2$ implies that $Msg(t_1)$ and $Msg(t_2)$. Finally, we require that $Msg(t)$ implies $t \doteq t$.

For example, the $\doteq$ relation could capture that $dec(enc(m, k), k) \doteq m$ for any m, k modelling that decryption cancels out encryption when the same key k is used; one may also define $Msg(dec(n, k))$ as false to express that decryption fails if the ciphertext argument is not an encryption with the matching key.

2.2. Syntax of the process calculus

The syntax for *processes* is defined as follows:

P, Q ::=	processes
0	nil
$\mathsf{in}(u, x); P$	output
$\mathsf{out}(u, v); P$	input
$P \mid Q$	parallel composition
$!P$	replication
$\mathsf{new}\ a; P$	restriction
$\mathsf{if}\ u = v\ \mathsf{then}\ P\ \mathsf{else}\ Q$	conditional
$P + Q$	non-deterministic choice
$P +_p Q$	probabilistic choice

where $u, v \in \mathcal{T}(\mathcal{F}, \mathcal{N} \cup \mathcal{X})$, $x \in \mathcal{X}$, $a \in \mathcal{N}$ and $p \in]0; 1[$. As usual, in examples we will omit trailing 0 processes and else 0 branches. A process P is closed when all variables in P are bound by an input.

Example 1. As an example, consider the process P:

$$(\mathsf{out}(c, k) +_{\frac{1}{3}} \mathsf{out}(c, a)) \mid \mathsf{in}(c, x); \mathsf{if}\ x = k\ \mathsf{then}\ \mathsf{out}(c, ok)$$

P consists of two parallel processes. The left process outputs on a channel c with probability $\frac{1}{3}$ the name k and with probability $\frac{2}{3}$ the name a. The right process inputs a value on channel c and binds this value to x. If x equals k then it outputs the constant ok.

We denote by $\mathcal{SP}$ the set of all processes in the probabilistic applied π-calculus, and by $\mathcal{MP}$ the set of all multisets over $\mathcal{SP}$.

2.3. Operational semantics

We will now define the semantics of the probabilistic applied π-calculus. We opt for a different presentation of the semantics than Goubault-Larrecq *et al.* [12] relying on existing formalisms for transition systems. Moreover, we allow for a more general class of schedulers.

Notation 1. Let $\mathcal{S}$ be an arbitrary set. We denote by $\mathcal{D}(\mathcal{S})$ the set of all finitely supported probability distributions over $\mathcal{S}$ and by $\mathcal{D}^{\leq 1}(\mathcal{S})$ the set of all sub-probability distributions over $\mathcal{S}$ (observe that $\mathcal{D}(\mathcal{S}) \subseteq \mathcal{D}^{\leq 1}(\mathcal{S})$). For $p, q \geq 0$, and D, E two sub-distributions, we define the measure

$$(p \cdot D + q \cdot E)(x) = p \cdot D(x) + q \cdot E(x).$$

When $q = 0$, the resulting sub-distribution does not depend on E, and we simply write $p \cdot D$ instead of $p \cdot D + 0 \cdot E$.

If $D \in \mathcal{D}(\mathcal{S})$, we denote by $\mathrm{supp}(D)$ the support of D, *i.e.*, the set of all elements $s \in \mathcal{S}$ such that $D(s) > 0$. If $\mathcal{S}' \subseteq \mathcal{S}$, we define $D(\mathcal{S}') = \sum_{s \in \mathcal{S}'} D(s)$. Finally, we denote by δ_x the Dirac distribution on x.

The operational semantics of processes is defined by a relation between multisets of processes and probability distributions on multisets of processes, denoted $\mathcal{P} \rightarrow_\tau \mu$. This relation is defined in Figure 2.

Remark 1. One may note that our calculus offers a non-deterministic choice operator that is resolved internally. This differs from the standard pi-calculus [29] where the non-deterministic choice operator is resolved externally. Note that the original applied pi calculus [11] does not contain non-deterministic choice.

In the following, we define the operational semantics of our calculus using well studied probabilistic systems. We choose the formalism of *non-deterministic probabilistic labelled transition systems* (NPLTS) used for instance in [16]. A NPLTS allows to represent states that allow both *internal* and *external* non-deterministic behavior. It can be noted that it coincides with the notion of *simple probabilistic automata* of Segala *et al.* [13].

Definition 1. A NPLTS is a triple $(\mathcal{S}, \mathcal{A}, \mathsf{trans})$, where

- $\mathcal{S}$ is a set of states,
- $\mathcal{A} = \{\tau\} \sqcup \mathcal{A}_{ext}$ is a set of labels, and

$$
\begin{aligned}
&\mathcal{P} \cup \{\!\!\{0\}\!\!\} \to_\tau \delta_{\mathcal{P}} \\
&\mathcal{P} \cup \{\!\!\{\mathsf{if}\ u = v\ \mathsf{then}\ P\ \mathsf{else}\ Q\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{P\}\!\!\}} \text{ if } u \doteq v \\
&\mathcal{P} \cup \{\!\!\{\mathsf{if}\ u = v\ \mathsf{then}\ P\ \mathsf{else}\ Q\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{Q\}\!\!\}} \text{ if } u \not\doteq v \\
&\mathcal{P} \cup \{\!\!\{\mathsf{out}(u,t).P, \mathsf{in}(v,x).Q\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{P, Q\{x \mapsto t\}\}\!\!\}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{if } Msg(t) \wedge u \doteq v \\
&\mathcal{P} \cup \{\!\!\{P \mid Q\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{P, Q\}\!\!\}} \\
&\mathcal{P} \cup \{\!\!\{!P\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{!P, P\}\!\!\}} \\
&\mathcal{P} \cup \{\!\!\{\mathsf{new}\ a; P\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{P\{^{a'}/_{a}\}\}\!\!\}} \\
&\qquad\qquad\qquad\qquad\qquad \text{where } a' \in \mathcal{N}_{priv} \text{ is fresh} \\
&\mathcal{P} \cup \{\!\!\{P + Q\}\!\!\}) \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{P\}\!\!\}} \\
&\mathcal{P} \cup \{\!\!\{P + Q\}\!\!\} \to_\tau \delta_{\mathcal{P} \cup \{\!\!\{Q\}\!\!\}} \\
&\mathcal{P} \cup \{\!\!\{P +_p Q\}\!\!\} \to_\tau p \cdot \delta_{\mathcal{P} \cup \{\!\!\{P\}\!\!\}} + (1-p) \cdot \delta_{\mathcal{P} \cup \{\!\!\{Q\}\!\!\}}
\end{aligned}
$$

Figure 2: Semantics of the calculus

- $\mathsf{trans} : \mathcal{S} \to \mathcal{A} \to \mathcal{P}(\mathcal{D}(\mathcal{S}))$ is a transition function: for each state in $\mathcal{S}$, and each label in $\mathcal{A}$, $\mathsf{trans}(s)(a)$ is a set of (finitely supported) distributions.

The label τ is the *internal action* and the labels in $\mathcal{A}_{ext}$ are the *external* actions. For $s \in \mathcal{S}, a \in \mathcal{A}$, we write $s \xrightarrow{a} D$ when $D \in \mathsf{trans}(s,a)$.

In the remaining of this paper, we may define a NPLTS by only its transition function, *i.e.*, we will say that $\mathsf{trans} : \mathcal{S} \to \mathcal{A} \to \mathcal{P}(\mathcal{D}(\mathcal{S}))$ is the NPLTS $(\mathcal{S}, \mathcal{A}, \mathsf{trans})$.

We now express our operational semantics as a NPLTS without external actions, *i.e.*, $\mathcal{A}_{ext} = \varnothing$. External actions will be used to express our labeled semantics in Section 4.1.

Definition 2. The *operational semantics* is the NPLTS $\mathsf{N}^o = (\mathcal{MP}, \{\tau\}, \mathsf{trans}^o)$ where for every $s \in \mathcal{MP}$, $\mathsf{trans}^o(s)(\tau) = \{D \mid s \to_\tau D\}$.

Note that the states of the NPLTS N^o contain *all* possible multisets of processes and how they are executed. Obviously, N^o is thus an infinite transition system. In examples illustrating transitions of a multiset of processes $\mathcal{P}$, we only show the relevant fragment of N^o that contains $\mathcal{P}$.

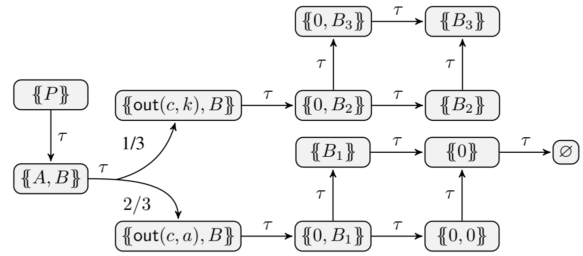
Example 2. The complete execution of the process P, introduced in Example 1 is given in Figure 3.

3. Behavioral equivalences

In this section we define probabilistic versions of two classical equivalences: *may-testing* and the stronger *observational equivalence*. In order to do so we first introduce the notion of *resolution* (also known as scheduler), *i.e.*, how internal non-determinism is resolved, and the notion of *barb*, which models an observational action.

3.1. Resolving the internal non-determinism

Resolutions express how internal non-determinism of a NPLTS is resolved. Intuitively, resolving the non-determinism means restricting the transition system by choosing for each state either one particular internal transition or leave the choice of a non-deterministic external action. The resulting transition system is called a *Reactive Probabilistic Labelled Transition System* (RPLTS) and has still external, but no internal, non-determinism. It can be noted that this model is equivalent to Labelled Markov Chains when extended with internal actions.



$A = \mathsf{out}(c,k) +_{\frac{1}{3}} \mathsf{out}(c,a)$ $B_2 = \mathsf{if}\ k = k\ \mathsf{then}\ \mathsf{out}(c, ok)$

$B = \mathsf{in}(c,x); \mathsf{if}\ x = k\ \mathsf{then}\ \mathsf{out}(c, ok)$ $B_3 = \mathsf{out}(c, ok)$

$B_1 = \mathsf{if}\ a = k\ \mathsf{then}\ \mathsf{out}(c, ok)$

Figure 3: Semantics of the process P from Example 2

Definition 3. A RPLTS is a triple $(\mathcal{S}, \mathcal{A}, \mathsf{trans})$, where

- $\mathcal{S}$ is a set of states,
- $\mathcal{A} = \{\tau\} \sqcup \mathcal{A}_{ext}$ a set of labels, and
- $\mathsf{trans} : \mathcal{S} \to \mathcal{D}(\mathcal{S}) \sqcup (\mathcal{A}_{ext} \to \mathcal{D}(\mathcal{S}) \cup \{\star\})$ is a transition function that assigns to each state in $\mathcal{S}$
  - either a unique distribution for the label τ (the deterministic internal action);
  - or a function mapping labels in $\mathcal{A}_{ext}$ to a failure ($\star$) or a distribution over $\mathcal{S}$ (the non deterministic external actions).

States $s \in \mathcal{S}$ such that $\mathsf{trans}(s) : \mathcal{A}_{ext} \to \mathcal{D}(\mathcal{S}) \cup \{\star\}$ are called *external states*, while the ones such that $\mathsf{trans}(s) : \mathcal{D}(\mathcal{S})$ are called *internal states*. Given a RPLTS R, we denote by $\mathcal{S}_{ext}(R)$ and $\mathcal{S}_{int}(R)$ the sets of external and internal states of R respectively. For a more homogeneous notation, when s is an internal state, we sometimes write $\mathsf{trans}(s)(\tau) = D$ instead of $\mathsf{trans}(s) = D$.

Remark 2. In the particular case of a NPLTS N with no external action, resolving the internal non-determinism results in a RPLTS *without any* non-determinism. This is the case of the operational semantics N^o. Such a purely probabilistic system is typically equivalent to the notion of Markov Chain. By abuse of notation, the transition function

$$\mathsf{trans} : \mathcal{S} \to \mathcal{D}(\mathcal{S}) \sqcup (\varnothing \to \mathcal{D}(\mathcal{S}) \cup \{\star\})$$

of such RPLTS is rewritten as

$$\mathsf{trans} : \mathcal{S} \to \mathcal{D}(\mathcal{S}) \sqcup \{\star\}$$

as for any set X, the cardinality of the set $(\varnothing \to X)$ is 1.

Before defining the notion of *resolution*–or schedulers–, we need to introduce two classical notions in probabilistic models: the *convex hull* of a set of distributions and the *probabilistic lifting* of a function.

Notation 2. Let $\mathcal{S}$ be a set of states. The *convex hull* of $\Delta \subseteq \mathcal{D}(\mathcal{S})$, denoted $\mathrm{conv}(\Delta)$, is the set of distributions $D \in \mathcal{D}(\mathcal{S})$ such that $\exists \alpha_1, \ldots, \alpha_n \in \mathbb{R}. \exists D_1, \ldots, D_n \in \Delta$.

$$\sum_{i=1}^{n} \alpha_i = 1 \text{ and } D = \sum_{i=1}^{n} \alpha_i \cdot D_i$$

Intuitively, rather than choosing one distribution in Δ, each element in $\mathrm{conv}(\Delta)$ corresponds to a distribution over the distributions in Δ. This will be useful for defining randomized schedulers.

Next, we lift functions to distributions: applying a function f to a distribution simply defines a new distribution that transfers, according to f, the probability weight of elements in the domain of f to its image, possibly summing these weights when f maps several inputs to a same output.

Notation 3. Let $\mathcal{S}, \mathcal{S}'$ be two sets of states and $f : \mathcal{S} \to \mathcal{S}'$. We define the function $\overline{f} : \mathcal{D}(\mathcal{S}) \to \mathcal{D}(\mathcal{S}')$ to be the *probabilistic lifting of* f, where

$$\overline{f}(D) = \sum_{s \in \mathcal{S}} D(s) \cdot \delta_{f(s)}$$

When obvious from context, we will overload the notation and write f instead of $\overline{f}$.

We now define *resolutions* for a NPLTS that allow to solve the *internal*, but not external, non-determinism: a resolution describes *one* of the possible ways of turning an NPLTS into a RPLTS. It means that for each state s, a resolution should choose whether s is an internal state or external state; in the first case, a *unique* post-transition distribution must be chosen; in the second case, for each external action a, the resolution must choose to either stop (*i.e.*, $\mathsf{trans}(s) = \star$) or a *unique* distribution D such that $s \xrightarrow{a} D$ (*i.e.*, $\mathsf{trans}(s) = D$). Due to the possible existence of cycles in the NPLTS, a scheduler that visit multiple times a certain state s must be able to choose differently how to resolve the non determinism every time it visits s. This leads to the notion of *correspondence function*.

Definition 4 ([16])**.** A *randomized resolution* for a NPLTS $\mathsf{N} = (\mathcal{S}, \mathcal{A}, \mathsf{trans})$ is a pair (corr, R) where

- $R = (\mathcal{S}', \mathcal{A}, \mathsf{trans}')$ is a RPLTS, and
- $\mathsf{corr} : \mathcal{S}' \to \mathcal{S}$ is the *correspondence function* such that for all states $s' \in \mathcal{S}'$, $\mathsf{trans}'(s')(a) = D$ implies $\mathsf{corr}(D) \in \mathrm{conv}(\mathsf{trans}(\mathsf{corr}(s'))(a))$.

Given a NPLTS N we denote by $\mathcal{R}_{\mathsf{r}}(\mathsf{N})$ the set of randomized resolutions. Additionally, we denote $\mathcal{R}_{\mathsf{r}}^{o} = \mathcal{R}_{\mathsf{r}}(\mathsf{N}^{o})$. Figure 4 shows an example of a resolution from $\mathcal{R}_{\mathsf{r}}^{o}$ for the process P from Example 2.

3.2. Computing the probability to reach a barb

The notion of *barb* is a classical way of expressing observables. Intuitively a state of N^{o}, *i.e.*, a multiset of processes, exhibits a barb c whenever an output on channel c is possible.

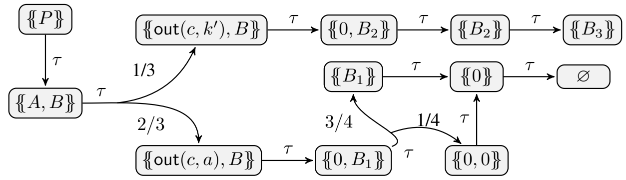


Figure 4: Example of a randomized resolution for process P from Example 2 where the correspondence function corr is the identity.

Definition 5. For $c \in \mathcal{N}_{pub}$ and $\mathcal{P} \in \mathcal{MP}$, we say that $\mathcal{P}$ *exhibits barb* c when there exists a process $\mathsf{out}(u, t).Q$ in $\mathcal{P}$ where $c \doteq u$ and $Msg(t)$. We denote by $\downarrow c$ the set of all multisets of processes that exhibit the barb c.

We next define the probability of reaching a state in a set of target states, in a *fully probabilistic* transition system, *i.e.*, in a transition system where all non-determinism–internal or external–has already been resolved. We first define the probability of reaching such a state in at most n steps, and then we take the probability of reaching them eventually as the limit of the n-step reaching probabilities.

Definition 6. Let $R = (\mathcal{S}, \mathcal{A}, \mathsf{trans})$ be a RPLTS, $\mathcal{T} \subseteq \mathcal{S}$ a set of states, and s an initial state. For every $n \in \mathbb{N}$ we define the *probability of reaching* $\mathcal{T}$ *from* s *in at most* n *steps* as:

$$\mathrm{RProb}_R^{\leq 0}(s, \mathcal{T}) = \begin{cases} 1 & \text{if } s \in \mathcal{T} \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathrm{RProb}_R^{\leq n+1}(s, \mathcal{T}) = \begin{cases} 1 & \text{if } s \in \mathcal{T} \\ 0 & \text{if } s \notin \mathcal{T} \wedge s \in \mathcal{S}_{ext}(R) \\ \sum_{u \in \mathsf{supp}(D)} D(u) \cdot \mathrm{RProb}_R^{\leq n}(u, \mathcal{T}) & \\ & \text{if } s \notin \mathcal{T} \wedge \mathsf{trans}(s)(\tau) = D \end{cases}$$

We define the *probability of reaching* $\mathcal{T}$ *from* s as:

$$\mathrm{RProb}_R(s, \mathcal{T}) = \lim_{n \to +\infty} \mathrm{RProb}_R^{\leq n}(s, \mathcal{T}).$$

Note that, as $\mathrm{RProb}_R^{\leq n}(s, \mathcal{T})$ is an increasing function in n we can replace the limit by the supremum on $n \in \mathbb{N}$.

Given $\mathsf{N} = (\mathcal{S}_{\mathsf{N}}, \mathcal{A}, \mathsf{trans}_{\mathsf{N}})$, we denote by $\mathrm{RProb}_{\mathcal{R}_{\mathsf{r}}(\mathsf{N})}(s, \mathcal{T})$ the probability:

$$\sup \left\{ \mathrm{RProb}_R(s', \mathsf{corr}^{-1}(\mathcal{T})) \;\middle|\; \begin{array}{l} (\mathsf{corr}, R) \in \mathcal{R}_{\mathsf{r}}(\mathsf{N}), \\ \mathsf{corr}(s') = s \end{array} \right\}$$

3.3. Defining May Testing Equivalence

Intuitively, two processes are may-testing equivalent if they exhibit the same observations when executed in the presence of any attacker process. This models the inability of an arbitrary process to distinguish them. More formally, two multisets of processes $\mathcal{P}$ and $\mathcal{Q}$ are may testing equivalent when the attacker has the same probability over all schedulers to exhibit the barb c in both $\mathcal{P}$ and $\mathcal{Q}$.

Definition 7 (May testing equivalence)**.** Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}$. We say that $\mathcal{P} \leq_{may} \mathcal{Q}$ iff:

$$\forall Adv \in \mathcal{MP} \text{ s.t. } fn(Adv) \subseteq \mathcal{N}_{pub}.\ \forall c \in \mathcal{N}_{pub}.$$
$$\mathrm{RProb}_{\mathcal{R}_r^\circ}(\mathcal{P} \cup Adv, \downarrow c) \leq \mathrm{RProb}_{\mathcal{R}_r^\circ}(\mathcal{Q} \cup Adv, \downarrow c)$$

We say that $\mathcal{P}, \mathcal{Q}$ are may testing equivalent, denoted $\mathcal{P} \approx_{may} \mathcal{Q}$, when $\mathcal{P} \leq_{may} \mathcal{Q}$ and $\mathcal{Q} \leq_{may} \mathcal{P}$.

One could also consider a more fine-grained definition of may testing pre-order that guarantees the equality of probabilities between two schedulers rather than comparing the probabilities over all schedulers. Formally, this pre-order, denoted $\leq'_{may}$, requires that for all resolutions $(\mathsf{corr}, R) \in \mathcal{R}_r(\mathsf{N})$ and state s of R such that $\mathsf{corr}(s) = \mathcal{P} \cup Adv$, there exist a resolution (corr', R') and a state s' of R' such that $\mathsf{corr}'(s') = \mathcal{Q} \cup Adv$ and

$$\mathrm{RProb}_R(s, \mathsf{corr}^{-1}(\downarrow c)) = \mathrm{RProb}_{R'}(s', \mathsf{corr}'^{-1}(\downarrow c))$$

However, the resulting relation is counter-intuitive and distinguishes processes

$$\mathcal{P} := \{\!\{\mathsf{out}(a, 0)\}\!\} \text{ and}$$
$$\mathcal{Q} := \{\!\{\mathsf{out}(a, 0) +_{\frac{1}{2}} (\mathsf{out}(a, 0) +_{\frac{1}{2}} \mathsf{out}(a, 0))\}\!\}.$$

Indeed, we can show that $\mathcal{Q} \not\leq'_{may} \mathcal{P}$: for $Adv = \{\!\{0\}\!\}$, there exists a resolution (corr, R) such that

$$\mathsf{corr}(s) = \mathcal{P} \cup Adv \quad \text{ and } \quad \mathrm{RProb}_R(\mathcal{Q}, \downarrow a) = \frac{1}{2}$$

but for every resolution (corr', R') such that $\mathsf{corr}'(s') = \mathcal{Q} \cup Adv$,

$$\mathrm{RProb}_{R'}(\mathcal{P}, \downarrow a) = 1.$$

3.4. Defining Observational Equivalence

In this section, we define *observational preorders* and *equivalence* which are stronger than may testing. When studying cryptographic protocols we suppose that internal actions are not observable and therefore only study *weak* equivalences hiding whether such internal actions take place or not. To define the observational preorder we need to introduce a weak relation for internal actions. In a purely non-deterministic system this simply corresponds to the reflexive, transitive closure $\xrightarrow{\tau}^*$. However, in our setting we need to compute the corresponding distributions.

Definition 8. Let N be a NPLTS and $D, E \in \mathcal{D}^{\leq 1}(\mathcal{S}_\mathsf{N})$. We write $D \stackrel{\tau}{\Longrightarrow}_{\mathcal{R}_r(\mathsf{N})} E$ when there exists $(\mathsf{corr}, R) \in \mathcal{R}_r(\mathsf{N})$ and $D', E' \in \mathcal{D}^{\leq 1}(\mathcal{S}_R)$ such that

- $\mathsf{corr}(D') = D$, $\mathsf{corr}(E') = E$, $\mathrm{supp}(E') \subseteq \mathcal{S}_{ext}(R)$,
- $\forall u \in \mathcal{S}_{ext}(R).\ E'(u) = \sum_{s' \in \mathcal{S}_R} D'(s') \cdot \mathrm{RProb}_R(s', \{u\})$.

To define the observational preorder, we additionally need to lift relations defined on a given set to relations on sub-distributions over this set.

Definition 9 (Lifting of a relation)**.** Let R be a binary relation on a discrete set $\mathcal{S}$. We define the *lifting of R to sub-distributions* as the binary relation on $\mathcal{D}^{\leq 1}(\mathcal{S})$, denoted $\widehat{R}$, defined as:

$$D\ \widehat{R}\ E \text{ when } \forall \mathcal{S}' \subseteq \mathcal{S}, D(\mathcal{S}') \leq E(R(\mathcal{S}'))$$

where $R(\mathcal{S}') = \{s \in \mathcal{S} \mid s' \in \mathcal{S}' \wedge s'\ R\ s\}$.

Using these notions of weak transition and lifting of relations to sub-distributions we can define observational equivalence.

Definition 10. The *observational preorder* $\leq_{obs}$ is the largest relation R on $\mathcal{MP}$ such that $\mathcal{P}\ R\ \mathcal{Q}$ implies :

- $\forall c \in \mathcal{N}_{pub}.\ \mathrm{RProb}_{\mathcal{R}_r^\circ}(\mathcal{P}, \downarrow c) \leq \mathrm{RProb}_{\mathcal{R}_r^\circ}(\mathcal{Q}, \downarrow c)$;
- if $\mathcal{P} \stackrel{\tau}{\Longrightarrow}_{\mathcal{R}_r^\circ} D$ and $D \in \mathcal{D}(\mathcal{S}_{\mathsf{N}^\circ})$ then $\mathcal{Q} \stackrel{\tau}{\Longrightarrow}_{\mathcal{R}_r^\circ} E$, $E \in \mathcal{D}(\mathcal{S}_{\mathsf{N}^\circ})$ and $D\ \widehat{R}\ E$;
- $\forall$ closed $Adv \in \mathcal{MP}$ such that $fn(Adv) \subseteq \mathcal{N}_{pub}$. $\{Adv\} \cup \mathcal{P}\ R\ \{Adv\} \cup \mathcal{Q}$.

The *observational equivalence* $\approx_{obs}$ is defined by additionally requiring R to be symmetric.

Remark 3. Note that we slightly diverge from the original definition of observational equivalence [11] where an evaluation context $C[\_]$ is of the form

$$\mathsf{new}\ n_1; \ldots; \mathsf{new}\ n_k; (\_ \mid A)$$

In our definition we simply consider a parallel process, and no additional name restriction. However, we prove in the full version that these two definitions coincide. Intuitively, restricting names whose scope includes the adversarial process A corresponds to making previously public channels invisible to the attacker at later steps, hence it does not provide additional distinguishing power.

4. Labelled semantics and equivalences

As usual in π calculi, and in the applied π-calculus, we define a *labelled* semantics. The intent of the labels is to capture adversarial actions and avoid the universal quantification over processes in equivalence definitions.

4.1. Labelled semantics

A state in this labeled semantics is defined by associating a multiset of processes with a *frame*, modeling the adversary's *knowledge*. We consider a new set of variables $\mathcal{AX} = \{\mathsf{ax}_1, \mathsf{ax}_2, \ldots\}$ distinct from $\mathcal{X}$ that will act as pointers to messages that were previously output.

Definition 11. An extended process is a pair $(\mathcal{P}, \phi)$, where $\mathcal{P} \in \mathcal{MP}$ and ϕ is a ground substitution

$$\{\mathsf{ax}_1 \mapsto t_1; \ \ldots\ ; \mathsf{ax}_n \mapsto t_n\}$$

such that $\mathsf{ax}_i \in \mathcal{AX}$, $t_i \in \mathcal{T}(\mathcal{F}, \mathcal{N})$ and $Msg(t_i)$ for $1 \leq i \leq n$.

We denote by $\mathcal{SP}_\ell$ the set of all extended processes.

A *recipe* is a term from $\mathcal{T}(\mathcal{F}, \mathcal{N}_{pub} \cup \mathcal{AX}\}$ representing how an attacker can deduce a message.

Notation 4. If D is a distribution over $\mathcal{MP}$, and ϕ a frame, we write (D, ϕ) for the distribution over extended processes defined as $(D, \phi) = \sum_{\mathcal{P} \in \mathsf{supp}(D)} D(\mathcal{P}) \cdot \delta_{(\mathcal{P}, \phi)}$.

We now define the NPLTS N^ℓ for the labelled semantics. External actions model interactions with the attacker.

Definition 12. The *labelled semantics* is the NPLTS $\mathsf{N}^\ell = (\mathcal{SP}_\ell, \{\tau\} \cup \mathcal{A}^\ell_{ext}, \mathsf{trans}^\ell)$ where

- $\mathcal{A}^\ell_{ext}$ is the set of labels $in(\xi, \zeta)$, $out(\xi, \mathsf{ax})$, $(\xi \stackrel{?}{=} \zeta)$ and $(\xi \stackrel{?}{\neq} \zeta)$ with ξ, ζ recipes and $\mathsf{ax} \in \mathcal{AX}$;
- $\mathsf{trans}^\ell((\mathcal{P}, \phi))(a) = \{D \mid (\mathcal{P}, \phi) \to_a D\}$ where $\to_a$ is defined in Figure 5.

Note that when we lift $\to_\tau$ to extended processes we suppose that the freshness requirement of a new name a' in the (NEW) rule of Figure 2 also applies to the frame ϕ, *i.e.*, a' must not appear in ϕ.

Remark 4. It should be noted that we deal with static equivalence in a different way as done usually in the applied π-calculus, or implicitly in the probabilistic applied π-calculus [12]: we encode static equivalence into the NPLTS N^ℓ by a countable set of actions–all the tests $(\xi \stackrel{?}{=} \zeta)$ and their negations– instead of just one action testing static equivalence. The motivation behind this choice is to be able to represent every *action* from the NPLTS by an *elementary action* of the adversary. As shown later, this choice has no effect on the definition of the simulation pre-orders or on bisimulation, but it leads to a slightly different notion of trace equivalence, that is closer to may testing equivalence.

4.2. Defining Trace Equivalence

We first define the probability of executing a trace for a given resolution. As we are interested in *weak* trace preorder (where internal actions cannot be observed), traces are sequences of external actions only. Our definition uses the previously introduced notation $\mathrm{RProb}_R(s, \{t\})$: recall that this denotes the probability of reaching state t from state s using only internal actions for some resolution (corr, R).

Definition 13. Let $R = (\mathcal{S}, \{\tau\} \sqcup \mathcal{A}_{ext}, \mathsf{trans})$ be a RPLTS. Let $w \in \mathcal{A}_{ext}{}^*$ be a trace, *i.e.*, a finite word on the alphabet $\mathcal{A}_{ext}$. For all states $s \in \mathcal{S}$, we define the *probability of executing w starting from s in R* as:

- $\mathrm{Prob}_R(s, \epsilon) = 1$
- $$\mathrm{Prob}_R(s, a.w) = \sum_{\substack{t \in \mathcal{S} \\ \mathsf{trans}(t)(a) = D}} \mathrm{RProb}_R(s, \{t\}) \cdot \sum_{s' \in \mathsf{supp}(D)} D(s') \cdot \mathrm{Prob}_R(s', w)$$

Given a NPLTS $\mathsf{N} = (\mathcal{S}, \mathcal{A}, \mathsf{trans})$, we denote by $\mathrm{Prob}_{\mathcal{R}_\mathsf{r}(\mathsf{N})}(s, w)$ the probability:

$$\sup\{\mathrm{Prob}_R(s', w) \mid (\mathsf{corr}, R) \in \mathcal{R}_\mathsf{r}(\mathsf{N}), \mathsf{corr}(s') = s\}$$

This allows us to define trace equivalence of $(\mathcal{P}, \phi)$ and $(\mathcal{P}', \phi')$: intuitively any trace that can be executed in $(\mathcal{P}, \phi)$ can be executed with at least the same probability in $(\mathcal{P}', \phi')$ and vice-versa.

Definition 14 (trace equivalence)**.** Let $(\mathcal{P}, \phi)$, $(\mathcal{P}', \phi') \in \mathcal{SP}_\ell$. We say that $(\mathcal{P}, \phi) \leq_{tr} (\mathcal{P}', \phi')$ iff for all $w \in \mathcal{A}^\ell_{ext}{}^*$,

$$\mathrm{Prob}_{\mathcal{R}_\mathsf{r}(\mathsf{N}^\ell)}((\mathcal{P}, \phi), w) \leq \mathrm{Prob}_{\mathcal{R}_\mathsf{r}(\mathsf{N}^\ell)}((\mathcal{P}', \phi'), w)$$

$(\mathcal{P}, \phi)$ and $(\mathcal{P}', \phi')$ are trace equivalent, denoted $(\mathcal{P}, \phi) \approx_{tr} (\mathcal{P}', \phi')$, when

$$(\mathcal{P}, \phi) \leq_{tr} (\mathcal{P}', \phi') \quad \text{and} \quad (\mathcal{P}', \phi') \leq_{tr} (\mathcal{P}, \phi).$$

Observe that in the NPLTS N^o, even though $s \xrightarrow{\tau} D$ implies that D has finite support, it is possible to have $\delta_s \stackrel{\tau}{\Longrightarrow}_\mathcal{R} E$, where the sub-distribution E has infinite support.

Unlike, in the purely possibilistic case, in our probabilistic setting trace preorder is strictly weaker than the may testing preorder.

Theorem 1. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}$ be two processes.

$$\mathcal{P} \leq_{may} \mathcal{Q} \quad \Rightarrow \quad (\mathcal{P}, \varnothing) \leq_{tr} (\mathcal{Q}, \varnothing)$$

Moreover, processes $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}$ defined in Figure 6 are such that $(\mathcal{P}, \varnothing) \leq_{tr} (\mathcal{Q}, \varnothing)$ and $\mathcal{P} \not\leq_{may} \mathcal{Q}$.

Figure 6 witnesses that the implication is strict. $\mathcal{P}$ and $\mathcal{Q}$ output each 3 encrypted bits (in a non-deterministic order). $\mathcal{P}$ outputs twice the encryption of 0; $\mathcal{Q}$ twice the encryption of 1. The (randomized) encryption ensures that these three values are indeed indistinguishable. We give the adversary a *single* access to a decryption oracle P_{dec}. Intuitively, trace equivalence holds, as the scheduler can ensure that matching encryptions are sent to P_{dec}. However, may-testing does not hold: the attacker chooses uniformly at random one of the three encryptions to submit. The probability to hit 0 will be $\frac{2}{3}$ in $\mathcal{P}$ and only $\frac{1}{3}$ in $\mathcal{Q}$.

Observe that Theorem 1 holds for any processes and does not require them to be image-finite, contrary to usual results in the literature, *e.g.*, [23]. This discrepancy comes from our choice of labelled actions for static equivalence (see Remark 4): a trace cannot test *directly* static equivalence, but can only do a *finite* numbers of recipe tests. We believe this variant definition of trace equivalence to be of independent interest as it provides an exact characterization of may testing in the purely non-deterministic case.

4.3. Defining Bisimulation

In this section, we define *simulations* on probabilistic processes and corresponding equivalences. Our definition of simulation preorder is similar to the definition of *randomized weak simulation preorder* introduced by Segala and Lynch for probabilistic automata [13]. We reuse the lifting of a relation and the weak relation for internal actions defined in Section 3.4 but applied to the NPLTS N^ℓ. In particular, given an action $a \in \mathcal{A}^\ell_{ext}$ and two distributions $D, D' \in$

$(\mathcal{P}, \phi) \rightarrow_\tau (D, \phi)$ if $\mathcal{P} \rightarrow_\tau D$

$(\{\!\{\mathsf{in}(u,x); P\}\!\} \cup \mathcal{P}, \phi) \rightarrow_{in(\xi,\zeta)} \delta_{(\{\!\{P\{^{\zeta\phi}/_x\}\}\!\} \cup \mathcal{P}, \phi)}$ if $u \doteq \xi\phi, Msg(\zeta\phi)$ and $vars(\xi,\zeta) \subseteq dom(\phi)$

$(\{\!\{\mathsf{out}(u,t); P\}\!\} \cup \mathcal{P}, \phi) \rightarrow_{out(\xi, \mathsf{ax}_{n+1})} \delta_{(\{\!\{P\}\!\} \cup \mathcal{P}, \phi\{ax_{n+1} \mapsto t\})}$ if $u \doteq \xi\phi, Msg(t), vars(\xi) \subseteq dom(\phi)$ and $|\phi| = n$

$(\mathcal{P}, \phi) \rightarrow_{(\xi \sim \zeta)} \delta_{(\mathcal{P}, \phi)}$ if $vars(\xi,\zeta) \subseteq dom(\phi)$ and $\xi\phi \sim \zeta\phi$ where $\sim \in \{\doteq, \not\doteq\}$

Figure 5: Labelled semantics: definition of $\rightarrow_a$

$$\begin{array}{rcl}\mathcal{P} & = & \{\!\{\mathsf{new}\ k;\ (P(0) \mid P(0) \mid P(1) \mid P_{dec})\}\!\} \\ \mathcal{Q} & = & \{\!\{\mathsf{new}\ k;\ (P(0) \mid P(1) \mid P(1) \mid P_{dec})\}\!\} \end{array} \quad \text{and}$$
$$\begin{array}{rcl} P(x) & = & \mathsf{new}\ r; \mathsf{out}(c, enc(x,r,k)) \\ P_{dec} & = & \mathsf{in}(d,y); \mathsf{out}(d, dec(y,k)) \end{array}$$

Figure 6: $\mathcal{P}, \mathcal{Q}$ such that $(\mathcal{P}, \varnothing) \leq_{tr} (\mathcal{Q}, \varnothing)$ and $\mathcal{P} \not\leq_{may} \mathcal{Q}$

$\mathcal{D}(\mathcal{S})$, we write $D \overset{a}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}} D'$ when $D \overset{\tau}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}} E_1$, $E_1 \xrightarrow{a} E_2$ and $E_2 \overset{\tau}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}} D'$ for some E_1, E_2 and where $\mathcal{R}^\ell_\mathsf{r}$ denotes $\mathcal{R}_\mathsf{r}(\mathsf{N}^\ell)$. Here $E_1 \xrightarrow{a} E_2$ is the natural lifting of the transition function of N^ℓ, *i.e.*, $E_2 = \sum_s E_1(s) \cdot D$ with $s \xrightarrow{a} D$.

Definition 15. A relation $R \subseteq (\mathcal{S}_{\mathsf{N}^\ell} \times \mathcal{S}_{\mathsf{N}^\ell})$ is

- a *simulation* if $s_1\ R\ s_2$ implies that for all $a \in \mathcal{A}^\ell_{ext} \cup \{\tau\}$, $D_1 \in \mathcal{D}(\mathcal{S}_{\mathsf{N}^\ell})$

$$\text{if} \quad s_1 \xrightarrow{a} D_1 \quad \text{then} \quad \begin{array}{l} s_2 \overset{a}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}} D_2,\ D_2 \in \mathcal{D}(\mathcal{S}_{\mathsf{N}^\ell}) \\ \text{and } D_1\ \widehat{R}\ D_2 \end{array}$$

- a *bisimulation* if R is a simulation and R is symmetric.

The *simulation preorder*, denoted $\leq_{sim}$, is the largest simulation and *bisimilarity*, denoted $\approx_{bi}$, is the largest bisimulation. We define *similarity*, denoted $\approx_{sim}$, as $\leq_{sim} \cap \leq_{sim}^{-1}$.

As usual in the field of (bi)simulation, it can be shown that $\leq_{sim}$, respectively $\approx_{bi}$, exists [19] and that it is a preorder, *i.e.*, a reflexive and transitive relation, respectively an equivalence, *i.e.*, a reflexive, symmetric and transitive relation [18].

The following proposition from [18] states that, as usual in the non-probabilistic case, the weak arrow $\overset{a}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}}$ can replace the single arrow $\xrightarrow{a}$ in the definition of simulation.

Proposition 1. Let R be the largest binary relation on $\mathcal{S}_{\mathsf{N}^\ell}$ such that $s_1\ R\ s_2$ implies that for every $a \in \mathcal{A}_{ext} \cup \{\tau\}$,

$$\text{if} \quad s_1 \overset{a}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}} D_1 \quad \text{then} \quad \begin{array}{l} s_2 \overset{a}{\Longrightarrow}_{\mathcal{R}^\ell_\mathsf{r}} D_2,\ D_2 \in \mathcal{D}(\mathcal{S}_{\mathsf{N}^\ell}) \\ \text{and } D_1\ \widehat{R}\ D_2 \end{array}$$

We have $R = \leq_{sim}$.

We now show that observational preorder and equivalence are exactly characterized by the simulation preorder and bisimilarity.

Theorem 2. Let $\mathcal{P}, \mathcal{Q}$ two processes in $\mathcal{MP}$.

$$\begin{array}{lll} \mathcal{P} \leq_{obs} \mathcal{Q} & \text{iff} & (\mathcal{P}, \varnothing) \leq_{sim} (\mathcal{Q}, \varnothing) \text{ and} \\ \mathcal{P} \approx_{obs} \mathcal{Q} & \text{iff} & (\mathcal{P}, \varnothing) \approx_{bi} (\mathcal{Q}, \varnothing) \end{array}$$

Proof sketch. We here provide the main intuitions of the proof that $\mathcal{P} \leq_{obs} \mathcal{Q}$ iff $(\mathcal{P}, \varnothing) \leq_{sim} (\mathcal{Q}, \varnothing)$.

($\Rightarrow$). To show that observational preorder implies simulation, we need to represent the frame of an extended process $(\mathcal{P}, \phi)$ as a process: we output in parallel the terms $\mathsf{ax}_i\phi$, with $\mathsf{ax}_i \in dom(\phi)$, on a public channel c_i, distinct for each i and not occurring anywhere in $(\mathcal{P}, \phi)$. Thus, we build the relation $\mathcal{R}$ such that $(\mathcal{P}, \phi)\,\mathcal{R}(\mathcal{Q}, \phi')$ if

$$\mathcal{P} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi); 0\}\!\}_{i=1}^n \leq_{obs} \mathcal{Q} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi').0\}\!\}_{i=1}^n$$

with $|dom(\phi)| = |dom(\phi')| = n$ and $c_1, \ldots, c_n \in \mathcal{N}_{pub}$ pairwise distinct and not occurring in $\mathcal{P}, \mathcal{Q}, \phi, \phi'$.

As the public channels c_i do not occur anywhere else, any internal transition on

$$\mathcal{P}_1 = \mathcal{P} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi); 0\}\!\}_{i=1}^n$$

must correspond to an internal transition on $\mathcal{P}$; and similarly for $\mathcal{Q}$.

For all visible actions, we rely on $\leq_{obs}$ being closed by composition with an adversarial process. For example, when the action is the test $\xi \overset{?}{=} \zeta$, we compose with the adversarial process that *(i)* reads the frame, *(ii)* applies the test, and *(iii)* outputs on a fresh public channel ok if the test succeeds:

$$\begin{array}{l} Adv = \mathsf{in}(c_1, x_i); \ldots; \mathsf{in}(c_n, x_n); \\ \qquad \mathsf{if}\ \xi\rho = \zeta\rho\ \mathsf{then}\ \mathsf{out}(ok, ok); 0\ \mathsf{else}\ 0 \end{array}$$

where $x_1, \ldots, x_n$ are fresh variables and $\rho = \{^{x_i}/_{\mathsf{ax}_i}\}_{i=1}^n$. We then consider the transition

$$\{\!\{Adv\}\!\} \cup \mathcal{P}_1 \overset{\tau}{\Longrightarrow}_{\mathcal{R}^o_\mathsf{r}} \delta_{\mathcal{P} \cup \mathsf{out}(ok,ok)}$$

and the fact that

$$\mathrm{RProb}_{\mathcal{R}_\mathsf{r}}(\mathcal{P} \cup \mathsf{out}(ok,ok), \downarrow ok) = 1$$

to conclude. Indeed, for

$$\{\!\{Adv\}\!\} \cup \mathcal{Q} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi').0\}\!\}_{i=1}^n \overset{\tau}{\Longrightarrow}_\mathsf{r} E$$

to exist with $\delta_{\mathcal{P} \cup \mathsf{out}(ok,ok)}\ \widehat{R}\ E$, $\{\!\{Adv\}\!\} \cup \mathcal{Q} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi'); 0\}\!\}_{i=1}^n$ must also have passed the test $\xi\rho = \zeta\rho$ in the conditional branching. Hence $\xi\phi' = \zeta\phi'$ and so $(\mathcal{Q}, \phi') \xrightarrow{\xi \overset{?}{=} \zeta} \delta_{(\mathcal{Q}, \phi')}$.

When the visible action is an output or an input, the process is more complicated. The adversarial process starts by reading the frame as before and executing the action. The last part of the adversarial process consists in outputting again the frame so that we re-enter the relation $\mathcal{R}$. Assume for instance the action $in(\xi, \zeta)$. By definition, $\mathcal{P} = \mathcal{P}_1 \cup \{\!\{\mathsf{in}(c, x); P\}\!\}$ with $\xi\phi \doteq c$, $(\mathcal{P}, \phi) \xrightarrow{in(\xi,\zeta)} \delta_{\mathcal{P}_1 \cup \{\!\{P\sigma\}\!\}}$ and $\sigma = \{^{\zeta\phi}/_x\}$.

We consider the following adversarial process Adv:

$$\begin{aligned} Adv = {} & \mathsf{in}(c_1, x_i); \dots; \mathsf{in}(c_n, x_n);\\ & \mathsf{out}(\xi\rho, \zeta\rho);\\ & (\mathsf{out}(ok, ok) +_{0.5} (\mathsf{out}(c'_1, x_1); 0 \mid \mathsf{out}(c'_n, x_n); 0)) \end{aligned}$$

where $c'_1, \dots, c'_n$ are fresh public names pairwise distinct not occurring anywhere else. We will conclude by considering the transition

$$\{\!\{Adv\}\!\} \cup \mathcal{P} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi); 0\}\!\}_{i=1}^n \xRightarrow{\tau}_{\mathsf{r}} D$$

where

$$\begin{aligned} D = {} & 0.5 \cdot \delta_{\mathcal{P}_1 \cup \{\!\{P\{^{\zeta\phi}/_x\}, \mathsf{out}(ok,ok).0\}\!\}}\\ & + 0.5 \cdot \delta_{\mathcal{P}_1 \cup \{\!\{P\{\zeta\phi/_x\}\}\!\} \cup \{\!\{\mathsf{out}(c'_i, \mathsf{ax}_i\phi); 0\}\!\}_{i=1}^n} \end{aligned}$$

Indeed, for

$$\{\!\{Adv\}\!\} \cup \mathcal{Q} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi'); 0\}\!\}_{i=1}^n \xRightarrow{\tau}_{\mathsf{r}} E$$

to exist with $D \,\widehat{R}\, E$, $\{\!\{Adv\}\!\} \cup \mathcal{Q} \cup \{\!\{\mathsf{out}(c_i, \mathsf{ax}_i\phi'); 0\}\!\}_{i=1}^n$ must also have applied an internal transition executing the construct $\mathsf{out}(\xi\rho, \zeta\rho)$ which allows for the labeled action $in(\xi, \zeta)$ to be executed on $(\mathcal{Q}, \phi')$.

($\Leftarrow$). Showing that simulation implies observational preorder is more straightforward. We build a relation $\mathcal{R}$ such that $\mathcal{P} \,\mathcal{R}\, \mathcal{Q}$ when there exist two extended processes $(\mathcal{P}_1, \phi), (\mathcal{Q}_1, \phi')$ with compatible frames (i.e. $dom(\phi) = dom(\phi')$), a renaming ρ from $\mathcal{N}_{pub}$ to $\mathcal{N}_{priv}$, and a multiset of adversarial processes $\mathcal{P}_{Att}$ such that:

- names in $img(\rho)$ do not occur in $\mathcal{P}_1, \phi, \mathcal{Q}_1$ and ϕ';
- $\mathcal{P} = \mathcal{P}_1\rho \cup \mathcal{P}_{Att}\{^{\mathsf{ax}_i\phi}/_{x_i}\}_{i=1}^n\rho$;
- $\mathcal{Q} = \mathcal{Q}_1\rho \cup \mathcal{P}_{Att}\{^{\mathsf{ax}_i\phi'}/_{x_i}\}_{i=1}^n\rho$;
- $(\mathcal{P}_1, \phi) \leq_{sim} (\mathcal{Q}_1, \phi')$.

The renaming ρ replaces the private names that are generated by the processes in $\mathcal{P}_{Att}$ (through the construct $\mathsf{new}\ a; P$) with public names that are chosen *fresh* (i.e. not in $\mathcal{P}_1, \phi, \mathcal{Q}_1$ and ϕ').

□

4.4. Randomized vs non-randomized schedulers

All our equivalence notions are based on randomized schedulers where the non-determinism is solved by picking a distribution from the convex hull of the available distributions. In the literature, more restrictive *non-randomized schedulers* have also been considered when defining observational equivalence and bisimilarity [12]. A non-randomized scheduler solves the non-determinism by choosing directly one of the available distributions. Formally, in Definition 4, instead of requiring that $\mathsf{trans}'(s')(a) = D$ implies $\mathsf{corr}(D) \in \mathsf{conv}(\mathsf{trans}(\mathsf{corr}(s'))(a))$, a non-randomized resolution requires that $\mathsf{trans}'(s')(a) = D$ implies $\mathsf{corr}(D) \in \mathsf{trans}(\mathsf{corr}(s'))(a)$ and corr is injective on the support of D.

Denoting by $\mathcal{R}_{\mathsf{nr}}(\mathsf{N})$ the set of all non-randomized schedulers of N, we can naturally update the notions used to define behavioural equivalences to non-randomized schedulers. For instance, we denote by $\mathrm{RProb}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N})}(s, \mathcal{T})$ the probability of reaching $\mathcal{T}$ from s over all non randomized schedulers $\mathcal{R}_{\mathsf{nr}}(\mathsf{N})$. Similarly, $\mathrm{Prob}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N})}(s, w)$ denotes the probability of executing the trace w from s over all schedulers from $\mathcal{R}_{\mathsf{nr}}(\mathsf{N})$. Updating the definitions results into may-testing and trace preorder for non-randomized schedulers, denoted $\leq_{may}^{\mathsf{nr}}$ and $\leq_{tr}^{\mathsf{nr}}$ respectively. We now show that $\leq_{may}$ and $\leq_{tr}$ do not depend on the whether schedulers are randomized or not (unlike simulation based notions as we will see below).

Lemma 1. May testing and trace preorders with randomized and non-randomized resolutions coincide:

$$\leq_{may} = \leq_{may}^{\mathsf{nr}} \quad \text{and} \quad \leq_{tr} = \leq_{tr}^{\mathsf{nr}}$$

Proof sketch. The core of the proof is the following fact: when we fix a randomized resolution R, an initial state s, and $n \in \mathbb{N}$, it is possible to *decompose* the behaviour of R from state s and during the n first execution steps into a weighted family of non-randomized resolution $(\alpha_i, R_i)_{i \in I}$ (where the weight α_i is a coefficient in $[0, 1]$, in the sense that for every set of processes $\mathcal{P}$, $\mathrm{RProb}_R^{\leq n}(s, \mathcal{P}) = \sum_{i \in I} \alpha_i \mathrm{RProb}_{R_i}^{\leq n}(s, \mathcal{P})$. The construction of this decomposition is defined inductively on n. □

This result is of interest as it is often easier to manipulate non-randomized schedulers, and we expect automated verification to be more convenient as well.

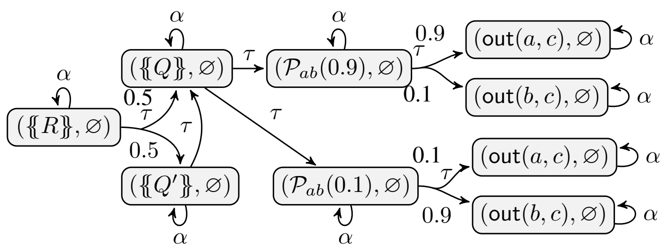
When considering observational equivalence, simulation and bisimulation, non-randomized schedulers raise a number of issues. First, as highlighted for instance in [19], [30], when considering bisimulation or simulation on general NPLTSs, non-randomized resolutions result into relations that are *not transitive*. We show that even on the specific NPLTS N^ℓ, simulation is not transitive.

Simulation for non-randomized scheduler is naturally defined by extending the notation $D \xRightarrow{\tau}_{\mathcal{R}_{\mathsf{r}}(\mathsf{N})} E$ to non-randomized scheduler, denoted $D \xRightarrow{\tau}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N})} E$: from Definition 8, we require (corr, R) to be in $\mathcal{R}_{\mathsf{nr}}(\mathsf{N})$ and additionally require an injectivity property on the correspondence function, *i.e.*, corr is injective on the support of D'. We denote the resulting simulation with non-randomized schedulers by $\leq_{sim}^{\mathsf{nr}}$ (and similarly for $\leq_{obs}^{\mathsf{nr}}$, $\approx_{obs}^{\mathsf{nr}}$ and $\approx_{bi}^{\mathsf{nr}}$).

Lemma 2. $\leq_{sim}^{\mathsf{nr}}$ is not transitive.

Proof sketch. Consider processes

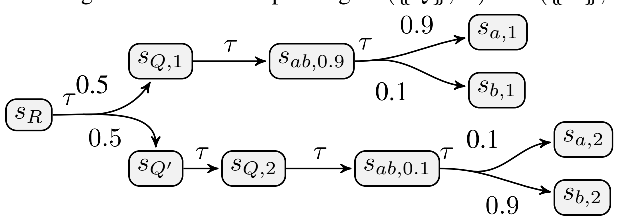
$$\begin{aligned} & P = \mathsf{out}(a, c) +_{0.5} \mathsf{out}(b, c)\\ & Q = (\mathsf{out}(a, c) +_{0.9} \mathsf{out}(b, c)) + (\mathsf{out}(a, c) +_{0.1} \mathsf{out}(b, c))\\ & Q' = \mathsf{if}\ c = c\ \mathsf{then}\ Q\ \mathsf{else}\ 0\\ & R = Q +_{0.5} Q' \end{aligned}$$



For readability, α stands for all labels $(\xi \sim \xi')$, with closed recipes ξ, ξ' such that $\xi \sim \xi'$ ($\sim \in \{\doteq, \not\doteq\}$) and

$$\mathcal{P}_{ab}(p) = \{\!\!\{\mathsf{out}(a,c) +_p \mathsf{out}(b,c)\}\!\!\}$$

(a) The fragment of N^ℓ corresponding to $(\{\!\!\{Q\}\!\!\}, \varnothing)$ and $(\{\!\!\{R\}\!\!\}, \varnothing)$.



(b) The resolution for $(\{\!\!\{R\}\!\!\}, \varnothing) \overset{\tau}{\Longrightarrow}_{\mathcal{R}_{\mathsf{nr}}} 0.5 \cdot \delta_{\mathsf{out}(a,c)} + 0.5 \cdot \delta_{\mathsf{out}(b,c)}$.

Figure 7: Fragments of N^ℓ showing $(\{\!\!\{P\}\!\!\}, \varnothing) \leq^{\mathsf{nr}}_{sim} (\{\!\!\{R\}\!\!\}, \varnothing) \leq^{\mathsf{nr}}_{sim} (\{\!\!\{Q\}\!\!\}, \varnothing)$ but $(\{\!\!\{P\}\!\!\}, \varnothing) \not\leq^{\mathsf{nr}}_{sim} (\{\!\!\{Q\}\!\!\}, \varnothing)$

The corresponding fragment of N^ℓ is displayed in Figure 7a. We will show that

$$(\{\!\!\{P\}\!\!\}, \varnothing) \leq^{\mathsf{nr}}_{sim} (\{\!\!\{R\}\!\!\}, \varnothing) \text{ and } (\{\!\!\{R\}\!\!\}, \varnothing) \leq^{\mathsf{nr}}_{sim} (\{\!\!\{Q\}\!\!\}, \varnothing)$$
$$\text{but } (\{\!\!\{P\}\!\!\}, \varnothing) \not\leq^{\mathsf{nr}}_{sim} (\{\!\!\{Q\}\!\!\}, \varnothing)$$

It is easy to see that $(\{\!\!\{Q\}\!\!\}, \varnothing) \approx^{\mathsf{nr}}_{bi} (\{\!\!\{Q'\}\!\!\}, \varnothing)$ and so $(\{\!\!\{R\}\!\!\}, \varnothing) \leq^{\mathsf{nr}}_{sim} (\{\!\!\{Q\}\!\!\}, \varnothing)$. The difficult part of the proof of $(\{\!\!\{P\}\!\!\}, \varnothing) \leq^{\mathsf{nr}}_{sim} (\{\!\!\{R\}\!\!\}, \varnothing)$ is to match

$$(\{\!\!\{P\}\!\!\}, \varnothing) \xrightarrow{\tau} 0.5 \cdot \delta_{(\{\!\!\{\mathsf{out}(a,c)\}\!\!\}, \varnothing)} + 0.5 \cdot \delta_{(\{\!\!\{\mathsf{out}(b,c)\}\!\!\}, \varnothing)}$$

This is achieved by the scheduler displayed in Figure 7b.

Finally, we prove $(\{\!\!\{P\}\!\!\}, \varnothing) \not\leq^{\mathsf{nr}}_{sim} (\{\!\!\{R\}\!\!\}, \varnothing)$ by showing that the transition $(\{\!\!\{P\}\!\!\}, \varnothing) \xrightarrow{\tau} 0.5 \cdot \delta_{(\{\!\!\{\mathsf{out}(a,c)\}\!\!\}, \varnothing)} + 0.5 \cdot \delta_{(\{\!\!\{\mathsf{out}(b,c)\}\!\!\}, \varnothing)}$ cannot be simulated in $(\{\!\!\{Q\}\!\!\}, \varnothing)$. $\square$

Note that the definitions of bisimilarity in [12] rely on non-randomized schedulers. Even though this does not necessarily imply that their relation is not transitive (as they focus directly on the semantics of processes) we show in the next lemma that $\approx^{\mathsf{nr}}_{bi}$ and $\approx^{\mathsf{nr}}_{obs}$ do *not* coincide, hence disproving [12, Theorem 2]. This reenforces our belief that it is preferable to use randomized schedulers in our definition.

Lemma 3. There exist processes $P, Q \in \mathcal{SP}$ such that

- $(\{\!\!\{Q\}\!\!\}, \varnothing) \approx_{bi} (\{\!\!\{P\}\!\!\}, \varnothing)$,
- $(\{\!\!\{Q\}\!\!\}, \varnothing) \approx^{\mathsf{nr}}_{bi} (\{\!\!\{P\}\!\!\}, \varnothing)$, and
- $\{\!\!\{Q\}\!\!\} \not\leq^{\mathsf{nr}}_{obs} \{\!\!\{P\}\!\!\}$.

Proof sketch. We consider the following processes:

$$\begin{aligned} P &= \mathsf{out}(d,c); (\mathsf{out}(a,c) +_{0.9} 0) + \\ &\quad\ \mathsf{out}(d,c); (\mathsf{out}(b,c) +_{0.9} 0) \\ P' &= \mathsf{if}\ c = c\ \mathsf{then}\ P\ \mathsf{else}\ 0 \\ Q &= P +_{\frac{1}{2}} P' \end{aligned}$$

Both $(\{\!\!\{Q\}\!\!\}, \varnothing) \approx^{\mathsf{nr}}_{bi} (\{\!\!\{P\}\!\!\}, \varnothing)$ and $(\{\!\!\{Q\}\!\!\}, \varnothing) \approx^{\mathsf{nr}}_{bi} (\{\!\!\{P\}\!\!\}, \varnothing)$ are proved by showing that the binary relation R, defined as the reflexive, symmetric and transitive closure of $\{((\{\!\!\{Q\}\!\!\}, \varnothing), (\{\!\!\{P\}\!\!\}, \varnothing)), ((\{\!\!\{P\}\!\!\}, \varnothing), (\{\!\!\{P'\}\!\!\}, \varnothing))\}$, is a bisimulation (see Figure 8a).

To prove that $\{\!\!\{Q\}\!\!\} \not\leq^{\mathsf{nr}}_{obs} \{\!\!\{P\}\!\!\}$, we show that $\{\!\!\{Q;\ \mathsf{in}(d,x).0\}\!\!\} \not\leq^{\mathsf{nr}}_{obs} \{\!\!\{P;\ \mathsf{in}(d,x).0\}\!\!\}$. In particular (see Figure 8b), we build a non-randomized scheduler such that $\{\!\!\{Q;\ \mathsf{in}(d,x).0\}\!\!\} \overset{\tau}{\Longrightarrow}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N}^o)} D$ where $D = 0.45 \cdot \delta_{\mathcal{P}_a} + 0.45 \cdot \delta_{\mathcal{P}_b} + 0.1 \cdot \delta_{\varnothing}$. However, there is no distribution E such that $\{\!\!\{P;\ \mathsf{in}(d,x).0\}\!\!\} \overset{\tau}{\Longrightarrow}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N}^o)} E$, and $D \mathrel{\widehat{\leq^{\mathsf{nr}}_{obs}}} E$. $\square$

Remark that we have cast the definitions of [12] in our own framework. In the full version [27] we show that processes P, Q in Lemma 3 can be adapted to obtain the counterpart of Lemma 3 in the exact framework of [12]. The failure of the proof of [12, Theorem 2] can be traced back to the auxilliary lemma [12, Lemma 3] that states that bisimilarity is closed under application of closing evaluation contexts. No proof of this lemma is however provided, and it is actually false: as shown in the proof of (our) Lemma 3, the extended processes $(\{\!\!\{Q\}\!\!\}, \varnothing)$ and $(\{\!\!\{P\}\!\!\}, \varnothing)$ defined there are bisimilar (with respect to non-randomized schedulers), but it is not the case of the extended processes $(\{\!\!\{Q \mid \mathsf{in}(d,x).0\}\!\!\}, \varnothing)$ and $(\{\!\!\{P \mid \mathsf{in}(d,x).0\}\!\!\}, \varnothing)$.

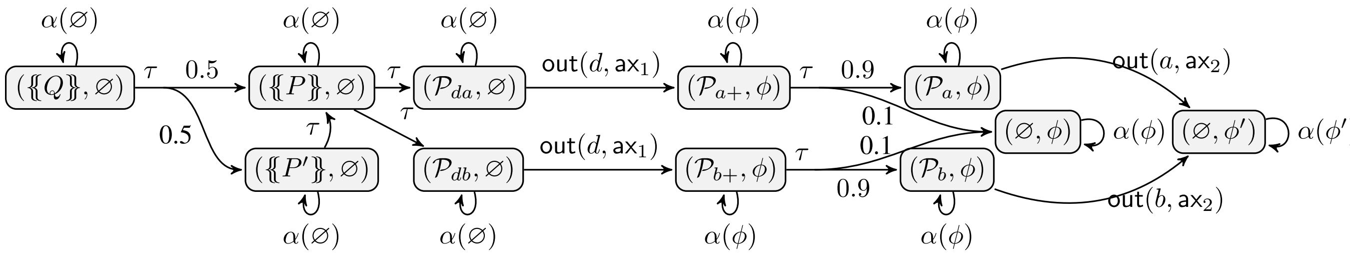
5. Well behaved subclasses of protocols

It is a well-known phenomenon that non-determinism and probabilistic choices do not interact well: a particular scheduler may for instance leak a secret probabilistic choice. Such schedulers are generally deemed unrealistic, and several papers aim at restricting schedulers, *e.g.*, [20], [21]. We illustrate this phenomenon on the following example.

Example 3. Consider the processes

$$\begin{aligned} P &:= (\mathsf{in}(c,x).\ \mathsf{if}\ x = 0\ \mathsf{then}\ \mathsf{out}(ok,1)\ \mathsf{else}\ \mathsf{out}(bad,1)) +_{\frac{1}{2}} \\ &\quad\ (\mathsf{in}(c,x).\ \mathsf{if}\ x = 0\ \mathsf{then}\ \mathsf{out}(bad,1)\ \mathsf{else}\ \mathsf{out}(ok,1)) \\ Q &:= \mathsf{in}(c,x).(\mathsf{out}(ok,1) +_{\frac{1}{2}} \mathsf{out}(bad,1)) \end{aligned}$$

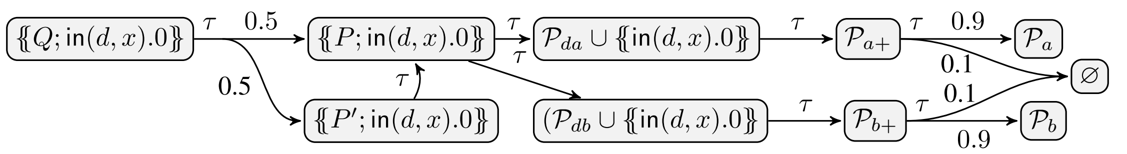
One may, intuitively, consider that these two processes exhibit the same behaviour. Q takes an input and then with probability $\frac{1}{2}$ decides to either output on ok or on bad. P on the other hand first choses a branch with probability $\frac{1}{2}$. Each branch performs an input and, depending on the input value, outputs either on ok or on bad. As the two branches make opposite choices on the output according to the input value, one might expect the probability to output on ok to be $\frac{1}{2}$.



For readability, $\alpha(\phi)$ stands for all labels $\xi \sim \xi'$ with ξ, ξ' closed recipes such that $\xi\phi \sim \xi'\phi$ and $\sim \in \{\doteq, \neq\}$.

(a) The fragment of N^ℓ corresponding to $(\{\!\{Q\}\!\}, \varnothing)$ and $(\{\!\{P\}\!\}, \varnothing)$

$$\begin{array}{llll}
\mathcal{P}_{da} = \mathsf{out}(d,c); (\mathsf{out}(a,c) +_{0.9} 0) & \mathcal{P}_{a+} = \{\!\{\mathsf{out}(a,c) +_{0.9} 0\}\!\} & \mathcal{P}_a = \{\!\{\mathsf{out}(a,c)\}\!\} & \phi = \{\mathsf{ax}_1 \to c\} \\
\mathcal{P}_{db} = \mathsf{out}(d,c); (\mathsf{out}(b,c) +_{0.9} 0) & \mathcal{P}_{b+} = \{\!\{\mathsf{out}(b,c) +_{0.9} 0\}\!\} & \mathcal{P}_b = \{\!\{\mathsf{out}(b,c)\}\!\} & \phi' = \{\mathsf{ax}_1 \to c, \mathsf{ax}_2 \to c\}
\end{array}$$



(b) The fragment of N^o corresponding to $\{\!\{P; \mathsf{in}(d,x).0\}\!\}$ and $\{\!\{Q; \mathsf{in}(d,x).0\}\!\}$

Figure 8: Fragments of NPLTS showing $(\{\!\{Q\}\!\}, \varnothing) \approx^{\mathsf{nr}}_{bi} (\{\!\{P\}\!\}, \varnothing)$ and $\{\!\{Q\}\!\} \not\lesssim^{\mathsf{nr}}_{obs} \{\!\{P\}\!\}$.

However, P and Q are not may testing equivalent and can be distinguished by the adversary $Adv = \{\!\{\mathsf{out}(c,0) \mid \mathsf{out}(c,1)\}\!\}$. Indeed, we can show that:

$$\mathrm{RProb}_{\mathcal{R}^o_\mathsf{r}}(\mathcal{P} \cup Adv, \downarrow ok) = 1 \quad \mathrm{RProb}_{\mathcal{R}^o_\mathsf{r}}(\mathcal{Q} \cup Adv, \downarrow ok) = \frac{1}{2}$$

Intuitively, this results from the fact that the resolution may *leak* the probabilistic choice through the non-deterministic choice of the attacker to output 0 or 1. The resolution chooses the attacker to output 0 in the first probabilistic branch of P and 1 in the second.

In this section we identify two subclasses of processes that avoid this problem. The first such subclass is that of non-probabilistic processes, *i.e.*, without the $+_p$ operator (we denote by $\mathcal{MP}^{\mathsf{np}}$ all the multisets of such processes). This is the class of the original applied π-calculus which also enjoys good tool support. Figure 6 already illustrated that even on non-probabilistic processes, *probabilistic* adversaries have a stronger distinguishing power for the may testing equivalence. We formally characterize this distinguishing power when restricting protocols to a bounded number of sessions (denoted $\mathcal{MP}^{<\infty,\mathsf{np}}$), *i.e.*, considering processes without replication: for this subclass, may-testing coincides with similarity. We therefore inherit from [22] the fact that deciding may-testing is coNEXP complete for a large class of cryptographic primitives.

The second subclass considers purely probabilistic processes with (nearly) no non-determinism. We show that trace equivalence in this class (as considered for instance in [25]) corresponds to may-testing with a restricted, *determinate* adversary process. We also sketch how the algorithms of the DeepSec prover [22] could be adapted to check trace equivalence in this probabilistic setting.

5.1. Non-probabilistic processes

May-testing with non-probabilistic adversaries and trace equivalence coincide. Our definitions of may testing and trace equivalence coincide with the classical definitions of the original, purely non-deterministic applied π-calculus when all processes are non probabilistic. As a first step, we observe that the weak operational semantics we defined in Section 4.3 is a conservative extension of the weak (non-probabilistic) operational semantics: indeed, when considering non-probabilistic processes, all distributions in the (labeled) operational semantics are Dirac distributions.

Notation 5. We write $\mathcal{SP}^{\mathsf{np}}_\ell$ for the set of all non-probabibilistic extended processes We write $\to_{\mathsf{np}}$, respectively $\xrightarrow{a}_{\mathsf{np}}$, for the one-step reduction relation we obtain when we restrict the NPLTS N^o to $\mathcal{MP}^{\mathsf{np}}$, respectively N^ℓ to $\mathcal{SP}^{\mathsf{np}}_\ell$. For $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{\mathsf{np}}$, we write $\mathcal{P} \Rightarrow_{\mathsf{np}} \mathcal{Q}$ when there exists a sequence $\mathcal{P} = \mathcal{P}_0 \to_{\mathsf{np}} \ldots \to_{\mathsf{np}} \mathcal{P}_n = \mathcal{Q}$.

Lemma 4. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{\mathsf{np}}$.

$$\mathcal{P} \Rightarrow_{\mathsf{np}} \mathcal{Q} \quad \text{iff} \quad \mathrm{RProb}_{\mathcal{R}_{\mathsf{nr}(\mathsf{N}^o)}}(\mathcal{P}, \{\mathcal{Q}\}) = 1$$

We look now at *preorder relations* between non-probabilistic processes. We first recall formally how may testing, trace equivalence and bisimulation are defined for non-probabilistic processes (Definition 16 below). Those definitions are coherent with those from the literature, e.g [23], (up to the difference on static equivalence, discussed in Remark 4).

Notation 6. If b is a barb, we write $\mathcal{P} \Downarrow_b$ when there exists $\mathcal{Q}$ such that $\mathcal{P} \Rightarrow_{\mathsf{np}} \mathcal{Q}$, and $\mathcal{Q} \downarrow_b$. For $a \in \mathcal{A}^{\ell}_{ext}$, we write $(\mathcal{P}, \phi) \xRightarrow{a}_{\mathsf{np}} (\mathcal{Q}, \psi)$ when $(\mathcal{P}, \phi) \xrightarrow{\tau}_{\mathsf{np}} \dots \xrightarrow{a}_{\mathsf{np}} \dots \xrightarrow{\tau}_{\mathsf{np}} (\mathcal{Q}, \psi)$. If $\alpha = a_1, \dots, a_n$ is a trace, we write $(\mathcal{P}, \phi) \xRightarrow{\alpha}_{\mathsf{np}} (\mathcal{Q}, \psi)$ when there exists a sequence $(\mathcal{P}, \phi) \xRightarrow{a_1}_{\mathsf{np}} \dots \xRightarrow{a_n}_{\mathsf{np}} (\mathcal{Q}, \psi)$.

Definition 16. We define the binary relations $\leq^{\mathsf{np}}_{may}$, $\leq^{\mathsf{np}}_{tr}$, $\leq^{\mathsf{np}}_{sim}$ on $\mathcal{MP}^{\mathsf{np}}$ as follows:

- $\mathcal{P} \leq^{\mathsf{np}}_{may} \mathcal{Q}$ when $\forall Adv \in \mathcal{MP}^{\mathsf{np}}$ s.t. $fn(Adv) \subseteq \mathcal{N}_{pub}$. $\forall c \in \mathcal{N}_{pub}$. $Adv \cup \mathcal{P} \Downarrow_c$ implies $Adv \cup \mathcal{Q} \Downarrow_c$;
- $(\mathcal{P}, \phi) \leq^{\mathsf{np}}_{tr} (\mathcal{Q}, \psi)$ when for every trace α, $(\mathcal{P}, \phi) \xRightarrow{\alpha}_{\mathsf{np}} (\mathcal{P}', \phi')$ implies $(\mathcal{Q}, \psi) \xRightarrow{\alpha}_{\mathsf{np}} (\mathcal{Q}', \psi')$;
- $\leq^{\mathsf{np}}_{sim}$ is the largest reflexive and transitive relation R such that $(\mathcal{P}, \phi)\ R\ (\mathcal{Q}, \psi)$ implies that for every $a \in \mathcal{A}^{\ell}_{ext} \cup \{\tau\}$, and $(\mathcal{P}, \phi) \xrightarrow{a}_{\mathsf{np}} (\mathcal{P}', \phi')$, there exists $(\mathcal{Q}', \psi')$ such that $(\mathcal{Q}, \psi) \xRightarrow{a}_{\mathsf{np}} (\mathcal{Q}', \psi')$ and $(\mathcal{P}', \phi')\ R\ (\mathcal{Q}', \psi')$.

The preorders $\leq_{sim}$ and $\leq_{tr}$–and the corresponding equivalence relations–are conservative extensions of $\leq^{\mathsf{np}}_{sim}$ and $\leq^{\mathsf{np}}_{tr}$. As expected, the preorder $\leq_{may}$ *is not* a conservative extension of $\leq^{\mathsf{np}}_{may}$, because of the additional expressive power of probabilistic adversaries. Nonetheless, we can recover $\leq^{\mathsf{np}}_{may}$ when we restrict the adversaries in the definition of $\leq_{may}$ to non-probabilistic adversaries.

Proposition 2. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{\mathsf{np}}$.

- $(\mathcal{P}, \varnothing) \leq^{\mathsf{np}}_{sim} (\mathcal{Q}, \varnothing)$ iff $(\mathcal{P}, \varnothing) \leq_{sim} (\mathcal{Q}, \varnothing)$;
- $(\mathcal{P}, \varnothing) \leq^{\mathsf{np}}_{tr} (\mathcal{Q}, \varnothing)$ iff $(\mathcal{P}, \varnothing) \leq_{tr} (\mathcal{Q}, \varnothing)$;
- $(\mathcal{P}, \varnothing) \leq^{\mathsf{np}}_{may} (\mathcal{Q}, \varnothing)$ iff

$$\forall Adv \in \mathcal{MP}^{\mathsf{np}} \text{ s.t. } fn(Adv) \subseteq \mathcal{N}_{pub}.\ \forall c \in \mathcal{N}_{pub}.$$
$$\mathrm{RProb}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N}^o)}(\mathcal{P} \cup Adv, \downarrow c) \leq \mathrm{RProb}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N}^o)}(\mathcal{Q} \cup Adv, \downarrow c)$$

Proof sketch. For may-testing and trace preorder, the proof uses crucially the fact that it is enough to consider non-randomized distributions (thanks to Lemma 1), and from there, we conclude using Lemma 4. Since it is not possible to consider only non-randomized schedulers in the definition of simulation, the proof of the first point in Proposition 2 is more subtle, and uses well-structured properties of the lifting of a relation from Definition 9. We need to show (1) that the (non-probabilistic) simulation $\leq^{\mathsf{np}}_{sim}$ is a probabilistic simulation in the sense of Definition 15, and (2) that $\leq_{sim}$ is also a non-probabilistic simulation in the sense of Definition 16.

- Suppose that $\mathcal{P} \leq^{\mathsf{np}}_{sim} \mathcal{Q}$ for $\mathcal{P}, \mathcal{Q} \in \mathcal{SP}^{\mathsf{np}}_{\ell}$. Let $a \in \mathcal{A}^{\ell}_{ext} \cup \{\tau\}$, $D \in \mathcal{D}(\mathcal{S}_{\mathsf{N}^{\ell}})$ such that $\mathcal{P} \xrightarrow{a}_{\mathsf{r}} D$. Looking at the way we defined $\xrightarrow{a}_{\mathsf{r}}$ (and since $\mathcal{P}$ is non-probabilistic), we also see that $\mathcal{P} \xrightarrow{a}_{\mathsf{np}} \mathcal{P}'_i$ for every $\mathcal{P}'_i$ is the support of D. From there, we obtain that for each i, there exists $\mathcal{Q}'_i$ such that $\mathcal{Q} \xRightarrow{a}_{\mathsf{np}} \mathcal{Q}'_i$, and $\mathcal{P}'_i \leq^{\mathsf{np}}_{sim} \mathcal{Q}'_i$. At that point, we build $E = \sum_i D(\mathcal{P}'_i) \cdot \delta_{\mathcal{Q}'_i}$, and we can see that $\mathcal{Q} \xRightarrow{a}_{\mathsf{r}} E$. Moreover, the structural properties of the lifting allows us to go from $(\forall i,\ \mathcal{P}'_i \leq^{\mathsf{np}}_{sim} \mathcal{Q}'_i)$ to $D\ \widehat{\leq^{\mathsf{np}}_{sim}}\ E$. Hence, we have shown that $\leq^{\mathsf{np}}_{sim}$ is indeed a probabilistic simulation in the sense of Definition 15.
- Suppose that $\mathcal{P} \leq_{sim} \mathcal{Q}$ for $\mathcal{P}, \mathcal{Q} \in \mathcal{SP}^{\mathsf{np}}_{\ell}$. Let $a \in \mathcal{A}^{\ell}_{ext} \cup \{\tau\}$, and $\mathcal{P}' \in \mathcal{SP}^{\mathsf{np}}_{\ell}$ such that $\mathcal{P} \xrightarrow{a}_{\mathsf{np}} \mathcal{P}'$. This transition carries over to N^{ℓ}, i.e., $\mathcal{P} \xrightarrow{a}_{\mathsf{r}} \delta_{\mathcal{P}'}$. We obtain that there exists a distribution E such that $\mathcal{Q} \xRightarrow{a}_{\mathsf{r}} E$, and $\widehat{\delta_{\mathcal{P}'} \leq_{sim}} E$. But by structural property of the lifting, we have that $\mathcal{P}' \leq_{sim} \mathcal{Q}'$ for every element $\mathcal{Q}'$ in the support of E. Moreover, since $\mathcal{Q}$ is non-probabilistic, it holds that $\mathcal{Q} \xRightarrow{a}_{\mathsf{np}} \mathcal{Q}'$ for every element $\mathcal{Q}'$ in the support of E. Since E is a distribution, there exists at least one such element $\mathcal{Q}'$, thus we can conclude. $\square$

The following result indicates that may testing and trace equivalence coincide in non-probabilistic settings. In particular, we recover the fact that for the classical definitions in non-probabilistic settings, trace equivalence implies may-testing, as shown in [23].

Proposition 3. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{\mathsf{np}}$. $(\mathcal{P}, \varnothing) \leq_{tr} (\mathcal{Q}, \varnothing)$ iff

$$\forall Adv \in \mathcal{MP}^{\mathsf{np}} \text{ s.t. } fn(Adv) \subseteq \mathcal{N}_{pub}.\ \forall c \in \mathcal{N}_{pub}.$$
$$\mathrm{RProb}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N}^o)}(\mathcal{P} \cup Adv, \downarrow c) \leq \mathrm{RProb}_{\mathcal{R}_{\mathsf{nr}}(\mathsf{N}^o)}(\mathcal{Q} \cup Adv, \downarrow c)$$

May-testing and simulation coincide for bounded processes. We rely on a modal characterization of strong simulation on *image finite labeled transition systems* (LTS) by a Hennessy-Milner logic [31] (HML).

We can rely on *strong* simulation as it is a well-known fact ([32]) that simulation for a LTS can be expressed as strong simulation on the corresponding *weak LTS*, that is, in our case all transitions $(\mathcal{P}, \phi) \xrightarrow{\tau}{}^{*} \xrightarrow{a} \xrightarrow{\tau}{}^{*} (\mathcal{Q}, \psi)$ are merged into a single transition.

A LTS is image finite when the LTS cannot infinitely branch from a state and a label. Therefore, as we consider only bounded processes and by denoting $\leq^{\mathsf{L}}_{ssim}$ the strong simulation relation on a LTS L, we can build an image finite LTS L^{ℓ} such that for all $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{\mathsf{np}}$:

$$(\mathcal{P}, \varnothing) \leq_{sim} (\mathcal{Q}, \varnothing) \quad \text{iff} \quad (\mathcal{P}, \varnothing) \leq^{\mathsf{L}^{\ell}}_{ssim} (\mathcal{Q}, \varnothing)$$

Our HML characterization consists in expressing strong simulation preorder by the means of satisfaction of *logical formulas* by the LTS.

Definition 17. Let $\mathcal{A}$ be a countable set of actions. We define the set of *logical formulas* as:

$$F \in \mathcal{F} := \top \mid a.F \mid F_1 \wedge F_2, \qquad \text{where } a \in \mathcal{A}$$

In our case, the set of actions corresponds to $\mathcal{A}^{\ell}_{ext}$ that is indeed countable. The satisfaction of such formulas by a LTS is defined as follows.

Definition 18. Let $\mathsf{L} = (\mathcal{S}, \mathcal{A}, \rightarrow)$ be a LTS. We say that L satisfies a formula F, written $s \models F$, if for all $s \in \mathcal{S}$,

- $s \models \top$;
- $s \models a.F$ when $s \xrightarrow{a} t$ and $t \models F$;

- $s \models F_1 \wedge F_2$ when $s \models F_1$ and $s \models F_2$.

The following proposition shows how to relate strong simulation with satisfiability of logical formulas.

Proposition 4 (HML caracterisation of simulation)**.** For an image finite LTS L,

$$s \leq^{\mathsf{L}}_{ssim} t \quad \text{iff} \quad \forall F \in \mathcal{F}.\ s \models F \text{ implies } t \models F$$

In order to prove that simulation coincides with may-testing for bounded non-probabilistic processes, we show that we can *emulate* any logical formula by a probabilistic adversary: for all formulas F, we build a probabilistic adversary Adv^c_F such that for all bounded non-probabilistic extended processes $(\mathcal{P}, \phi)$,

$$(\mathcal{P}, \varnothing) \models F \quad \text{iff} \quad \mathrm{RProb}_{\mathcal{R}_r}(\mathcal{P} \cup \{\!\{ Adv^c_F \}\!\}, \downarrow c) = 1$$

We illustrate the construction of Adv^c_F on a few selected formulas:

$$\begin{array}{rcl} Adv^c_{in(\xi,\zeta).F'} & = & \mathsf{out}(\xi,\zeta).Adv^c_{F'} \\ Adv^c_{(\xi \overset{?}{=} \zeta).F'} & = & \mathsf{if}\ \xi = \zeta\ \mathsf{then}\ Adv^c_{F'} \\ Adv^c_{F_1 \wedge F_2} & = & Adv^c_{F_1} +_{\frac{1}{2}} Adv^c_{F_2} \\ Adv^c_{\top} & = & \mathsf{out}(c,c) \end{array}$$

In particular, conjunction is encoded by probabilistic choice and on formula $\top$, the adversary process exhibits the barb c. The main result of this section follows almost directly.

Proposition 5. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{<\infty,\mathsf{np}}$.

$$(\mathcal{P}, \varnothing) \leq_{sim} (\mathcal{Q}, \varnothing) \quad \text{iff} \quad \mathcal{P} \leq_{may} \mathcal{Q}$$

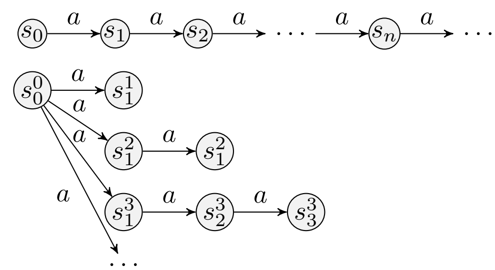
Cheval *et al.* have shown [22] that both deciding trace equivalence and bisimilarity is coNEXPTIME complete when cryptographic primitives are modelled by a subterm convergent destructor rewrite system and the number of sessions is bounded. (We refer the reader to [22] for a precise definition of this class of rewrite systems.) The hardness proof reduces SUCCINT 3SAT to both trace equivalence and bisimilarity using a same encoding which also proves hardness of similarity. The coNEXPTIME decision procedure for bisimilarity can be directly adapted to the case of similarity, hence, we have the following result.

Corollary 1. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{<\infty,\mathsf{np}}$. Deciding $\mathcal{P} \approx_{may} \mathcal{Q}$ is coNEXPTIME complete when $\doteq$ is defined by a subterm convergent destructor rewrite system.

May-testing and simulation do not coincide for unbounded processes. We consider the following two simple LTS L_1 and L_2, that are an asymetric variant of the LTSs used in [33] to show that the finitary HML does not characterize bisimulation.

L_1:

L_2:



Though both L_1 and L_2 may produce an unbounded number of a transitions, the initial transition in L_2 decides on an arbitrary, but fixed number of a transitions in the rest of the execution. This in particular shows that L_2 does not simulate L_1, *i.e.*, $\mathsf{L}_1 \not\leq_{sim} \mathsf{L}_2$.

In the applied π-calculus, L_1 can be represented by the process $!\mathsf{out}(c,a)$. Modeling L_2 is more complex: we non-deterministically choose an integer $n \geq 0$, and output $n+1$ times a. The non deterministic choice of n is realized through the following process that outputs the unary encoding $h^n(b)$ of n:

$$\begin{array}{rl} Q_{count}(e) = & \mathsf{new}\ d.(\mathsf{out}(d,b) \mid \\ & \quad !\mathsf{in}(d,x).(\mathsf{out}(e,x) + \mathsf{out}(d,h(x)))) \end{array}$$

Channel d represents a memory cell initiated with a public name b (encoding 0). When reading on d the current value x, the process non deterministically chooses to increment x (updating the cell with $h(x)$) or to *select* the value x by outputting it on channel e.

It remains to model the process that given an integer x, produces $x+1$ outputs of a:

$$\begin{array}{rl} Q_{out}(e) = & \mathsf{new}\ d'.\mathsf{in}(e,x).(\mathsf{out}(d',b) \mid \\ & \quad !\mathsf{in}(d',y).\mathsf{out}(c,a).\mathsf{if}\ x = y\ \mathsf{then}\ 0 \\ & \quad\ \mathsf{else}\ \mathsf{out}(d',h(y))) \end{array}$$

Similarly to $Q_{count}(e)$, $Q_{out}(e)$ relies on a private d' to increment b until reaching the value x read on e. It is easy to see that the process outputs $x+1$ times a.

Lemma 5. Let $Q = \mathsf{new}\ e.(Q_{count}(e) \mid Q_{out}(e))$. We have:

$$(!\mathsf{out}(c,a), \varnothing) \not\leq^{\mathsf{N}^\ell}_{sim} (Q, \varnothing) \quad \text{but} \quad !\mathsf{out}(c,a) \leq_{may} Q$$

5.2. Purely probabilistic processes

At the opposite of the spectrum of purely non-non-deterministic processes, we study *purely probabilistic* processes with (nearly) no non-determinism. A similar class of processes has been considered in [25], [26] to model various protocols relying on randomization (*e.g.*, Crowds [10], mix-net [9], electronic voting [34]). They consider systems that are built as the parallel composition of independant agents, called *roles*, and where all communications are mediated by the adversary. Moreover, the internal behavior of each role is deterministic; the only non-determinism is controlled by the adversary–thus external–and consists in the adversary's choice for scheduling the communications. We model purely probabilistic processes as follows.

Definition 19. A process is *fully determinate* if it does not contain the operators $+$, $|$, nor $!$.

$\mathcal{P} = \{P_1, \ldots, P_n\} \in \mathcal{MP}$ is *purely probabilistic* when:

- each P_i is fully determinate;
- there exist distinct public channels $c_1, \ldots, c_n \in \mathcal{N}_{pub}$ such that for all $i \in \{1, \ldots, n\}$, all input and output actions in P_i are on c_i.

The class of purely probabilistic multisets of processes is denoted by $\mathcal{MP}^{\mathsf{pp}}$.

$\mathcal{MP}^{\mathsf{pp}}$ can also be seen as a probabilistic extension of the class of *simple processes* introduced in [35] to show that trace equivalence coincides with observational equivalence for such processes.

Removing scheduling of τ-actions. In [25], [26], the authors consider trace equivalence for a slightly restricted fragment of purely probabilistic processes. More precisely, all processes have exactly the same control structure which removes the necessity of scheduling honest τ-actions and allows to directly consider *strong* trace equivalence, as exactly the same τ-actions occur. In this work, we lift this restriction on the shape of the processes and show instead that the non-determinism related to honest τ actions is inconsequential when deciding trace equivalence. Indeed, in a multiset of processes $\{P_1, \ldots, P_n\}$, a τ action may be available simultaneously in multiple components. However, all such τ-action are in fact purely deterministic (*e.g.*, conditional branching, probabilistic choice). Moreover, as the P_is do not contain parallel composition and all input and output occur on distinct channels c_i, no internal communication between processes P_i and P_j is possible.

We show that to compute the probability of executing a trace w, we only need to consider a single *maximal* resolution on the NPLTS N^ℓ. Such a resolution always executes an action when at least one is available. Formally, a resolution (corr, R) with $R = (S, \mathcal{A}, \mathsf{trans})$ on purely probabilistic processes is maximal when for all $s \in S$, if there exists $\mathsf{corr}(s) \xrightarrow{a} D$ in N^ℓ for some a, D then $\mathsf{trans}(s)(a) = D'$ for some D'.

Proposition 6. Let $(\mathcal{P}, \phi)$ be an extended purely probabilistic process. For all maximal resolutions (corr, R) on N^ℓ, for all $s \in \mathcal{S}(R)$ with $\mathsf{corr}(s) = (\mathcal{P}, \phi)$,

$$\forall w \in \mathcal{A}^\star_{ext}.\ \mathrm{Prob}_R(s, w) = \mathrm{Prob}_{\mathcal{R}^\ell_\mathsf{r}}((\mathcal{P}, \phi), w).$$

May-testing and trace equivalence coincide for fully determinate adversaries. As illustrated in Example 3, may-testing is strictly stronger than trace equivalence even on purely probabilistic processes due to the non-determinism in the adversarial process. However, by restricting the adversarial process to be fully determinate, we can show that may-testing and trace equivalence coincide. We define the resulting *determinate may testing preorder*, denoted $\leq_{d\text{-}may}$, exactly as in Definition 7 but additionally restrict the adversary process Adv to be a singleton $\{\!\{A\}\!\}$ where A is fully determinate.

Theorem 3. Let $\mathcal{P}, \mathcal{Q} \in \mathcal{MP}^{\mathsf{pp}}$.

$$\mathcal{P} \leq_{d\text{-}may} \mathcal{Q} \quad \text{iff} \quad (\mathcal{P}, \varnothing) \leq_{tr} (\mathcal{Q}, \varnothing)$$

Proof sketch. To prove that $(\mathcal{P}, \varnothing) \leq_{d\text{-}may} (\mathcal{Q}, \varnothing)$ implies $\mathcal{P} \leq_{tr} \mathcal{Q}$ we encode any trace w into a determinate adversary Adv^c_w where c is fresh. Adv^c_w is defined in a similar way as Adv^c_F (Section 5.1), *e.g.*,

$$\begin{array}{rcl} Adv^c_{in(\xi,\zeta).w'} & = & out(\xi,\zeta); Adv^c_{w'} \text{ and} \\ Adv^c_{(\xi \stackrel{?}{=} \zeta).w'} & = & \mathsf{if}\ \xi = \zeta\ \mathsf{then}\ Adv^c_{w'} \end{array}$$

In particular, on the empty trace the adversary process exhibits the barb c: $Adv^c_\varepsilon = \mathsf{out}(c, c)$. We obtain that

$$\mathrm{Prob}_{\mathcal{R}_\mathsf{r}(\mathsf{N}^\ell)}((\mathcal{P}, \varnothing), w) = \mathrm{RProb}_{\mathcal{R}^o_\mathsf{r}}(\mathcal{P} \cup \mathcal{A}_{dv}, \downarrow c)$$

The other implication is more difficult as the adversarial process Adv is allowed to use probabilistic choices which cannot be directly encoded in a trace. Instead, we show that any adversarial process Adv aiming to exhibit a barb c corresponds to a multiset of weighted traces $\mathsf{Tr}(Adv)$, built inductively on Adv. For instance, when $Adv = Adv_1 +_p Adv_2$ and

$$\mathsf{Tr}(Adv_i) = \{\!\{(p^i_k, w^i_k)\}\!\}^{n_i}_{k=1} \quad \text{for } i = 1, 2$$

then

$$\mathsf{Tr}(Adv) = \{\!\{(p \cdot p^1_k, w^1_k)\}\!\}^{n_1}_{k=1} \cup \{\!\{((1-p) \cdot p^2_k, w^2_k)\}\!\}^{n_2}_{k=1}$$

For other constructs, the set of weighted traces is built as expected, *e.g.*, $\mathsf{Tr}(0) = \varnothing$ and $\mathsf{Tr}(\mathsf{new}\ a; Adv') = \mathsf{Tr}(Adv')$. For outputs or inputs, we additionally test if the channel corresponds to the barb c, *i.e.*, when $Adv = \mathsf{out}(\xi, \zeta); Adv'$ and $\mathsf{Tr}(Adv') = \{\!\{(p_k, w_k)\}\!\}^n_{k=1}$,

$$\mathsf{Tr}(Adv) = \{\!\{(p_k, (\xi \stackrel{?}{\neq} c).in(\xi, \zeta).w_k)\}\!\}^n_{k=1} \cup \{\!\{(1, \xi \stackrel{?}{=} c)\}\!\}$$

This construction yields the following property:

$$\begin{aligned} &\mathrm{RProb}_{\mathcal{R}_\mathsf{nr}(\mathsf{N}^o)}(P \cup Adv, \downarrow c) = \mathrm{RProb}_{\mathcal{R}_\mathsf{nr}(\mathsf{N}^o)}(P, \downarrow c) \\ &\quad + (1 - \mathrm{RProb}_{\mathcal{R}_\mathsf{nr}(\mathsf{N}^o)}(P, \downarrow c)) \cdot \\ &\quad \sum_{(\alpha, w) \in \mathsf{Tr}(Adv)} \alpha \cdot \mathrm{Prob}_{\mathcal{R}_\mathsf{nr}(\mathsf{N}^\ell)}((P, \varnothing), w) \end{aligned}$$

Intuitively, exhibiting the barb on c does not require any interaction with the adversary, or this interaction is correctly encoded in $\mathsf{Tr}(Adv)$. Using this property we easily conclude. $\square$

Towards deciding trace equivalence. As previous mentioned, Cheval *et al.* [22] designed a decision procedure for trace equivalence when cryptographic primitives are modelled by a subterm convergent destructor rewrite system and a bounded number of sessions. This procedure is based on constraint solving techniques that represent in a finite symbolic tree the infinite set of all possible concrete executions of the processes and an arbitrary attacker. Intuitively, each node of this symbolic tree represents the state of the two processes after executing a trace tr. Due to non-determinism, a node may contain several constraint systems corresponding to every possible interleaving allowing the execution of such trace tr. Deciding trace equivalence between processes A and B, in the original, non-probabilistic setting, requires to check that each node of the symbolic tree contains at leat one constraint system derived from process A and one from process B; or the node is empty.

For purely probabilistic processes, as there is no non-determinism, we can associate to each constraint system the probability of executing this trace, depending on the probabilistic choices taken during the execution of the processes. Thus, using the same constraint solving techniques, we can compute a symbolic tree representing the state of the processes after executing trace tr but where the constraint systems of a node correspond to the possible probabilistic choices. Deciding trace equivalence between processes A and B thus reduces to checking that in every node, the sum of the probabilities of the constraint systems derived from the process A is equal to the sum of the probabilities of the constraint systems derived from B.

6. Conclusion and future work

In this paper we introduced a framework to reason about indistinguishability properties, modelled as process equivalences, in symbolic models enhanced with probabilities. Defining such a framework turns out to rely on subtle technicalities such as the need for randomized schedulers, overseen in previous attempts. In addition to solving technical problems, we believe that randomized schedulers capture more faithfully the idea that one cannot predict how non-determinism is resolved. Randomized schedulers generalize the idea that one distribution is chosen non-deterministically by allowing an arbitrary combination (in the convex hull) of the available distributions.

We define different, classical behavioral and labelled equivalences and show their precise relations. As usual in models mixing non-determinism and probabilities, the resulting equivalences may be considered as too strong: indeed arbitrary schedulers may leak the (private) probabilistic choices of the processes and give the attacker an unrealistically strong distinguishing power. Defining more restricted schedulers that are only allowed partial knowledge of the current state, such as in [20], is orthogonal to our work. We however believe that our work provides a convenient framework for defining such more fine-grained notions of schedulers and consider this an interesting direction for future work.

We therefore study two classes of protocols where this problem is avoided. First, we study protocols that do not make probabilistic choices, but allow the adversary to do so. This class of non-probabilistic protocols corresponds to the classical setting and captures all major case studies performed in the context of symbolic models. Our results highlight that the classical notion of may-testing, considered rather intuitive as it models an arbitrary attacker running in parallel, does not take into account attackers that make probabilistic choices. Interestingly, when bounding the number of sessions, (non-probabilistic) similarity exactly captures such probabilistic attackers and offers an attractive target for automated analysis. Second, we study a class of fully probabilistic protocols, also considered in [25], and show that trace equivalence on such protocols coincides with may testing in the presence of a (syntactic) class of determinate attackers. One may indeed argue that determinacy removes artificial non-deterministic choices that the attacker could exploit and that correspond to unrealistic behaviors. When protocols can be expressed in the class of purely probabilistic processes, from a formal analysis point, it seems appealing to do so as it also simplifies the analysis.

Our work paves the road towards several future work, in addition to exploring restricted schedulers mentioned above. We sketched a decision procedure in the DeepSec tool for the class of fully probabilistic adversaries. A thorough treatment of this idea and an implementation are a natural next step. Also, the insight that (purely possibilistic) similarity takes into account probabilistic adversaries (as it coincides with may testing) when the number of sessions is bounded and protocols are non-deterministic motivates adding support for (bi)similarity in a tool such as DeepSec (which currently only verifies trace equivalence). A different direction going beyond the subclasses considered in this paper is to investigate restrictions of the scheduler (building, *e.g.*, on ideas from [20], [21]) in our framework to limit the adversary's power without restricting the class of protocols. Finally, a more prospective direction is the use of more quantitative equivalences, *i.e.*, distances between processes, that might be interesting to compare different protocols that try to achieve a same property.

Acknowledgments. We thank Alwen Tiu and the anonymous reviewers for their helpful comments and suggestions. This work has been partly supported by the ANR Research and teaching chair in AI ASAP and by ANR TECAP (decision number ANR-17-CE39-0004-03).

References

[1] D. Dolev and A. C. Yao, "On the security of public key protocols," *IEEE Trans. Inf. Theory*, vol. 29, no. 2, pp. 198–207, 1983.

[2] K. Bhargavan, B. Blanchet, and N. Kobeissi, "Verified models and reference implementations for the tls 1.3 standard candidate," in *IEEE Symposium on Security and Privacy (S&P'17)*. IEEE, 2017, pp. 483–502.

[3] C. Cremers, M. Horvat, J. Hoyland, S. Scott, and T. van der Merwe, "A comprehensive symbolic analysis of TLS 1.3," in *ACM SIGSAC Conference on Computer and Communications Security (CCS'17)*. ACM, 2017, pp. 1773–1788.

[4] K. Cohn-Gordon, C. Cremers, B. Dowling, L. Garratt, and D. Stebila, "A formal security analysis of the signal messaging protocol," *Journal of Cryptology*, vol. 33, no. 4, pp. 1914–1983, 2020.

[5] D. A. Basin, J. Dreier, L. Hirschi, S. Radomirovic, R. Sasse, and V. Stettler, "A formal analysis of 5g authentication," in *ACM SIGSAC Conference on Computer and Communications Security (CCS'18)*, D. Lie, M. Mannan, M. Backes, and X. Wang, Eds. ACM, 2018, pp. 1383–1396.

[6] D. A. Basin, R. Sasse, and J. Toro-Pozo, "The EMV standard: Break, fix, verify," in *42nd IEEE Symposium on Security and Privacy (S&P'21)*. IEEE, 2021, pp. 1766–1781.

[7] V. Cortier, S. Kremer, and B. Warinschi, “A survey of symbolic methods in computational analysis of cryptographic systems,” *Journal of Automated Reasoning*, vol. 46, no. 3-4, pp. 225–259, Apr. 2010.

[8] D. Chaum, “The dining cryptographers problem: Unconditional sender and recipient untraceability,” *J. Cryptol.*, vol. 1, no. 1, pp. 65–75, 1988.

[9] ——, “Untraceable electronic mail, return addresses, and digital pseudonyms,” *Commun. ACM*, vol. 24, no. 2, pp. 84–88, 1981.

[10] M. K. Reiter and A. D. Rubin, “Crowds: Anonymity for web transactions,” *ACM Trans. Inf. Syst. Secur.*, vol. 1, no. 1, pp. 66–92, 1998.

[11] M. Abadi, B. Blanchet, and C. Fournet, “The applied pi calculus: Mobile values, new names, and secure communication,” *Journal of the ACM (JACM)*, 2017.

[12] J. Goubault-Larrecq, C. Palamidessi, and A. Troina, “A probabilistic applied pi–calculus,” in *Asian Symposium on Programming Languages and Systems*. Springer, 2007, pp. 175–190.

[13] R. Segala and N. Lynch, “Probabilistic simulations for probabilistic processes,” *Nordic Journal of Computing*, vol. 2, no. 2, pp. 250–273, 1995.

[14] A. Parma and R. Segala, “Logical characterizations of bisimulations for discrete probabilistic systems,” in *International Conference on Foundations of Software Science and Computational Structures*. Springer, 2007, pp. 287–301.

[15] V. Castiglioni, “Trace and testing metrics on nondeterministic probabilistic processes,” in *Combined 25th International Workshop on Expressiveness in Concurrency and 15th Workshop on Structural Operational Semantics and 15th Workshop on Structural Operational Semantics,(EXPRESS/SOS) 2018*, vol. 276, 2018, pp. 19–36.

[16] F. Bonchi, A. Sokolova, and V. Vignudelli, “The theory of traces for systems with nondeterminism and probability,” in *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, 2019, pp. 1–14.

[17] Y. Deng, R. Van Glabbeek, M. Hennessy, and C. Morgan, “Testing finitary probabilistic processes,” in *International Conference on Concurrency Theory*. Springer, 2009, pp. 274–288.

[18] M. Stoelinga, “Alea jacta est: verification of probabilistic, real-time and parametric systems,” Ph.D. dissertation, University of Nijmegen, the Netherlands, Apr. 2002.

[19] C. G. Eisentraut, “Principles of markov automata,” Ph.D. dissertation, Saarländische Universitäts-und Landesbibliothek, 2017.

[20] K. Chatzikokolakis and C. Palamidessi, “Making random choices invisible to the scheduler,” *Inf. Comput.*, vol. 208, no. 6, pp. 694–715, 2010.

[21] M. S. Alvim, M. E. Andrés, C. Palamidessi, and P. van Rossum, “Safe equivalences for security properties,” in *Theoretical Computer Science - 6th IFIP TC 1/WG 2.2 International Conference, TCS 2010, Held as Part of WCC 2010*, ser. IFIP Advances in Information and Communication Technology, C. S. Calude and V. Sassone, Eds., vol. 323. Springer, 2010, pp. 55–70.

[22] V. Cheval, S. Kremer, and I. Rakotonirina, “Deepsec: Deciding equivalence properties in security protocols - theory and practice,” in *Proceedings of the 39th IEEE Symposium on Security and Privacy (S&P’18)*. San Francisco, CA, USA: IEEE Computer Society Press, May 2018, pp. 525–542.

[23] V. Cheval, V. Cortier, and S. Delaune, “Deciding equivalence-based properties using constraint solving,” *Theor. Comput. Sci.*, vol. 492, pp. 1–39, 2013.

[24] H. Comon-Lundh and V. Cortier, “Computational soundness of observational equivalence,” in *ACM Conference on Computer and Communications Security (CCS’08)*, P. Ning, P. F. Syverson, and S. Jha, Eds. ACM, 2008, pp. 109–118.

[25] R. Chadha, A. P. Sistla, and M. Viswanathan, “Verification of randomized security protocols,” in *32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS’17)*. IEEE, 2017, pp. 1–12.

[26] M. S. Bauer, R. Chadha, A. P. Sistla, and M. Viswanathan, “Model checking indistinguishability of randomized security protocols,” in *International Conference on Computer Aided Verification (CAV’18)*. Springer, 2018, pp. 117–135.

[27] V. Cheval, R. Crubillé, and S. Kremer, “Symbolic protocol verification with dice: process equivalences in the presence of probabilities (extended version),” Jun. 2022, working paper or preprint. [Online]. Available: https://hal.inria.fr/hal-03683907

[28] B. Blanchet, “Modeling and verifying security protocols with the applied pi calculus and proverif,” *Foundations and Trends in Privacy and Security*, 2016.

[29] R. Milner, J. Parrow, and D. Walker, “A calculus of mobile processes, I,” *Inf. Comput.*, vol. 100, no. 1, pp. 1–40, 1992.

[30] Y. Deng, “Axiomatisations and types for probabilistic and mobile processes,” Ph.D. dissertation, École des Mines de Paris, 2005.

[31] M. Hennessy and R. Milner, “On observing nondeterminism and concurrency,” in *Automata, Languages and Programming (ICALP’80)*, J. de Bakker and J. van Leeuwen, Eds. Springer, 1980, pp. 299–309.

[32] R. Amadio, “Operational methods in semantics,” 2016.

[33] R. J. van Glabbeek, “Bounded nondeterminism and the approximation induction principle in process algebra,” in *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, 1987, pp. 336–347.

[34] D. Chaum, P. Y. A. Ryan, and S. A. Schneider, “A practical voter-verifiable election scheme,” in *10th European Symposium on Research in Computer Security (ESORICS’05)*, ser. Lecture Notes in Computer Science, S. D. C. di Vimercati, P. F. Syverson, and D. Gollmann, Eds., vol. 3679. Springer, 2005, pp. 118–139.

[35] V. Cortier and S. Delaune, “A method for proving observational equivalence,” in *22nd IEEE Computer Security Foundations Symposium, (CSF’09)*. IEEE Computer Society, 2009, pp. 266–276.