Foundations and Trends® in Programming Languages

Probabilistic Trace and Testing Semantics: The Importance of Being Coherent

Suggested Citation: Marco Bernardo (2022), "Probabilistic Trace and Testing Semantics: The Importance of Being Coherent", Foundations and Trends® in Programming Languages: Vol. xx, No. xx, pp 1–81. DOI: 10.1561/XXXXXXXXX.

Marco Bernardo
Università di Urbino
marco.bernardo@uniurb.it

now
the essence of knowledge
Boston — Delft

Contents

Probabilistic Trace and Testing Semantics: The Importance of Being Coherent

Marco Bernardo[1]

[1] *Università di Urbino, Italy; marco.bernardo@uniurb.it*

ABSTRACT

It is well known that trace and testing semantics over nondeterministic and probabilistic processes are influenced by the class of schedulers used to resolve nondeterministic choices. In particular, it is the capability of suitably limiting the power of the considered schedulers that guarantees the validity of a number of desirable properties of those semantics. Among such properties we mention the fact of being coarser than bisimulation semantics, the fact of being a congruence with respect to typical process operators, and the fact of coinciding with the corresponding semantics when restricting to fully nondeterministic or fully probabilistic processes.

In this monograph, we recall various approaches against almighty schedulers appearing in the literature, we survey structure-preserving and structure-modifying resolutions of nondeterminism by providing a uniform definition for them, and we present an overview of behavioral equivalences for nondeterministic and probabilistic processes along with some anomalies affecting trace and testing semantics. We then introduce the notion of coherent resolution, which

Marco Bernardo (2022), "Probabilistic Trace and Testing Semantics: The Importance of Being Coherent", Foundations and Trends® in Programming Languages: Vol. xx, No. xx, pp 1–81. DOI: 10.1561/XXXXXXXXX.

prevents a scheduler from selecting different continuations in equivalent states of a process, so that the states to which they correspond in any resolution of the process have equivalent continuations too.

We show that coherency avoids anomalies related to the discriminating power, the compositionality, and the backward compatibility of probabilistic trace post-equivalence and pre-equivalence, which are variants of trace semantics. Moreover, we exhibit an alternative characterization of the former based on coherent trace distributions and an alternative characterization of the latter relying on coherent weighted trace sets. We finally extend the notion of coherent resolution by adding suitable transition decorations and prove that this ensures the insensitivity of probabilistic testing equivalence to the moment of occurrence of nondeterministic or probabilistic choices among identical actions, thus enhancing the backward compatibility of testing semantics.

1

Introduction

1.1 Probabilistic Behavioral Models and Relations

Quantitative models of computer, communication, and software systems combine, among others, functional and extra-functional aspects of system behavior. On the one hand, these models describe system operations and their execution order, possibly admitting nondeterminism in case of concurrency phenomena or to support implementation freedom. On the other hand, they include some information about the probabilities or the durations of activities and events in which the system is involved.

In the probabilistic setting, a particularly expressive model is given by probabilistic automata (Segala, 1995a), because they encompass as special cases fully nondeterministic models like labeled transition systems (Keller, 1976), fully probabilistic models like action-labeled variants of discrete-time Markov chains (Kemeny and Snell, 1960), and reactive probabilistic models like Markov decision processes (Derman, 1970). In a probabilistic automaton, which consists of states and transitions, the choice among the transitions departing from the current state is nondeterministic and can be influenced by the external environment, while the choice of the next state reached by the selected transition is probabilistic and is made internally by the process.

Behavioral relations (van Glabbeek, 2001; Jou and Smolka, 1990; Huynh and Tian, 1992; Baier *et al.*, 2005; Bernardo, 2007; Bernardo *et al.*, 2014b) play a fundamental role in the analysis of probabilistic models. They formalize observational mechanisms that permit relating models that, despite their different representations in the same mathematical domain, cannot be distinguished by external entities when abstracting from certain internal details. Moreover, they support system modeling and verification by providing a means to relate system descriptions expressed at different levels of abstraction, as well as to reduce the size of a system representation while preserving specific properties to be assessed later.

From the first comparative work (De Nicola, 1987) to the elaboration of the full spectrum (van Glabbeek, 2001), a number of equivalences have emerged over fully nondeterministic models, which range from the branching-time – i.e., (bi)simulation-based – endpoint (Park, 1981; Milner, 1989) to the linear-time – i.e., trace-based – endpoint (Brookes *et al.*, 1984) passing through testing relations (De Nicola and Hennessy, 1984). The spectrum becomes simpler when considering fully probabilistic models (Jou and Smolka, 1990; Huynh and Tian, 1992; Baier *et al.*, 2005; Bernardo, 2007), whereas as shown in (Bernardo *et al.*, 2014b) it is much more variegated in the case of models with nondeterminism and probabilities like probabilistic automata. The reason is that the probability of equivalence-specific events can be calculated only after removing nondeterminism. Examples of such events are the reachability via given actions of certain sets of equivalent states (bisimulation semantics), the execution of specific action sequences (trace semantics), and the passing of tests (testing semantics), with states/traces being possibly enriched with additional information.

In this monograph, we focus on trace and testing semantics for nondeterministic and probabilistic processes represented by simple probabilistic automata (Segala, 1995a).

A trace is a sequence of activities labeling a sequence of transitions performed by a process, thus abstracting from branching points in the process behavior. Several execution probabilities may be associated with the same trace, each corresponding to a different resolution of nondeterminism. Although the discriminating power of probabilistic trace

equivalences depends on how nondeterminism is resolved, in general this power turns out to be excessive, which hampers the achievement of a number of desirable properties.

A test is formalized as a nondeterministic and probabilistic process extended with success states or success actions, which is run in parallel with the process under test thus resulting in an interaction or testing system. The probability of reaching success is not unique, but depends on the specific resolution of nondeterminism considered within the interaction system. Also in the testing approach, the resulting probabilistic behavioral equivalences tend to be overdiscriminating.

1.2 Struggling Against Demonic Schedulers

Nondeterminism is resolved by resorting to *policies*, according to the terminology of (Bellman, 1957), or *schedulers*, according to the terminology of (Vardi, 1985). They establish which is the next transition or combination of transitions to be executed, possibly based on the sequence of states traversed so far.

The problem with almighty schedulers yielding a demonic view of nondeterminism is well known for both trace and testing semantics. In the case of a process given by the parallel composition of several subprocesses, or in a testing scenario where a process is composed in parallel with a test, schedulers come into play *after* the various components have been assembled together. As a consequence, schedulers can solve both choices local to the individual components and choices arising from their interleaving execution. In other words, this *centralized* approach enables any scheduler to make decisions in one component on the basis of those made in other components, especially in the case of history-dependent schedulers (Vardi, 1985).

To cope with the aforementioned information leakage, the idea of *distributed* scheduling was proposed in (de Alfaro *et al.*, 2001), which is akin to partial-information policies (de Alfaro, 1999). Given a number of modules, i.e., of variable-based versions of automata, that interact *synchronously* by updating all variables during every round, for each module there are several schedulers. One of them chooses the initial values and the updated values for the module external variables; for each

atom, intended as a cluster of variables of the module, a further scheduler chooses the initial values and the updated values for the private and interface variables controlled by that atom. Compose-and-schedule is thus replaced by schedule-and-compose.

Distributed scheduling was then applied in (Cheung *et al.*, 2006) to the *asynchronous* model of switched probabilistic input/output automata. Following the terminology of (van Glabbeek *et al.*, 1995), given a reactive interpretation to input actions and a generative interpretation to output actions, an input scheduler and an output scheduler are considered for each automaton occurring in a system. A token passing mechanism among the automata eliminates global choices by ensuring that a single automaton at a time can select a generative output action, to which the other automata can respond with reactive input actions having the same name.

Both (de Alfaro *et al.*, 2001) and (Cheung *et al.*, 2006) guarantee the compositionality of the probabilistic trace-distribution equivalence of (Segala, 1995b), which is *not* a congruence with respect to parallel composition under centralized scheduling. As shown in (Lynch *et al.*, 2003), the coarsest congruence contained in that linear-time equivalence turns out to be a variant of the simulation equivalence of (Segala and Lynch, 1994), which is a branching-time equivalence.

Distributed scheduling was further studied in (Giro and D'Argenio, 2007; Giro and D'Argenio, 2009) for interleaved probabilistic input/output automata, a variant of switched ones in which an interleaving scheduler replaces the token passing mechanism. The examined problem was the attainment of the extremal probabilities of satisfying reachability properties under different classes of distributed schedulers (memoryless vs. history-dependent, deterministic vs. randomized), knowing that in the centralized case those probabilities are obtained when using memoryless deterministic schedulers (Bianco and de Alfaro, 1995).

The overwhelming power of schedulers already shows up in the *memoryless* case, i.e., when neglecting the path followed to reach the current state. Under memoryless schedulers, a different definition of probabilistic trace equivalence allows compositionality to be recovered without resorting to distributed scheduling.

In the probabilistic trace-distribution equivalence of (Segala, 1995b), for each resolution of either process there must exist a resolution of the other process such that the two resolutions are *fully matching*, in the sense that, for every trace, both resolutions feature the same probability of executing that trace. This is called probabilistic trace *post*-equivalence as the quantification over traces occurs *after* the quantifications over resolutions, which is a source of overdiscrimination.

In (Bernardo *et al.*, 2014a) it was proposed to exchange the order of those quantifications, which avoids hardly justifiable process distinctions and regains compositionality. Given an arbitrary trace, for each resolution of either process there must exist a resolution of the other process such that both of them exhibit the same probability of executing that trace. In this case, resolutions are *partially matching*, as a resolution of either process can be matched by different resolutions of the other process with respect to different traces. The resulting relation is called probabilistic trace *pre*-equivalence because the quantification over traces occurs *before* the quantifications over resolutions.

On the other hand, the probabilistic testing equivalences of (Yi and Larsen, 1992; Jonsson and Yi, 1995; Segala, 1996) are *not* backward compatible with testing equivalences for simpler processes such as fully nondeterministic ones (De Nicola and Hennessy, 1984) and fully probabilistic ones (Cleaveland *et al.*, 1999).

Indeed, in (Jonsson and Yi, 2002; Deng *et al.*, 2008) it was shown that those equivalences can be characterized in terms of branching-time, simulation-like relations, which is consistent with the fact that they are *not* insensitive to the moment of occurrence of nondeterministic or probabilistic choices among identical actions. In addition to centralized scheduling, this is a consequence of a special instance of the *copying capability* (Abramsky, 1987), which shows up in the presence of a nondeterministic choice in either component that synchronizes with a probabilistic choice in the other, thus creating copies of a state possessing several outgoing transitions, where different decisions can be made.

Under centralized scheduling, in (Georgievska and Andova, 2012) additional labels were used so that the same decision is made by schedulers in distinct copies of the same state of a testing system, which weakens the discriminating power of the probabilistic testing equiva-

lences of (Yi and Larsen, 1992; Jonsson and Yi, 1995; Segala, 1996). An analogous weakening result under the same class of schedulers was obtained in (Bernardo *et al.*, 2014a) by means of a different definition of probabilistic testing equivalence, in which success probabilities are compared in a trace-by-trace fashion rather than cumulatively. Instead of the overall success probability, the probability of reaching success is examined separately for each possible trace.

1.3 Coherent Resolutions of Nondeterminism

Being a congruence with respect to parallel composition, which is ensured by distributed scheduling (de Alfaro *et al.*, 2001; Cheung *et al.*, 2006) as well as partially matching resolutions (Bernardo *et al.*, 2014a), is not the only desirable property of probabilistic trace equivalences. In addition to compositionality with respect to other typical process operators, it is necessary to address the inclusion of the probabilistic bisimilarity of (Segala and Lynch, 1994) together with the backward compatibility with respect to trace equivalences over less expressive models, such as fully nondeterministic processes (Brookes *et al.*, 1984) and fully probabilistic processes (Jou and Smolka, 1990).

We will see that the validity of the aforementioned properties of trace semantics, as well as the possibility of enhancing the backward compatibility of testing semantics, critically depend on the capability of limiting the freedom of schedulers and can be achieved if we restrict ourselves to *coherent resolutions* of nondeterminism. Similar to (Georgievska and Andova, 2012), the basic idea is that schedulers cannot select different continuations in states of a process that are equivalent to each other, so that the states to which they correspond in any resolution of the process also have equivalent continuations.

As a preliminary step towards the study of the impact of resolution coherency on the discriminating power, on the compositionality, and on the backward compatibility of probabilistic trace and testing equivalences, we will provide a uniform way of defining the resolutions induced by different subclasses of centralized, memoryless schedulers. In particular, we formalize any resolution as a fully probabilistic automaton, which we equip with a *correspondence function* from the acyclic

state space of the resolution to the possibly cyclic state space of the original automaton. This technique was introduced for the first time in (Jonsson *et al.*, 1994) for deterministic schedulers.

We divide resolutions into *structure preserving* and *structure modifying*, depending on whether they respect or alter the structure of the automaton from which they are obtained. A structure-preserving resolution is produced by a *deterministic scheduler*, which selects at the current state one of the transitions departing from that state or no transitions at all. A structure-modifying resolution is derived via a *randomized scheduler* (Segala, 1995a), which probabilistically combines the transitions departing from the current state, or an *interpolating scheduler* (Deng *et al.*, 2007), which splits the current state into copies, each having at most one outgoing transition and whose probabilities sum up to the probability of the original state.

We will then present a number of anomalies affecting the probabilistic trace equivalences of (Segala, 1995b) and (Bernardo *et al.*, 2014a), mostly arising under deterministic schedulers. More precisely, we show that they do not contain probabilistic bisimilarity, are not congruences with respect to action prefix, and are not backward compatible with their versions for fully probabilistic models. The reason is that schedulers have the freedom to make *different* decisions in *equivalent* states occurring in the target distribution of a transition, with these decisions not necessarily replicable in equivalent distributions of distinct automata. This is especially true for deterministic schedulers, as the resolutions they induce must be structure preserving.

Such anomalies can be avoided by employing coherent resolutions in the definition of probabilistic trace equivalences. If several states in the target distribution of a transition are equivalent, then the states to which they correspond in a resolution must be equivalent as well. The coherency constraints can be formalized by reasoning on *coherent trace distributions*, i.e., suitable families of sets of traces weighted with their execution probabilities in a given resolution.

In the case of testing semantics, coherency will be accompanied by additional transition decorations, so that the same decisions are made by schedulers in distinct copies of the same state of a process or a test occurring in a choice within the testing system. This is similar to the

technique employed in (Georgievska and Andova, 2012) for processes in which branchings based on actions, nondeterminism, and probabilities alternate, with the remarkable difference that our decoration procedure turns out to be much simpler.

The resulting probabilistic testing equivalence retrieves insensitivity to the moment of occurrence of nondeterministic or probabilistic choices among identical actions, thus enhancing backward compatibility with respect to (Yi and Larsen, 1992; Jonsson and Yi, 1995; Segala, 1996). Consistent with the ready-trace semantics characterization of (Georgievska and Andova, 2012), a counterexample inspired by failure semantics for fully nondeterministic processes shows that complete backward compatibility cannot be achieved in the presence of certain synchronizations among external choices, a fact that has nothing to do with coherency.

1.4 Alternative Characterizations

In a fully nondeterministic setting, two processes are trace equivalent if and only if, for each trace α, both processes can perform α or neither can. An immediate alternative characterization is that two trace equivalent processes possess the same trace set (Brookes *et al.*, 1984), where this set can be viewed as the language accepted by the automata underlying those processes. Likewise, two fully probabilistic processes are trace equivalent if and only if, for each trace α, both processes can perform α with the same probability, which amounts to possessing the same set of traces each weighted with its execution probability (Jou and Smolka, 1990), i.e., the same probabilistic language. In either case, process equivalence reduces to (possibly weighted) trace set equality.

Straightforward characterizations of that form are not possible in the case of nondeterministic and probabilistic processes, because (i) traces can have different execution probabilities in different coherent resolutions and (ii) trace semantics can be defined according to different approaches leading to probabilistic trace post-/pre-equivalences. This motivates the investigation of alternative characterizations for the two aforementioned equivalences under coherent resolutions arising from centralized, memoryless schedulers. We will see that the coherency-based

variant of the probabilistic trace post-equivalence of (Segala, 1995b) can be characterized in terms of the coherent trace distributions used for defining the coherency constraints. In contrast, since it treats traces individually without keeping track of the resolutions in which they can be executed, the coherency-based variant of the probabilistic trace pre-equivalence of (Bernardo *et al.*, 2014a) can be characterized by something weaker, which is constituted by coherent weighted trace sets.

1.5 Outline

This work is an extended, revised, and integrated version of (Bernardo, 2019a; Bernardo, 2020a; Bernardo, 2020b), which is organized as follows. In Sect. 2 we recall the simple probabilistic automaton model and its specializations to fully nondeterministic and fully probabilistic models. In Sect. 3 we survey different ways of resolving nondeterminism in the aforementioned model, which preserve or modify the model structure, and provide a uniform manner of defining all of them. In Sect. 4 we present an overview of different approaches to probabilistic behavioral equivalences and then recall the formal definitions of probabilistic bisimulation equivalence, probabilistic trace post-/pre-equivalences, and probabilistic testing equivalence. In Sect. 5 we illustrate three anomalies of the two probabilistic trace equivalences related to their discriminating power, their compositionality, and their backward compatibility. In Sect. 6 we show how to avoid those anomalies by resorting to coherent resolutions, which are formulated in terms of coherency constraints based on coherent trace distributions. In Sect. 7 we develop alternative characterizations of the coherency-based variants of the two probabilistic trace equivalences, respectively relying on coherent trace distributions and coherent weighted trace sets, and use them to express some considerations about congruence with respect to parallel composition. In Sect. 8 we illustrate that the backward compatibility of probabilistic testing equivalence is only partial due to the sensitivity to the moment of occurrence of nondeterministic or probabilistic choices among identical actions. In Sect. 9 we show how to enhance compatibility through the combined use of coherent resolutions and suitable transition decorations. Finally, in Sect. 10 we provide some concluding remarks.

2

Nondeterministic and Probabilistic Models

Processes featuring nondeterminism and probability can be described by extending the labeled transition system (LTS) model of (Keller, 1976), in such a way that every action-labeled transition goes from a source state to a probability distribution over target states (Larsen and Skou, 1991; Segala, 1995a) rather than to a single target state. The resulting models are essentially Markov decision processes (Derman, 1970), or probabilistic automata in the sense of (Rabin, 1963), that additionally allow for internal nondeterminism, i.e., the presence of equally labeled transitions departing from the same state.

In the literature, these models have been represented through a number of slightly different computational entities such as, e.g., concurrent Markov chains (Vardi, 1985), strictly alternating models (Hansson and Jonsson, 1990), probabilistic automata in the sense of (Segala, 1995a), and the denotational probabilistic models of (Jifeng *et al.*, 1997); see (Sokolova and de Vink, 2004) for an overview. We formalize them through a variant of simple probabilistic automata (Segala, 1995a), in which we do not distinguish between external and internal actions.

Definition 2.1. A *nondeterministic and probabilistic labeled transition system*, NPLTS for short, is a triple $(S, A, \longrightarrow)$ where:

- $S \neq \emptyset$ is an at most countable set of states.
- $A \neq \emptyset$ is a countable set of transition-labeling actions.
- $\longrightarrow \subseteq S \times A \times \mathit{Distr}(S)$ is a transition relation, with $\mathit{Distr}(S) = \{\Delta : S \to \mathbb{R}_{[0,1]} \mid \sum_{s \in S} \Delta(s) = 1\}$ being the set of discrete probability distributions over S. ■

A transition (s, a, Δ) is written $s \xrightarrow{a} \Delta$. We say that $s' \in S$ is not reachable from s via that a-transition if $\Delta(s') = 0$, otherwise we say that it is reachable with probability $p = \Delta(s')$. The reachable states form the support of the target distribution Δ, i.e., $\mathit{supp}(\Delta) = \{s' \in S \mid \Delta(s') > 0\}$. An NPLTS can be depicted as a directed graph in which vertices represent states and action-labeled edges represent transitions, with states in the support of the same target distribution being linked by a dashed line and decorated with the respective probabilities when these are different from 1.

For instance, in the forthcoming Fig. 3.1, the NPLTS with initial state s has an a-transition to a distribution whose support includes s'_1 and s'_2, each of which is reachable with probability 0.5. Then s'_1 has a b-transition to a distribution whose support includes only s''_1, whereas s'_2 has a c-transition to a distribution whose support includes only s''_2, hence both s''_1 and s''_2 are reachable with probability 1.

The nondeterministic choice among all the transitions departing from a state can be influenced by the external environment, while the probabilistic choice among the target states of the selected transition takes place internally. An NPLTS represents:

- A fully nondeterministic process when every transition has a target distribution with a singleton support (see, e.g., the leftmost and the rightmost NPLTS in Fig. 3.2).
- A fully probabilistic process when every state has at most one outgoing transition (see, e.g., the leftmost NPLTS in Fig. 3.1).
- A Markov decision process when, for each action, any state has at most one outgoing transition labeled with that action (see, e.g., the rightmost NPLTS in Fig. 4.1), implying the absence of internal nondeterminism (present in the other two NPLTS models).

In this setting, a computation is a sequence of state-to-state steps, each denoted by $s \stackrel{a}{\longmapsto} s'$ and derived from a state-to-distribution transition $s \stackrel{a}{\longrightarrow} \Delta$. Let A^* be the set of traces, i.e., finite action sequences, and ε be the empty trace.

Definition 2.2. Let $\mathcal{L} = (S, A, \longrightarrow)$ be an NPLTS and $s, s' \in S$. We say that the finite sequence of steps:

$$c \equiv s_0 \stackrel{a_1}{\longmapsto} s_1 \stackrel{a_2}{\longmapsto} s_2 \ldots s_{n-1} \stackrel{a_n}{\longmapsto} s_n$$

is a *computation* of $\mathcal{L}$ of length $n \in \mathbb{N}$ from $s = s_0$ to $s' = s_n$ *compatible* with trace $\alpha = a_1\, a_2 \ldots a_n \in A^*$, written $c \in \mathcal{CC}(s, \alpha)$, iff for each step $s_{i-1} \stackrel{a_i}{\longmapsto} s_i$ in c there exists a transition $s_{i-1} \stackrel{a_i}{\longrightarrow} \Delta_i$ in $\mathcal{L}$ such that $s_i \in \mathit{supp}(\Delta_i)$, $1 \leq i \leq n$, where:

- $\Delta_i(s_i)$ is the execution probability of step $s_{i-1} \stackrel{a_i}{\longmapsto} s_i$ conditioned on the selection of transition $s_{i-1} \stackrel{a_i}{\longrightarrow} \Delta_i$ at state s_{i-1}, or simply the execution probability of that step if $\mathcal{L}$ is fully probabilistic.
- $\mathit{prob}(c) = \prod_{1 \leq i \leq n} \Delta_i(s_i)$ is the execution probability of c if $\mathcal{L}$ is fully probabilistic, assuming that $\mathit{prob}(c) = 1$ when $n = 0$.
- $\mathit{prob}(C) = \sum_{c \in C} \mathit{prob}(c)$ for $C \subseteq \bigcup_{\alpha \in A^*} \mathcal{CC}(s, \alpha)$ if $\mathcal{L}$ is fully probabilistic, provided that no computation in C is a proper prefix of one of the others. ■

For example, again in Fig. 3.1, the initial state s of the leftmost NPLTS features the empty computation with probability 1, computations $s \stackrel{a}{\longmapsto} s'_1$ and $s \stackrel{a}{\longmapsto} s'_2$ each with probability 0.5, and computations $s \stackrel{a}{\longmapsto} s'_1 \stackrel{b}{\longmapsto} s''_1$ and $s \stackrel{a}{\longmapsto} s'_2 \stackrel{c}{\longmapsto} s''_2$ each with probability $0.5 \cdot 1 = 0.5$. Note that $s \stackrel{a}{\longmapsto} s'_1$ (resp. $s \stackrel{a}{\longmapsto} s'_2$) is a proper prefix of $s \stackrel{a}{\longmapsto} s'_1 \stackrel{b}{\longmapsto} s''_1$ (resp. $s \stackrel{a}{\longmapsto} s'_2 \stackrel{c}{\longmapsto} s''_2$) as well as the empty computation is a proper prefix of any other computation.

3

An Overview of Resolutions of Nondeterminism

When several transitions depart from a state s of an NPLTS $\mathcal{L}$, they describe a nondeterministic choice among different behaviors. The choice is called internal or external depending on whether all actions labeling those transitions are equal or not. Probabilistic behavioral equivalences compare numeric values extracted from the processes at hand after resolving every nondeterministic choice.

A *resolution* of s is the outcome of a possible way of resolving nondeterministic choices starting from s, as if a *scheduler* were applied that decides which activity has to be performed next. A resolution of nondeterminism can thus be formalized as a fully probabilistic NPLTS $\mathcal{Z}$ with a tree-like structure, whose branching points come from the target distributions of the transitions selected among those of $\mathcal{L}$.

We now present an overview of the various ways to resolve nondeterminism according to a *centralized, memoryless strategy*, i.e., when using a single scheduler whose choices do not depend on those made in the past. This will be accomplished by providing a uniform technique for defining all the aforementioned ways based on *correspondence functions*, so as to facilitate their comparison.

In particular, we address the notions of resolution arising from two different approaches, respectively preserving or modifying the structure of the original NPLTS. The idea underlying the former approach is to construct a resolution by *importing states and transitions* from the original model (Sect. 3.1). The idea at the basis of the latter approach is that (i) a transition of a resolution can be produced by *probabilistically combining transitions* of the original model (Sect. 3.2) or (ii) a state of a resolution can be obtained by *probabilistically splitting states* of the original model (Sect. 3.3).

3.1 Structure-Preserving Resolutions via Deterministic Schedulers

A *deterministic scheduler* selects one of the transitions departing from the current state, or no transitions at all thus stopping the execution. As a consequence, the resulting resolution is isomorphic to a submodel of the original model (or of its unfolding, should cycles be present), thereby *preserving* the structure of the original model (or of its unfolding). If the model is fully nondeterministic, each of its resolutions corresponds to a computation of the model; if the model is fully probabilistic, its maximal resolution corresponds to the entire model.

For instance, starting from the initial state s' of the rightmost NPLTS in Fig. 3.1, a deterministic scheduler may decide to immediately stop the execution or select the leftmost, the central, or the rightmost a-transition. If the central a-transition is chosen, the scheduler may then decide to stop the execution at that point or perform the b-transition, the c-transition, or both (which is possible as the two transitions depart from two different states).

In (Yi and Larsen, 1992) a resolution was defined as a maximal subtree of the unfolding of the considered model – with the unfolding yielding a potentially infinite tree – in which every state has at most one outgoing transition. Resolutions were defined as fully probabilistic maximal subtrees also in (Jonsson and Yi, 1995), but the considered models were finite trees in lieu of directed graphs. Subtree maximality was required in those works because of their focus on testing semantics.

The paper (Jonsson *et al.*, 1994), instead of reasoning in terms of unfoldings and submodels, introduced for the first time a *correspondence*

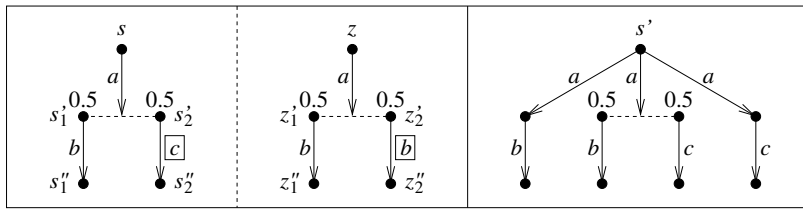

Figure 3.1: Lack of injectivity leads to structure violation and wrong identifications

function $corr_{\mathcal{Z}} : Z \to S$ from the acyclic state space of the resolution $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}})$ being built, to the possibly cyclic state space of the considered model $\mathcal{L} = (S, A, \longrightarrow)$. For each transition $z \xrightarrow{a}_{\mathcal{Z}} \Delta$, the function $corr_{\mathcal{Z}}$ had to preserve the probabilities of all the states corresponding to those in $supp(\Delta)$. In other words, it had to satisfy the following constraint on transitions: if $z \xrightarrow{a}_{\mathcal{Z}} \Delta$ then $corr_{\mathcal{Z}}(z) \xrightarrow{a} \Gamma$, with $\Delta(z') = \Gamma(corr_{\mathcal{Z}}(z'))$ for all $z' \in supp(\Delta)$.

The correspondence function with its constraint as defined in (Jonsson *et al.*, 1994) and reused in (Bernardo *et al.*, 2014a; Bernardo *et al.*, 2014b) has the drawback of not being structure preserving in the case that the target distribution of a transition assigns the same probability to several inequivalent states.

Let us see for instance the three NPLTS models in Fig. 3.1, where s'_1 and s'_2 enable different actions. The correspondence function that maps z to s, z'_1 and z'_2 to s'_1, and z''_1 and z''_2 to s''_1 causes the central NPLTS to be considered a resolution of the leftmost NPLTS, although the former is not isomorphic to any submodel of the latter because z'_1 and z'_2 enable the same action. This may have no consequences on the discriminating power of testing equivalences, the subject of (Jonsson *et al.*, 1994), if all transitions of testing systems are identically labeled. However, it would lead to consider the leftmost NPLTS and the rightmost NPLTS as trace equivalent, because also the leftmost one would have a resolution in which trace $a\,b$ (resp. trace $a\,c$) is executable with probability 1.

The constraint was rectified in (Bernardo *et al.*, 2014c) by additionally requiring the *injectivity* of $corr_{\mathcal{Z}}$ over $supp(\Delta)$, so that in Fig. 3.1 z'_1 and z'_2 can no longer be both mapped to s'_1. We also point out that in (Bernardo, 2019b) it was further observed that *bijectivity* between

$supp(\Delta)$ and $supp(\Gamma)$, rather than injectivity over $supp(\Delta)$, is necessary to preserve the overall reachability mass of the target of any transition – which boils down to the total probability 1 in the NPLTS model – in a more general setting like the ULTRAS metamodel.

Below is the rectified definition of (Bernardo *et al.*, 2014c) in the style of (Jonsson *et al.*, 1994), i.e., based on a correspondence function from the acyclic state space of the resolution to the possibly cyclic state space of the considered model, which is required to be injective in the first clause. The second clause ensures that the resolution is a fully probabilistic NPLTS, i.e., a model without nondeterministic choices.

Definition 3.1. Let $\mathcal{L} = (S, A, \longrightarrow)$ be an NPLTS and $s \in S$. An acyclic NPLTS $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}})$ is a *structure-preserving resolution* of s, written $\mathcal{Z} \in Res_{\rm sp}(s)$, iff there exists a correspondence function $corr_{\mathcal{Z}} : Z \to S$ such that $s = corr_{\mathcal{Z}}(z_s)$, for some $z_s \in Z$ acting as the initial state of $\mathcal{Z}$, and for all $z \in Z$ it holds that:

- If $z \xrightarrow{a}_{\mathcal{Z}} \Delta$ then $corr_{\mathcal{Z}}(z) \xrightarrow{a} \Gamma$, with $corr_{\mathcal{Z}}$ being injective over $supp(\Delta)$ and satisfying $\Delta(z') = \Gamma(corr_{\mathcal{Z}}(z'))$ for all $z' \in supp(\Delta)$.
- At most one transition departs from z. ■

For example, since the leftmost NPLTS in Fig. 3.1 is fully probabilistic and acyclic, its maximal structure-preserving resolution coincides with the considered NPLTS. This is shown by the correspondence function that maps every state to itself, which in particular is injective over the support of the target distribution of the only transition whose target support is not a singleton, i.e., the a-transition.

3.2 Structure-Modifying Resolutions via Randomization

Randomized schedulers, proposed in (Segala, 1995a) and applied to the definition of probabilistic trace (Segala, 1995b) and testing (Segala, 1996) semantics, probabilistically combine transitions of the original model. Therefore, the resulting resolutions are not necessarily isomorphic to submodels of the original model (or of its unfolding) because a *modification* of the original model structure may have taken place.

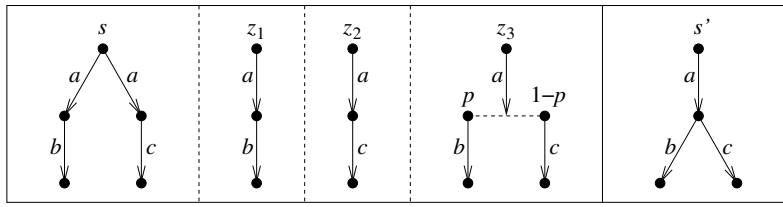

Figure 3.2: An example of structure modification induced by a randomized scheduler

An example of this phenomenon is shown in Fig. 3.2, where the NPLTS in the leftmost part admits under randomized schedulers the three maximal resolutions depicted next to it in the figure. The resolution starting with z_3 is obtained as the combination of the two a-transitions departing from s by using probabilities p and $1 - p$ respectively.

In general, if the current state has $n \in \mathbb{N}_{\geq 1}$ outgoing transitions, a *randomized scheduler* generates the numeric values $p_i \in \mathbb{R}_{[0,1]}$ for $i = 1, \ldots, n$ such that $\sum_{i=1}^{n} p_i \leq 1$ and then selects transition i with probability p_i, or stops the execution with probability $1 - \sum_{i=1}^{n} p_i$. A deterministic scheduler is a special case of randomized scheduler in which $p_i = 1$ for some i or $p_i = 0$ for each i.

The formalization via a correspondence function of a resolution stemming from a randomized scheduler is not an easy task. The reason is that, according to (Segala, 1995a), a combined transition may derive from several *differently labeled* transitions, as shown in the central part of the forthcoming Fig. 3.3. In other words, a resolution of a simple probabilistic automaton (Segala, 1995a), in which every transition has a single label, may have a transition with *several* labels, thereby deviating from a simple probabilistic automaton and hence from an NPLTS.

Similar to (Bernardo *et al.*, 2014a), below we formalize a resolution induced by a variant of randomized scheduler in accordance with the definition of probabilistic bisimilarity given in (Segala and Lynch, 1994) for simple probabilistic automata. At the current state, the scheduler decides to stop or to perform a certain action among the available ones; in the latter case, it takes a convex combination (i.e., the sum of the values p_i is 1) of the outgoing transitions *identically labeled* with that action. To compensate for the impossibility of combining differently labeled

transitions, we admit self-combinations; e.g., in Fig. 3.3 a combination of the a-transition departing from s with itself n times is able to reproduce the situation in the rightmost part of the same figure, which is equivalent to the one in the central part.

Definition 3.2. Let $\mathcal{L} = (S, A, \longrightarrow)$ be an NPLTS and $s \in S$. An acyclic NPLTS $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}})$ is a *structure-modifying resolution via randomization* of s, written $\mathcal{Z} \in Res_{\mathrm{sm,r}}(s)$, iff there exists a correspondence function $corr_{\mathcal{Z}} : Z \to S$ such that $s = corr_{\mathcal{Z}}(z_s)$, for some $z_s \in Z$ acting as the initial state of $\mathcal{Z}$, and for all $z \in Z$ it holds that:

- If $z \xrightarrow{a}_{\mathcal{Z}} \Delta$ then there exist $n \in \mathbb{N}_{\geq 1}$, $p_i \in \mathbb{R}_{]0,1]}$ for $1 \leq i \leq n$ summing up to 1, and $corr_{\mathcal{Z}}(z) \xrightarrow{a} \Gamma_i$ for $1 \leq i \leq n$, with $corr_{\mathcal{Z}}$ being injective when considered from $supp(\Delta)$ to the disjoint union of the sets $supp(\Gamma_i)$ and satisfying $\Delta(z') = \sum_{i=1}^{n} p_i \cdot \Gamma_i(corr_{\mathcal{Z}}(z'))$ for all $z' \in supp(\Delta)$.
- At most one transition departs from z. ■

For instance, in Fig. 3.2 the NPLTS with initial state z_3 is a structure-modifying resolution via randomization of both the NPLTS with initial state s and the NPLTS with initial state s'. In the former case, the a-transition in the resolution comes from the convex combination of the two a-transitions in the original model, respectively taken with probabilities p and $1 - p$. In the latter case, the a-transition in the resolution instead comes from the self-combination of the only a-transition of the original model, taken with the same two probabilities.

In Def. 3.2 injectivity cannot be directly imposed as in Def. 3.1 – the disjoint union of the supports of the target distribution of the original transitions has to be considered – otherwise in Fig. 3.2 the NPLTS model starting with z_3 would not be a legal resolution induced by the self-combination of the a-transition departing from s', and hence s' would not be deemed trace equivalent to s.

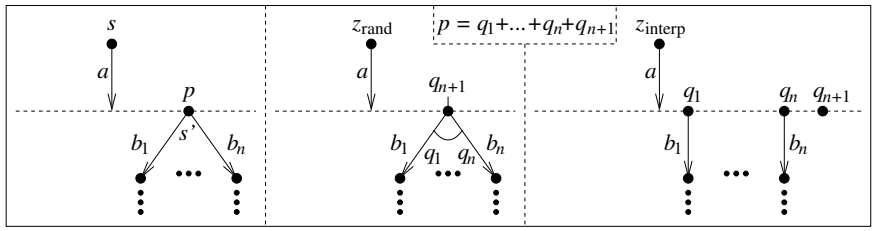

Figure 3.3: Equivalent resolutions induced by randomized/interpolating schedulers

3.3 Structure-Modifying Resolutions via Interpolation

Interpolating schedulers, proposed in (Deng *et al.*, 2007), probabilistically split states of the original model thereby inducing resolutions possibly modifying the structure of the original model. As mentioned in (Deng *et al.*, 2007), interpolating and randomized schedulers are closely related: for each resolution obtained from an interpolating (resp. randomized) scheduler, there exists a resolution obtained from a randomized (resp. interpolating) scheduler with the same trace distribution.

This can be seen in Fig. 3.3, where in the leftmost part we have a state s' reached with probability p in the target distribution of an a-transition. The resolution in the central part, induced by a randomized scheduler that combines the transitions departing from s', is trace equivalent to the resolution in the rightmost part, induced by an interpolating scheduler that splits state s', where $\sum_{i=1}^{n+1} q_i = p$.

For every state in the support of the target distribution of the current transition, an *interpolating scheduler* splits the state into $n \in \mathbb{N}_{\geq 1}$ copies, each having at most one outgoing transition, to which probabilities are assigned whose sum is the overall probability of the original state, and then selects one of the copies based on its probability. A deterministic scheduler is a special case of interpolating scheduler in which $n = 1$.

Resolutions arising from interpolating schedulers were natively defined in (Deng *et al.*, 2007) through a correspondence function that maps all split states to the original state from which they derive. Unlike Defs. 3.1 and 3.2, the constraint on transitions is formulated with respect to the states in the support of the corresponding transition of the original model – rather than the states in the support of the transition of the

resolution – and the preservation of the overall probability associated with each such state makes injectivity requirements unnecessary.

Definition 3.3. Let $\mathcal{L} = (S, A, \longrightarrow)$ be an NPLTS and $s \in S$. An acyclic NPLTS $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}})$ is a *structure-modifying resolution via interpolation* of s, written $\mathcal{Z} \in Res_{\mathrm{sm,i}}(s)$, iff there exists a correspondence function $corr_{\mathcal{Z}} : Z \to S$ such that $s = corr_{\mathcal{Z}}(z_s)$, for some $z_s \in Z$ acting as the initial state of $\mathcal{Z}$, and for all $z \in Z$ it holds that:

- If $z \xrightarrow{a}_{\mathcal{Z}} \Delta$ then $corr_{\mathcal{Z}}(z) \xrightarrow{a} \Gamma$, with $corr_{\mathcal{Z}}$ satisfying $\Gamma(s) = \sum_{z' \in supp(\Delta)}^{corr_{\mathcal{Z}}(z')=s} \Delta(z')$ for all $s \in supp(\Gamma)$.
- At most one transition departs from z. ■

For instance, in Fig. 3.2 the NPLTS with initial state z_3 is a structure-modifying resolution via interpolation of the NPLTS with initial state s'. In particular, the state with two outgoing transitions reached by the a-transition departing from s' has been split into two states with a single transition each, which respectively correspond to the two states reached by the a-transition departing from z_3. Formally, the correspondence function maps the last two states to the state being split. Note that the function is not injective, but this is not a problem because the constraint in the first clause checks for all states s whether $\Gamma(s)$ is the sum of all values $\Delta(z')$ for every z' to which s corresponds.

A variant of the notion of structure-modifying resolution in Def. 3.3 has been proposed in (Bonchi *et al.*, 2019), which combines the effects of interpolating and randomized schedulers.

4

Behavioral Equivalences for NPLTS Models

Many approaches to the definition of behavioral relations have appeared in the literature, together with the investigation of their compositional, equational, and logical characteristics. They have been the subject of comparative concurrency theory (van Glabbeek, 2001), which studies the discriminating power and the mutual relationships of behavioral relations. In the specific case of nondeterministic and probabilistic processes, the spectrum of behavioral equivalences has been examined in (Bernardo *et al.*, 2014b).

When applied to NPLTS models, the three major approaches – bisimilarity (Park, 1981; Milner, 1989), trace semantics (Brookes *et al.*, 1984), and testing semantics (De Nicola and Hennessy, 1984) – rely on comparing the probabilities of equivalence-specific events after removing nondeterminism. The aforementioned events are the reachability via given actions of certain sets of equivalent states (bisimilarity, Sect. 4.1), the execution of specific action sequences (trace semantics, Sect. 4.2), and the passing of tests (testing semantics, Sect. 4.3).

Regardless of the approach, unlike fully nondeterministic and fully probabilistic processes, there are at least three alternative ways of applying a behavioral equivalence to nondeterministic and probabilistic

processes, based on how the resolutions of nondeterminism of those processes are employed:

- The first option, coming from (Segala and Lynch, 1994; Segala, 1995b), examines the probability distributions of *all* equivalence-specific events calculated over resolutions. Two processes are considered equivalent if, for each resolution of either process, there exists a resolution of the other process such that the probability of *each* equivalence-specific event is the same in the two resolutions (*fully matching resolutions*). The resulting portion of the spectrum closely resembles the spectrum for fully probabilistic processes. This option will be exemplified in Defs. 4.1 and 4.2.

- The second option, deriving from (Tracol *et al.*, 2011; Song *et al.*, 2013; Bernardo *et al.*, 2014a; Bernardo *et al.*, 2015), compares resolutions on the basis of the probabilities of *individual* equivalence-specific events. A resolution of either process can be matched, with respect to *different* equivalence-specific events, by *different* resolutions of the other process (*partially matching resolutions*). The resulting equivalences are less discriminating than those arising from fully matching resolutions, retrieve simple logical characterizations for bisimulation semantics and useful compositionality properties for trace semantics, and yield a portion of the spectrum featuring many analogies with the spectrum for fully nondeterministic processes. This option will be exemplified in Def. 4.3.

- The third option, stemming from the testing theories in (Yi and Larsen, 1992; Jonsson and Yi, 1995; Segala, 1996) and adapted to other semantics in (Bernardo *et al.*, 2014b), instead of comparing individual resolutions, takes into account only the *extremal probabilities* of equivalence-specific events computed over all the resolutions of either process (*max-min matching resolutions*). The resulting equivalences are less discriminating than the ones originated from partially matching resolutions, but induce a similar portion of the spectrum. This option will be exemplified in Def. 4.6.

4.1 Bisimulation Semantics: $\sim_{\mathbf{PB}}$

Although the focus of this monograph is on trace and testing semantics, we start by recalling the definition of bisimulation semantics because one of the anomalies we will examine has to do with the inclusion of the last semantics in the first two.

According to (Larsen and Skou, 1991), probabilistic bisimilarity requires that two processes are able to mimic each other behavior stepwise, in terms of the probability of reaching the same class of equivalent states when executing the same action. Its application to the NPLTS model does not need to explicitly resort to resolutions, as these are implicitly built while selecting a transition from each considered state (Segala and Lynch, 1994). In the following, we let $\Delta(C) = \sum_{s \in C} \Delta(s)$ for $C \subseteq S$.

Definition 4.1. Let $(S, A, \longrightarrow)$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim_{\mathrm{PB}} s_2$ iff there exists a probabilistic bisimulation $\mathcal{B}$ over S such that $(s_1, s_2) \in \mathcal{B}$. An equivalence relation $\mathcal{B}$ over S is a *probabilistic bisimulation* iff, whenever $(s_1, s_2) \in \mathcal{B}$, then for all $a \in A$ it holds that for each $s_1 \xrightarrow{a} \Delta_1$ there exists $s_2 \xrightarrow{a} \Delta_2$ such that for all equivalence classes $C \in S/\mathcal{B}$:

$$\Delta_1(C) \;=\; \Delta_2(C)$$

■

For example, in the forthcoming Fig. 5.1 the NPLTS with initial state s_1 is probabilistic bisimilar to the NPLTS with initial state s_2. This is witnessed by the probabilistic bisimulation $\mathcal{B}$ resulting from the reflexive, symmetric, and transitive closure of the relation containing the pairs $(s_1, s_2), (s_1', s_2'), (s_1', s_2'')$ along with the pairs formed by a terminal state of the former NPLTS reached by b (resp. c) and a terminal state of the latter NPLTS reached by b (resp. c). In particular, when s_1 performs its a-transition and reaches the equivalence class $\{s_1', s_2', s_2''\}$ with probability 1, then s_2 can respond with its a-transition and reach the same equivalence class with overall probability $p + (1 - p)$, i.e., with the same probability as s_1, and vice versa.

4.2 Trace Semantics: $\sim^{\mathbf{post}}_{\mathbf{PTr}}$ and $\sim^{\mathbf{pre}}_{\mathbf{PTr}}$

Unlike bisimulation semantics, trace semantics abstracts from branching points of process behavior and explicitly rely on $Res(\_)$, with which we denote any of the sets of resolutions introduced in Defs. 3.1 to 3.3. While there is only one way of defining trace semantics for fully nondeterministic processes (Brookes *et al.*, 1984) and for fully probabilistic processes (Jou and Smolka, 1990), this is not the case with processes featuring both nondeterminism and probabilities.

The first probabilistic trace equivalence that we consider is the one of (Segala, 1995b). Two states are deemed equivalent when every resolution of either state is matched by a resolution of the other, in the sense that for each trace both resolutions execute that trace with the same probability. We call it probabilistic trace *post*-equivalence because the quantification over traces occurs *after* selecting the two *fully matching* resolutions, as underlined in the definition below where z_{s_i} denotes both the initial state of $\mathcal{Z}_i$ and the state to which s_i corresponds.

Definition 4.2. Let $(S, A, \longrightarrow)$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim^{\text{post}}_{\text{PTr}} s_2$ iff for each $\mathcal{Z}_1 \in Res(s_1)$ there exists $\mathcal{Z}_2 \in Res(s_2)$ such that for all $\alpha \in A^*$:

$$prob(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; prob(\mathcal{CC}(z_{s_2}, \alpha))$$

and the same condition holds when exchanging $\mathcal{Z}_1$ with $\mathcal{Z}_2$. ■

The second probabilistic trace equivalence is the one of (Bernardo *et al.*, 2014a). It is less restrictive than the previous equivalence, hence it avoids some hardly justifiable distinctions and, most importantly, turns out to be a congruence with respect to parallel composition. Two states are deemed equivalent when a resolution of either state can be matched by possibly different resolutions of the other with respect to different traces. We call it probabilistic trace *pre*-equivalence because traces are fixed *before* selecting the two *partially matching* resolutions.

Definition 4.3. Let $(S, A, \longrightarrow)$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim^{\text{pre}}_{\text{PTr}} s_2$ iff, for all $\alpha \in A^*$, for each $\mathcal{Z}_1 \in Res(s_1)$ there exists $\mathcal{Z}_2 \in Res(s_2)$ such that:

$$prob(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; prob(\mathcal{CC}(z_{s_2}, \alpha))$$

and the same condition holds when exchanging $\mathcal{Z}_1$ with $\mathcal{Z}_2$. ■

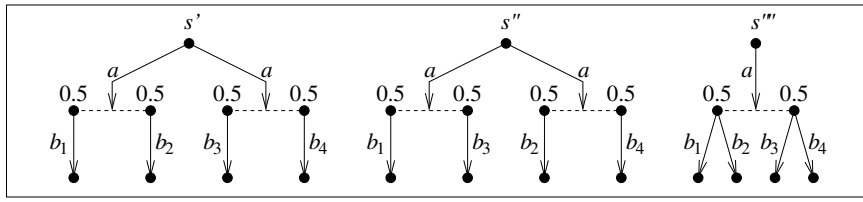

Figure 4.1: $\sim^{\text{post}}_{\text{PTr}}$ strictly finer than $\sim^{\text{pre}}_{\text{PTr}}$: $s' \not\sim^{\text{post}}_{\text{PTr}} s'' \not\sim^{\text{post}}_{\text{PTr}} s'''$, $s' \sim^{\text{pre}}_{\text{PTr}} s'' \sim^{\text{pre}}_{\text{PTr}} s'''$

Note that in Def. 4.2 we have that $z_{s_1} \sim^{\text{post}}_{\text{PTr}} z_{s_2}$ too, whilst in Def. 4.3 it is not necessarily the case that $z_{s_1} \sim^{\text{pre}}_{\text{PTr}} z_{s_2}$. Furthermore, the relation $\sim^{\text{post}}_{\text{PTr}}$ is trivially contained in $\sim^{\text{pre}}_{\text{PTr}}$.

The strictness of the inclusion, i.e., the difference between the two equivalences in terms of their discriminating power, is exemplified in Fig. 4.1. The initial states s', s'', s''' of the three NPLTS models are pairwise distinguished by $\sim^{\text{post}}_{\text{PTr}}$. For instance, the resolution of s' including the leftmost a-transition followed by the b_1-transition and the b_2-transition cannot be matched by any resolution of s'' and s''', because in no resolution of s'' and s''' traces $a\,b_1$ and $a\,b_2$ are both executable. Likewise, the resolution of s''' including the a-transition followed by the b_1-transition and the b_4-transition cannot be matched by any resolution of s'', because in no resolution of s'' traces $a\,b_1$ and $a\,b_4$ are both executable. On the other hand, the three initial states are identified by $\sim^{\text{pre}}_{\text{PTr}}$, because for all $i = 1, \ldots, 4$ the probability of executing trace $a\,b_i$ is the same in all the three NPLTS models.

4.3 Testing Semantics: $\sim_{\text{PTe-}\sqcup\sqcap}$

The testing theories developed in (Yi and Larsen, 1992; Jonsson and Yi, 1995; Segala, 1996) for nondeterministic and probabilistic processes are based on comparing the *extremal probabilities* of passing a test.

We formalize both processes and tests as NPLTS models, with the difference that a test has finitely many states and transitions, features an acyclic graph structure, and may contain occurrences of a success state that has no outgoing transitions.

Definition 4.4. A *nondeterministic and probabilistic test*, NPT for short, is an acyclic NPLTS $\mathcal{T} = (O, A, \longrightarrow)$ where both O and $\longrightarrow$ are finite, with O containing a distinguished success state denoted by ω having no outgoing transitions. We say that a computation of $\mathcal{T}$ is *successful* iff its last state is ω. ■

A test is passed by a process with a certain probability if there exists a resolution of nondeterminism of the parallel composition of the process and the test, with synchronization being enforced on any action, in which the probability of reaching a state having success in its test component is equal to the given probability.

Definition 4.5. Let $\mathcal{L} = (S, A, \longrightarrow_{\mathcal{L}})$ be an NPLTS and let $\mathcal{T} = (O, A, \longrightarrow_{\mathcal{T}})$ be an NPT. The *interaction system* of $\mathcal{L}$ and $\mathcal{T}$ is the NPLTS $\mathcal{I}(\mathcal{L}, \mathcal{T}) = (S \times O, A, \longrightarrow)$ where:

- Every state $(s, o) \in S \times O$ is called a *configuration*, which is *successful* iff $o = \omega$.
- $(s, o) \xrightarrow{a} \Delta$ iff $s \xrightarrow{a}_{\mathcal{L}} \Delta_1$ and $o \xrightarrow{a}_{\mathcal{T}} \Delta_2$ with $\Delta(s', o') = \Delta_1(s') \cdot \Delta_2(o')$ for all $(s', o') \in S \times O$.
- A computation of $\mathcal{I}(\mathcal{L}, \mathcal{T})$ is *successful* iff so is its last configuration.

■

We observe that $\mathcal{I}(\mathcal{L}, \mathcal{T})$ and any $\mathcal{Z} \in \mathit{Res}(s, o)$ have finitely many computations due to the structure of $\mathcal{T}$. We denote by $\mathcal{SC}(z_{s,o})$ the set of successful computations from the initial state $z_{s,o}$ of $\mathcal{Z}$.

Only *maximal resolutions* of nondeterminism, whose set is denoted by $\mathit{Res}_{\max}(\_)$, are taken into account within interaction systems, meaning that whenever $z \in Z$ has no outgoing transitions, then $\mathit{corr}_{\mathcal{Z}}(z)$ has no outgoing transitions either. The reason for this restriction is that resolutions that are not maximal do not expose all successful computations and hence may erroneously lead to conclude that the minimal success probability is zero.

As shown in Thm. 4.4 of (Bernardo *et al.*, 2014a), the discriminating power of the probabilistic testing equivalence $\sim_{\text{PTe-}\sqcup\sqcap}$ below, where $\sqcup$ and $\sqcap$ respectively denote the supremum and infimum of a set of

numbers, does not change if structure-modifying resolutions are used in place of structure-preserving ones.

Definition 4.6. Let $\mathcal{L} = (S, A, \longrightarrow_{\mathcal{L}})$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim_{\text{PTe-}\sqcup\sqcap} s_2$ iff for every NPT $\mathcal{T} = (O, A, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$ it holds that:

$$\bigsqcup_{\mathcal{Z}_1 \in Res_{\max}(s_1,o)} prob(\mathcal{SC}(z_{s_1,o})) = \bigsqcup_{\mathcal{Z}_2 \in Res_{\max}(s_2,o)} prob(\mathcal{SC}(z_{s_2,o}))$$

$$\bigsqcap_{\mathcal{Z}_1 \in Res_{\max}(s_1,o)} prob(\mathcal{SC}(z_{s_1,o})) = \bigsqcap_{\mathcal{Z}_2 \in Res_{\max}(s_2,o)} prob(\mathcal{SC}(z_{s_2,o}))$$

■

For instance, in the forthcoming Fig. 8.1 the NPLTS with initial state s_1 is not probabilistic testing equivalent to the NPLTS with initial state s_2. This is witnessed by the NPT with initial state o, which results in the two interaction systems whose initial states are (s_1, o) and (s_2, o) respectively. The former interaction system has two maximal resolutions whose initial states are $z'_{s_1,o}$ and $z''_{s_1,o}$ respectively, in which it holds that $\sqcup\{p, 1-p\} = p$ and $\sqcap\{p, 1-p\} = 1-p$ if we assume that $p \geq 1-p$. The latter interaction system has four maximal resolutions whose initial states are $z'_{s_2,o}$, $z''_{s_2,o}$, $z'''_{s_2,o}$, and $z''''_{s_2,o}$ respectively, in which it holds that $\sqcup\{p, 1, 0, 1-p\} = 1$ and $\sqcap\{p, 1, 0, 1-p\} = 0$ instead.

5

Anomalies of Probabilistic Trace Equivalences

Deterministic schedulers are very intuitive, but they cause the two probabilistic trace equivalences in Defs. 4.2 and 4.3 to be overdiscriminating, thereby violating desirable properties. On the one hand, this is due to the rigid preservation of the original model structure ensured by deterministic schedulers. On the other hand, as we will see shortly, it stems from the freedom of these schedulers of performing inconsistent choices in states with equivalent continuations. This also happens, to a much lesser extent though, with randomized and interpolating schedulers.

The resulting anomalies consist of $\sim_{\mathrm{PTr}}^{\mathrm{post}}$ and $\sim_{\mathrm{PTr}}^{\mathrm{pre}}$ *not* being:

- coarser than $\sim_{\mathrm{PB}}$ under deterministic schedulers;
- congruences w.r.t. action prefix under deterministic schedulers;
- compatible with their version for fully probabilistic processes.

The first anomaly is illustrated by the two NPLTS models in the leftmost part of Fig. 5.1. It holds that $s_1 \sim_{\mathrm{PB}} s_2$ – because, as shown in the example right after Def. 4.1, s_2' and s_2'' belong to the same bisimulation class thus causing their probabilities to be summed up – whereas $s_1 \not\sim_{\mathrm{PTr}}^{\mathrm{post}} s_2$ and $s_1 \not\sim_{\mathrm{PTr}}^{\mathrm{pre}} s_2$ – hence the two probabilistic trace

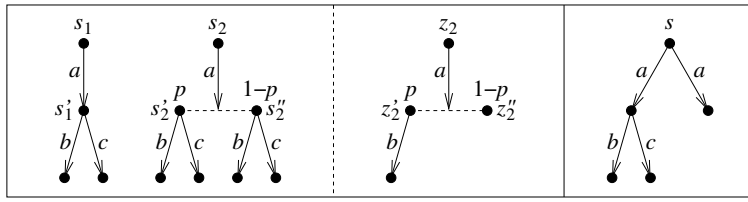

Figure 5.1: Violation of $s_1 \sim_{\mathrm{PB}} s_2 \Longrightarrow s_1 \sim_{\mathrm{PTr}} s_2$ (maximality does not help)

equivalences do not include probabilistic bisimilarity – because of the resolution whose initial state is z_2. This belongs to $Res_{\mathrm{sp}}(s_2) \setminus Res_{\mathrm{sp}}(s_1)$, as it does not preserve the structure of the NPLTS whose initial state is s_1, and cannot be matched by any resolution of that NPLTS, as trace $a\,b$ is executable from z_2 with probability p instead of 1. Notice that the same resolution belongs to $Res_{\mathrm{sm,r}}(s_1)$, if the a-transition of s_1 is combined with itself, as well as to $Res_{\mathrm{sm,i}}(s_1)$, if z_2' and z_2'' are both mapped to s_1'.

One may be tempted to admit only *maximal* resolutions in the definition of the two probabilistic trace equivalences. However, the problem would still be there if a c-transition departed from z_2'', i.e., if the scheduler selected b in s_2' and c in s_2'' in spite of the fact that s_2' and s_2'' are isomorphic. Moreover, by restricting to maximal resolutions, the probabilistic trace equivalences would no longer be compatible with trace equivalence. For instance, the former would not identify the two trace equivalent, fully nondeterministic NPLTS models in Fig. 5.1 whose initial states are s_1 and s, because the maximal resolution of s with an a-transition only – featuring traces ε and a – is not matched by the two maximal resolutions of s_1 – respectively featuring also $a\,b$ and $a\,c$.

The second anomaly is illustrated by the two NPLTS models in the leftmost part of Fig. 5.2. After the initial a-transitions, two distributions are reached that are probabilistic trace equivalent, in the sense that for each class of equivalent states both distributions assign the same probability to that class. However, it holds that $s_3 \not\sim^{\mathrm{post}}_{\mathrm{PTr}} s_4$ and $s_3 \not\sim^{\mathrm{pre}}_{\mathrm{PTr}} s_4$, hence the two probabilistic trace equivalences are not congruences with respect to action prefix, which is the operator that concatenates the execution of an action with a process distribution. The distinction is witnessed by the resolution whose initial state is z_3, which belongs to

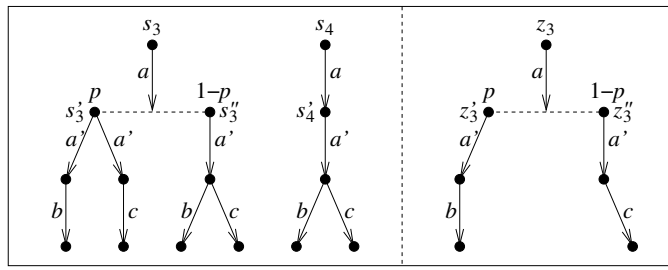

Figure 5.2: Violation of congruence with respect to action prefix: $s_3 \not\sim_{\mathrm{PTr}} s_4$

$Res_{\mathrm{sp}}(s_3) \setminus Res_{\mathrm{sp}}(s_4)$, as it does not preserve the structure of the NPLTS whose initial state is s_4, and cannot be matched by any resolution of that NPLTS, as trace $a\,a'\,b$ is executable from z_3 with probability p instead of 1. Notice that the same resolution belongs to $Res_{\mathrm{sm,r}}(s_4)$, if the a-transition of s_4 is combined with itself, and to $Res_{\mathrm{sm,i}}(s_4)$, if z_3' and z_3'' are both mapped to s_4'.

It is worth recalling that trace equivalence for fully nondeterministic processes is a congruence with respect to action prefix (Brookes *et al.*, 1984). The difference with the fully nondeterministic setting is that in our setting the continuation after an action is *not a single process*, but a probability distribution over processes. The problem arises when several equivalent states are in the support of the same distribution, as in the case of the target distribution of the a-transition of s_3 in Fig. 5.2, thereby allowing the scheduler to act inconsistently.

The third anomaly is illustrated by the two NPLTS models in the leftmost part of Fig. 5.3. They are identified by the trace equivalence for fully probabilistic processes of (Jou and Smolka, 1990), which does not use schedulers at all as in those processes there is no nondeterminism. However, it turns out that $s_5 \not\sim^{\mathrm{post}}_{\mathrm{PTr}} s_6$ and $s_5 \not\sim^{\mathrm{pre}}_{\mathrm{PTr}} s_6$ – hence the two probabilistic trace equivalences are not backward compatible with the one for fully probabilistic processes – because $\sim^{\mathrm{post}}_{\mathrm{PTr}}$ and $\sim^{\mathrm{pre}}_{\mathrm{PTr}}$ do make use of schedulers, and schedulers may decide of stopping the execution. This is witnessed by the resolution whose initial state is z_6, which belongs to $Res_{\mathrm{sp}}(s_6) \setminus Res_{\mathrm{sp}}(s_5)$, as it does not preserve the structure of the NPLTS whose initial state is s_5, and cannot be matched by any resolution of that NPLTS, as the scheduler has decided to stop at z_6''

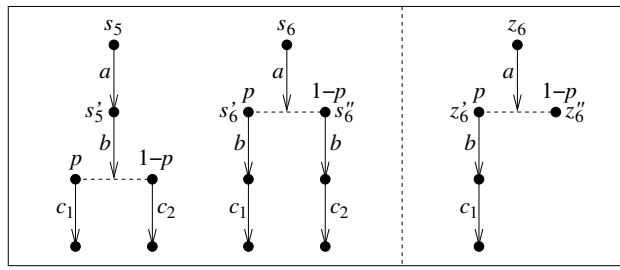

Figure 5.3: Lack of backward compatibility: $s_5 \not\sim_{\mathrm{PTr}} s_6$ (levelwise coherency)

so that trace $a\,b$ is executable from z_6 with probability p instead of 1.

The resolution does not even belong to $Res_{\mathrm{sm,r}}(s_5) \cup Res_{\mathrm{sm,i}}(s_5)$. After performing the a-transition and the b-transition, the c_1-transition in the NPLTS starting with s_5 can be executed with probability p, while the c_1-transition in the resolution can be executed with probability 1 and hence its source state cannot be mapped to the source state of the former c_1-transition.

This third example highlights that schedulers inducing structure-modifying resolutions are *not* exempt from shortcomings despite their greater flexibility with respect to deterministic schedulers. The considered resolution would be ruled out by imposing maximality but, as we have seen with the first example depicted in Fig. 5.1, this may generate other anomalies.

6

Anomaly Avoidance via Coherent Resolutions

The anomalies illustrated in Figs. 5.1 to 5.3 are mostly due to the freedom of schedulers of making different decisions in states enabling the same actions followed by the same continuations. We consequently propose to limit the excessive power of schedulers by restricting them to yield *coherent resolutions*. Intuitively, if several states in the support of the target distribution of a transition are equivalent, then the decisions made by the scheduler in those states have to be coherent with each other, so that the states to which they correspond in any resolution are equivalent as well.

Our proposal will be implemented in two steps. First, we introduce the notion of coherency for trace distributions by means of suitable operations on them relying on the intuition above (Sect. 6.1) and we reason on how to set up coherency constraints based on the counterexamples in Figs. 5.1 to 5.3 as well as two further counterexamples (Sect. 6.2). Then, we strengthen the previous construction by ensuring that coherent decisions are not forgotten when extending coherent trace distributions to longer ones (Sect. 6.3). Finally, we formalize the coherency constraints yielding coherent resolutions and show that they avoid the examined anomalies of $\sim_{\mathrm{PTr}}^{\mathrm{post}}$ and $\sim_{\mathrm{PTr}}^{\mathrm{pre}}$ (Sect. 6.4).

6.1 Coherent Trace Distributions

The *trace distribution* of a state is a family of sets of traces weighted with their execution probabilities in a given resolution of that state, with each such set containing all the weighted traces up to a certain length in a different resolution. The coherency constraints implementing our proposal will be expressed by reasoning on *coherent* trace distributions built through the following operations, where TD denotes a trace distribution whilst T denotes a weighted trace set.

Definition 6.1. Let $A \neq \emptyset$ be a countable set. For $a \in A$, $p \in \mathbb{R}$, $TD \subseteq 2^{A^* \times \mathbb{R}}$, and $T \subseteq A^* \times \mathbb{R}$ we define:

$$\begin{aligned} a \,.\, TD &= \{a \,.\, T \mid T \in TD\} \\ a \,.\, T &= \{(a\,\alpha, p') \mid (\alpha, p') \in T\} \\ p \cdot TD &= \{p \cdot T \mid T \in TD\} \\ p \cdot T &= \{(\alpha, p \cdot p') \mid (\alpha, p') \in T\} \\ tr(TD) &= \{tr(T) \mid T \in TD\} \\ tr(T) &= \{\alpha \in A^* \mid \exists p' \in \mathbb{R}.\, (\alpha, p') \in T\} \end{aligned}$$

while for $TD_1, TD_2 \subseteq 2^{A^* \times \mathbb{R}}$ we define:

$$TD_1 + TD_2 = \begin{cases} \{T_1 + T_2 \mid T_1 \in TD_1 \wedge T_2 \in TD_2 \wedge tr(T_1) = tr(T_2)\} \\ \qquad\qquad \text{if } tr(TD_1) = tr(TD_2) \\ \{T_1 + T_2 \mid T_1 \in TD_1 \wedge T_2 \in TD_2\} \\ \qquad\qquad \text{otherwise} \end{cases}$$

where for $T_1, T_2 \subseteq A^* \times \mathbb{R}$:

$$\begin{aligned} T_1 + T_2 = {} & \{(\alpha, p_1 + p_2) \mid (\alpha, p_1) \in T_1 \wedge (\alpha, p_2) \in T_2\} \cup \\ & \{(\alpha, p) \in T_1 \cup T_2 \mid \alpha \notin tr(T_1) \cap tr(T_2)\} \end{aligned}$$

■

The coherent addition $T_1 + T_2$ of weighted trace sets is commutative and associative, with probabilities of identical traces in the two summands being always added up for coherency purposes. In contrast, the coherent addition $TD_1 + TD_2$ of trace distributions is only commutative. Essentially, the two summands represent two families of sets of weighted traces executable in two resolutions of two states in the support of a target distribution. Every weighted trace set $T_1 \in TD_1$ is summed with every weighted trace set $T_2 \in TD_2$ – so as to characterize an overall resolution – unless TD_1 and TD_2 have the same family of trace sets (regardless of their weights), in which case summation is restricted to weighted trace sets featuring the same traces for the sake of coherency.

Due to the lack of associativity, in the definition below all trace distributions $\Delta(s') \cdot TD^{\mathrm{c}}_{n-1}(s')$ exhibiting the same family Θ of trace sets have to be summed up first, which is ensured by the presence of a double summation.

Definition 6.2. Let $(S, A, \longrightarrow)$ be an NPLTS and $s \in S$. The *coherent trace distribution* of s is the subset of $2^{A^* \times \mathbb{R}_{]0,1]}}$ defined as follows:

$$TD^{\mathrm{c}}(s) \;=\; \bigcup_{n \in \mathbb{N}} TD^{\mathrm{c}}_n(s)$$

with $TD^{\mathrm{c}}_n(s)$, the coherent trace distribution of s whose traces have length at most n, being defined as:

$$\begin{cases} (\varepsilon,1) \dagger \bigcup\limits_{s \xrightarrow{a} \Delta} a \,.\, \left(\sum\limits_{\Theta \in tr(\Delta, n-1)} \; \sum\limits_{s' \in supp(\Delta)}^{tr(TD^{\mathrm{c}}_{n-1}(s')) = \Theta} \Delta(s') \cdot TD^{\mathrm{c}}_{n-1}(s') \right) \\ \qquad\qquad \text{if } n > 0 \text{ and } s \text{ has outgoing transitions} \\ \{\{(\varepsilon,1)\}\} \\ \qquad\qquad \text{otherwise} \end{cases}$$

where $tr(\Delta, n-1) = \{tr(TD^{\mathrm{c}}_{n-1}(s')) \mid s' \in supp(\Delta)\}$ and the operator $(\varepsilon,1) \dagger \_$ is such that $(\varepsilon,1) \dagger TD = \{\{(\varepsilon,1)\} \cup T \mid T \in TD\}$. ■

For example, in Fig. 5.1 we have that $TD^{\mathrm{c}}_n(s_1) = TD^{\mathrm{c}}_n(s_2)$ for all $n \in \mathbb{N}$, where due to Defs. 6.2 and 6.1:

- $TD^{\mathrm{c}}_0(s_2) = \{\{(\varepsilon,1)\}\}$
- $\begin{aligned} TD^{\mathrm{c}}_1(s_2) &= (\varepsilon,1) \dagger a \,.\, (p \cdot TD^{\mathrm{c}}_0(s_2') + (1-p) \cdot TD^{\mathrm{c}}_0(s_2'')) \\ &= (\varepsilon,1) \dagger a \,.\, (p \cdot \{\{(\varepsilon,1)\}\} + (1-p) \cdot \{\{(\varepsilon,1)\}\}) \\ &= (\varepsilon,1) \dagger a \,.\, (\{\{(\varepsilon,p)\}\} + \{\{(\varepsilon,1-p)\}\}) \\ &= (\varepsilon,1) \dagger a \,.\, \{\{(\varepsilon,1)\}\} \\ &= (\varepsilon,1) \dagger \{\{(a,1)\}\} \\ &= \{\{(\varepsilon,1),(a,1)\}\} \end{aligned}$
- $\begin{aligned} TD^{\mathrm{c}}_2(s_2) &= (\varepsilon,1) \dagger a \,.\, (p \cdot TD^{\mathrm{c}}_1(s_2') + (1-p) \cdot TD^{\mathrm{c}}_1(s_2'')) \\ &= (\varepsilon,1) \dagger a \,.\, (p \cdot \{\{(\varepsilon,1),(b,1)\},\{(\varepsilon,1),(c,1)\}\} + \\ &\qquad\qquad (1-p) \cdot \{\{(\varepsilon,1),(b,1)\},\{(\varepsilon,1),(c,1)\}\}) \\ &= (\varepsilon,1) \dagger a \,.\, (\{\{(\varepsilon,p),(b,p)\},\{(\varepsilon,p),(c,p)\}\} + \\ &\qquad\qquad \{\{(\varepsilon,1-p),(b,1-p)\},\{(\varepsilon,1-p),(c,1-p)\}\}) \\ &= (\varepsilon,1) \dagger a \,.\, \{\{(\varepsilon,1),(b,1)\},\{(\varepsilon,1),(c,1)\}\} \\ &= (\varepsilon,1) \dagger \{\{(a,1),(a\,b,1)\},\{(a,1),(a\,c,1)\}\} \\ &= \{\{(\varepsilon,1),(a,1),(a\,b,1)\},\{(\varepsilon,1),(a,1),(a\,c,1)\}\} \end{aligned}$

so that $TD^c(s_1) = TD^c(s_2) = TD^c_0(s_2) \cup TD^c_1(s_2) \cup TD^c_2(s_2) = \{\{(\varepsilon,1)\},\{(\varepsilon,1),(a,1)\},\{(\varepsilon,1),(a,1),(a\,b,1)\},\{(\varepsilon,1),(a,1),(a\,c,1)\}\}$.

Let us investigate the properties of the construction in Def. 6.2. First of all, in the absence of nondeterminism, like in the case of a fully probabilistic NPLTS, any coherent trace distribution $TD^c_n(s)$ contains a single weighted trace set. This holds in particular for resolutions.

Proposition 6.1. Let $(S, A, \longrightarrow)$ be a fully probabilistic NPLTS, $s \in S$, and $n \in \mathbb{N}$. Let $A^{\leq n} = \{\alpha \in A^* \mid |\alpha| \leq n\}$. Then $TD^c_n(s) = \{\{(\alpha,p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s,\alpha)) = p\}\}$.

Proof. We proceed by induction on $n \in \mathbb{N}$:

- For $n = 0$ we have $TD^c_n(s) = \{\{(\varepsilon,1)\}\} = \{\{(\alpha,p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s,\alpha)) = p\}\}$.
- Let $n = m+1$ for some $m \in \mathbb{N}$ and suppose that the result holds for the coherent trace distribution $TD^c_m(s')$ of any state $s' \in S$. If s has no outgoing transitions, we proceed like in the case $n = 0$. If s has an outgoing transition $s \xrightarrow{a} \Delta$, then this must be unique as the considered NPLTS is fully probabilistic. Therefore:
$$TD^c_n(s) = (\varepsilon,1) \dagger a \,.\, \left(\sum_{\Theta \in tr(\Delta,m)} \sum_{s' \in supp(\Delta)}^{tr(TD^c_m(s'))=\Theta} \Delta(s') \cdot TD^c_m(s') \right)$$
From the induction hypothesis, for each $s' \in supp(\Delta)$ it follows that:
$$TD^c_m(s') = \{\{(\alpha',p') \in A^{\leq m} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s',\alpha')) = p'\}\}$$
and hence:
$$\begin{aligned} TD^c_n(s) &= (\varepsilon,1) \dagger \sum_{s' \in supp(\Delta)} \{\{(a\,\alpha', \Delta(s') \cdot p') \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid \\ &\qquad\qquad prob(\mathcal{CC}(s',\alpha')) = p'\}\} \\ &= \{\{(\alpha,p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s,\alpha)) = p\}\} \end{aligned}$$
where the summation indexed by Θ has disappeared after applying the induction hypothesis because when summing up singleton trace distributions there is no difference, according to Def. 6.1, between the case in which they share the same family of trace sets and the case in which they do not. □

As for the relationship between $TD^c_n(s)$ and $TD^c_{n-1}(s)$, it turns out that every element of the former contains, among its traces, those of

an element of the latter. As we will see in Sect. 6.3, the probabilities of common traces may differ in the two sets.

Proposition 6.2. Let $(S, A, \longrightarrow)$ be an NPLTS, $s \in S$, and $n \in \mathbb{N}_{\geq 1}$. Then for all $T \in TD^{c}_{n}(s)$ there exists $T' \in TD^{c}_{n-1}(s)$ such that $tr(T') \subseteq tr(T)$.

Proof. If s has no outgoing transitions, then $TD^{c}_{n}(s) = TD^{c}_{n-1}(s) = \{\{(\varepsilon, 1)\}\}$ by Def. 6.2 and hence the result trivially holds, otherwise we proceed by induction on $n \in \mathbb{N}_{\geq 1}$:

- For $n = 1$ we have that $TD^{c}_{n}(s) = \{\{(\varepsilon, 1), (a, 1)\} \mid s \xrightarrow{a} \Delta\}$, with each of its elements T including as a subset the only element $T' = \{(\varepsilon, 1)\}$ of $TD^{c}_{n-1}(s)$.

- Let $n = m + 1$ for some $m \in \mathbb{N}_{\geq 1}$ and suppose that the result holds for the coherent trace distributions $TD^{c}_{m}(s')$ and $TD^{c}_{m-1}(s')$ of any state $s' \in S$. Consider an arbitrary element T of $TD^{c}_{n}(s)$ originated from some transition departing from s, say $s \xrightarrow{a} \Delta$. Then by virtue of Def. 6.2:
$$T \;\in\; (\varepsilon, 1) \dagger a \,. \left(\sum_{\Theta \in tr(\Delta, m)} \; \sum_{s' \in supp(\Delta)}^{tr(\,TD^{c}_{m}(s')) = \Theta} \Delta(s') \cdot TD^{c}_{m}(s') \right)$$
Since T is obtained by summing up a suitable element $T_{s'}$ of $TD^{c}_{m}(s')$ for every $s' \in supp(\Delta)$, we have that:
$$tr(T) \;=\; \{\varepsilon\} \cup a \,. \bigcup_{s' \in supp(\Delta)} tr(T_{s'}) \text{ [maximum trace length is } n\text{]}$$
From the induction hypothesis, for each such $T_{s'} \in TD^{c}_{m}(s')$ there exists $T'_{s'} \in TD^{c}_{m-1}(s')$ such that $tr(T'_{s'}) \subseteq tr(T_{s'})$. Using these sets $T'_{s'} \in TD^{c}_{m-1}(s')$ in the first formula above deriving from Def. 6.2, we assemble a set $T' \in TD^{c}_{m}(s)$, originated from the same aforementioned transition $s \xrightarrow{a} \Delta$, such that:
$$tr(T') \;=\; \{\varepsilon\} \cup a \,. \bigcup_{s' \in supp(\Delta)} tr(T'_{s'}) \text{ [maximum trace length is } m\text{]}$$
which thus satisfies $tr(T') \subseteq tr(T)$. □

6.2 Towards Coherency Constraints

Let us reconsider the three counterexamples in Figs. 5.1 to 5.3 by examining the coherent trace distributions for states (of models or resolutions) in the support of the target distribution of some transitions:

- In Fig. 5.1 it holds that in the NPLTS with initial state s_2:
  $TD^c(s_2') = \{\{(\varepsilon,1)\},\{(\varepsilon,1),(b,1)\},\{(\varepsilon,1),(c,1)\}\} = TD^c(s_2'')$
  whilst in the resolution with initial state z_2:
  $TD^c(z_2') = \{\{(\varepsilon,1)\},\{(\varepsilon,1),(b,1)\}\} \neq \{\{(\varepsilon,1)\}\} = TD^c(z_2'')$
  In other words, s_2' and s_2'' have the same coherent trace distribution, but the states to which they correspond in the resolution, i.e., z_2' and z_2'', have not.

- In Fig. 5.2 it holds that in the NPLTS with initial state s_3:
  $TD^c(s_3') = \{\{(\varepsilon,1)\},\{(\varepsilon,1),(a',1)\},\{(\varepsilon,1),(a',1),(a'\,b,1)\},$
  $\{(\varepsilon,1),(a',1),(a'\,c,1)\}\} = TD^c(s_3'')$
  whereas in the resolution with initial state z_3:
  $TD^c(z_3') = \{\{(\varepsilon,1)\},\{(\varepsilon,1),(a',1)\},\{(\varepsilon,1),(a',1),(a'\,b,1)\}\} \neq$
  $\{\{(\varepsilon,1)\},\{(\varepsilon,1),(a',1)\},\{(\varepsilon,1),(a',1),(a'\,c,1)\}\} = TD^c(z_3'')$
  Again, the relationships between coherent trace distributions of states in the model and states to which they correspond in the resolution reveal inconsistent choices made by the scheduler at resolution construction time.

- In Fig. 5.3 it holds that in the NPLTS with initial state s_6:
  $TD^c(s_6') = \{\{(\varepsilon,1)\},\{(\varepsilon,1),(b,1)\},\{(\varepsilon,1),(b,1),(b\,c_1,1)\}\} \neq$
  $\{\{(\varepsilon,1)\},\{(\varepsilon,1),(b,1)\},\{(\varepsilon,1),(b,1),(b\,c_2,1)\}\} = TD^c(s_6'')$
  However, like s_2' and s_2'' as well as s_3' and s_3'', also s_6' and s_6'' enable the same action set, which is $\{b\}$. Indeed, if we limit the length of the considered traces to 1, s_6' and s_6'' turn out to have the same trace distribution because:
  $$TD_1^c(s_6') = \{\{(\varepsilon,1),(b,1)\}\} = TD_1^c(s_6'')$$
  but in the resolution with initial state z_6:
  $$TD_1^c(z_6') = \{\{(\varepsilon,1),(b,1)\}\} \neq \{\{(\varepsilon,1)\}\} = TD_1^c(z_6'')$$
  This shows that, instead of a single constraint based on TD^c sets, we should set up *separate coherency constraints* relying on TD_n^c sets for every $n \in \mathbb{N}$.

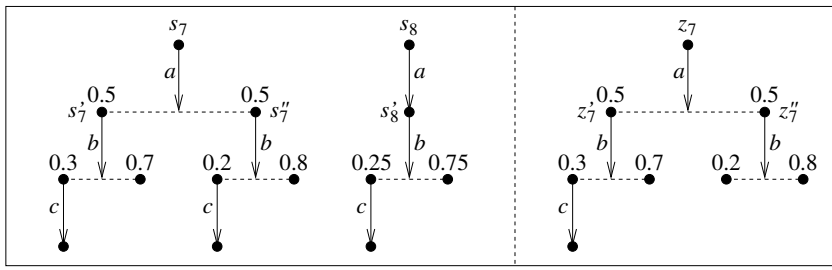

Figure 6.1: Lack of backward compatibility: $s_7 \not\sim_{\mathrm{PTr}} s_8$ (probability abstraction)

An additional counterexample, depicted in Fig. 6.1, indicates that the coherency constraints should be based on TD_n^{c} sets *up to the probabilities* they contain, i.e., the constraints should rely on $tr(\,TD_n^{\mathrm{c}})$ sets.

The two fully probabilistic NPLTS models on the left are identified by the trace equivalence for fully probabilistic processes of (Jou and Smolka, 1990) – and indeed $TD_3^{\mathrm{c}}(s_7) = \{\{(\varepsilon,1),(a,1),(a\,b,1),(a\,b\,c,0.25)\}\} = TD_3^{\mathrm{c}}(s_8)$ – but $s_7 \not\sim_{\mathrm{PTr}}^{\mathrm{post}} s_8$ and $s_7 \not\sim_{\mathrm{PTr}}^{\mathrm{pre}} s_8$ as witnessed by the resolution whose initial state is z_7. This resolution belongs to $Res_{\mathrm{sp}}(s_7) \setminus Res_{\mathrm{sp}}(s_8)$, as it does not preserve the structure of the NPLTS whose initial state is s_8, and cannot be matched by any resolution of that NPLTS. It does not even belong to $Res_{\mathrm{sm,r}}(s_8) \cup Res_{\mathrm{sm,i}}(s_8)$ because, after performing the a-transition and the b-transition, the c-transition in the NPLTS starting with s_8 can be executed with probability 0.25, while the c-transition in the resolution can be executed with probability 0.3 and hence its source state cannot be mapped to the source state of the former c-transition. It holds that $TD_2^{\mathrm{c}}(s_7') = \{\{(\varepsilon,1),(b,1),(b\,c,0.3)\}\} \neq \{\{(\varepsilon,1),(b,1),(b\,c,0.2)\}\} = TD_2^{\mathrm{c}}(s_7'')$. However, s_7' and s_7'' enable the same action set, which is $\{b\}$. Moreover, if we restrict ourselves to traces without their weights, s_7' and s_7'' turn out to have the same trace set:

$$tr(\,TD_2^{\mathrm{c}}(s_7')) \;=\; \{\{\varepsilon, b, b\,c\}\} \;=\; tr(\,TD_2^{\mathrm{c}}(s_7''))$$

but in the resolution:

$$tr(\,TD_2^{\mathrm{c}}(z_7')) \;=\; \{\{\varepsilon, b, b\,c\}\} \;\neq\; \{\{\varepsilon, b\}\} \;=\; tr(\,TD_2^{\mathrm{c}}(z_7''))$$

The violations, depicted in Figs. 5.3 and 6.1, of backward compatibility with respect to the trace equivalence of (Jou and Smolka, 1990) share a common characteristic about the choices made in states having the same traces of a certain length. The lack of coherency is

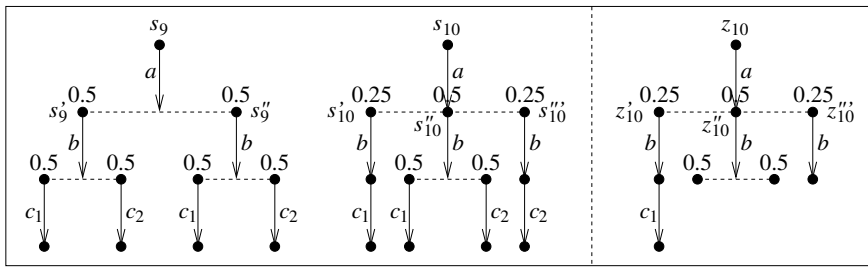

Figure 6.2: Lack of backward compatibility: $s_9 \not\sim_{\mathrm{PTr}} s_{10}$ (levelwise completeness)

a consequence of the fact that, in both resolutions of those figures, at a certain point the scheduler selects a transition along one direction while it stops the execution along the other direction. This is even more evident with the two fully probabilistic NPLTS models in the leftmost part of Fig. 6.2, which are identified by (Jou and Smolka, 1990) but told apart by $\sim^{\mathrm{post}}_{\mathrm{PTr}}$ and $\sim^{\mathrm{pre}}_{\mathrm{PTr}}$ due to the resolution whose initial state is z_{10}. States s'_{10}, s''_{10}, s'''_{10} enable the same action set, which is $\{b\}$, but $tr(TD^{\mathrm{c}}_2(s'_{10}))$, $tr(TD^{\mathrm{c}}_2(s''_{10}))$, $tr(TD^{\mathrm{c}}_2(s'''_{10}))$ are pairwise different, hence we cannot proceed like in the previous cases to detect inconsistent choices. However, we observe that trace $a\,b\,c_1$ is executable with probability 0.25 in the resolution, whereas in every resolution of s_9 without inconsistent choices it can be executed only with probability 0 or 0.5.

This further counterexample calls for the presence in each resolution of all the computations of length n if any, for every $n \in \mathbb{N}$, and of all possible shorter maximal computations. Note that *trace completeness up to length n* as a coherency constraint is an obligation looser than resolution maximality. Moreover, it can be easily formalized, as each set in the trace distribution of the original model contains all the weighted traces up to a certain length.

6.3 Making Coherent Trace Distributions Memoryful

The construction in Def. 6.2 alone is not enough because coherent decisions made in the past may be forgotten when extending coherent trace distributions to longer trace distributions due to the presence of longer traces that differ in their last action.

For example, consider the leftmost NPLTS in Fig. 6.3. We have:

$$TD_1^c(r_1) \;=\; \{\{(\varepsilon,1),(b,1)\}\} \;=\; TD_1^c(r_2)$$

and also:

$$TD_2^c(r_1) \;=\; \{\{(\varepsilon,1),(b,1),(b\,c,1)\},\{(\varepsilon,1),(b,1),(b\,d,1)\}\} \;=\; TD_2^c(r_2)$$

because in the complete submodel rooted at r_1 it holds that:

$$TD_1^c(r_1') \;=\; \{\{(\varepsilon,1),(c,1)\},\{(\varepsilon,1),(d,1)\}\} \;=\; TD_1^c(r_1'')$$

and hence, when applying Def. 6.2 to compute $TD_2^c(r_1)$, according to Def. 6.1 the summation is restricted to weighted trace sets featuring the same traces as:

$$tr(\,TD_1^c(r_1')) \;=\; \{\{\varepsilon,c\},\{\varepsilon,d\}\} \;=\; tr(\,TD_1^c(r_1''))$$

Nevertheless, when considering traces of length 3, which differ in their last action, since:

$$\begin{aligned} TD_2^c(r_1') \;&=\; \{\{(\varepsilon,1),(c,1),(c\,e_1,1)\},\{(\varepsilon,1),(d,1),(d\,e_2,1)\}\} \\ TD_2^c(r_1'') \;&=\; \{\{(\varepsilon,1),(c,1),(c\,e_3,1)\},\{(\varepsilon,1),(d,1),(d\,e_4,1)\}\} \end{aligned}$$

with:

$$\begin{aligned} tr(\,TD_2^c(r_1')) \;&=\; \{\{\varepsilon,c,c\,e_1\},\{\varepsilon,d,d\,e_2\}\} \;\neq \\ &\neq\; \{\{\varepsilon,c,c\,e_3\},\{\varepsilon,d,d\,e_4\}\} \;=\; tr(\,TD_2^c(r_1'')) \end{aligned}$$

we subsequently derive that:

$$\begin{aligned} TD_3^c(r_1) \;=\;& (\varepsilon,1)\,\dagger \\ & (\{\{(b,p),(b\,c,p),(b\,c\,e_1,p)\}, \\ & \quad \{(b,p),(b\,d,p),(b\,d\,e_2,p)\}\} + \\ & \{\{(b,1-p),(b\,c,1-p),(b\,c\,e_3,1-p)\}, \\ & \quad \{(b,1-p),(b\,d,1-p),(b\,d\,e_4,1-p)\}\}) \\ =\;& \{\{(\varepsilon,1),(b,1),(b\,c,1),(b\,c\,e_1,p),(b\,c\,e_3,1-p)\}, \\ & \{(\varepsilon,1),(b,1),\underline{(b\,c,p)},\underline{(b\,d,1-p)},(b\,c\,e_1,p),(b\,d\,e_4,1-p)\}, \\ & \{(\varepsilon,1),(b,1),\underline{(b\,d,p)},\underline{(b\,c,1-p)},(b\,d\,e_2,p),(b\,c\,e_3,1-p)\}, \\ & \{(\varepsilon,1),(b,1),(b\,d,1),(b\,d\,e_2,p),(b\,d\,e_4,1-p)\}\} \end{aligned}$$

whereas:

$$\begin{aligned} TD_3^c(r_2) \;=\;& \{\{(\varepsilon,1),(b,1),(b\,c,1),(b\,c\,e_1,p),(b\,c\,e_3,1-p)\}, \\ & \{(\varepsilon,1),(b,1),(b\,d,1),(b\,d\,e_2,p),(b\,d\,e_4,1-p)\}\} \end{aligned}$$

Therefore, in the calculation of $TD_4^c(r)$ we cannot simply sum up weighted trace sets in $TD_3^c(r_1)$ and weighted trace sets in $TD_3^c(r_2)$ that exhibit the same traces. This is due to the presence in $TD_3^c(r_1)$ of the following two weighted trace sets:

$$\begin{aligned} &\{(\varepsilon,1),(b,1),\underline{(b\,c,p)},\underline{(b\,d,1-p)},(b\,c\,e_1,p),(b\,d\,e_4,1-p)\} \\ &\{(\varepsilon,1),(b,1),\underline{(b\,d,p)},\underline{(b\,c,1-p)},(b\,d\,e_2,p),(b\,c\,e_3,1-p)\} \end{aligned}$$

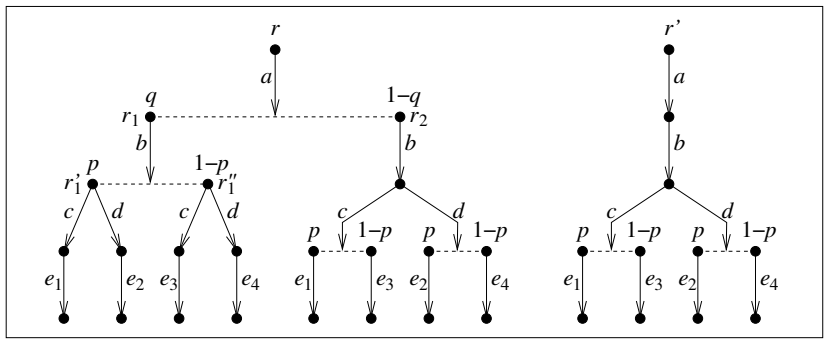

Figure 6.3: Memoryful coherency is necessary to reconcile $TD^{\mathrm{c}}_3(r_1)$ and $TD^{\mathrm{c}}_3(r_2)$

which cannot be exposed by any coherent resolution. The key observation is that coherency constraints avoiding a mix of shorter traces like $b\,c$ and $b\,d$ have been ignored by those two weighted trace sets in $TD^{\mathrm{c}}_3(r_1)$, hence they cannot be *extensions* of weighted trace sets in $TD^{\mathrm{c}}_2(r_1)$. Indeed, neither of those weighted trace sets in $TD^{\mathrm{c}}_3(r_1)$ includes as a subset a weighted trace set in $TD^{\mathrm{c}}_2(r_1)$ because of the different probabilities of traces $b\,c$ and $b\,d$ in the considered sets (see the sentence before Prop. 6.2). If we now consider the rightmost NPLTS in Fig. 6.3, then from $TD^{\mathrm{c}}_3(r_1) \neq TD^{\mathrm{c}}_3(r_2)$ it also follows that r' would be distinguished from r instead of being identified with it.

This example reveals that the construction of Def. 6.2, together with the coherent additions of weighted trace sets and of trace distributions in Def. 6.1, is not enough to set up the coherency constraints. What is missing is that every set $TD^{\mathrm{c}}_n(s)$, with $n > 0$ and s having outgoing transitions, should *incrementally* build on $TD^{\mathrm{c}}_{n-1}(s)$, in the sense that every weighted trace set in the former should include as a subset a weighted trace set in the latter. This is a *monotonicity*-like property stronger than the one of Prop. 6.2 – as probabilities are now included – which causes longer trace distributions to remember coherent decisions made in the past within shorter trace distributions. We thus introduce a variant of coherent trace distribution, which we call *memoryfully* coherent trace distribution.

Definition 6.3. Let $(S, A, \longrightarrow)$ be an NPLTS and $s \in S$. The *memoryfully coherent trace distribution* of s is the subset of $2^{A^* \times \mathbb{R}_{]0,1]}}$ defined as follows:

$$TD^{\mathrm{mc}}(s) \;=\; \bigcup_{n \in \mathbb{N}} TD^{\mathrm{mc}}_n(s)$$

with $TD^{\mathrm{mc}}_n(s)$, the memoryfully coherent trace distribution of s whose traces have length at most n, being the subset of $TD^{\mathrm{c}}_n(s)$ defined as:

$$\begin{cases} \{T \in TD^{\mathrm{c}}_n(s) \mid \exists T' \in TD^{\mathrm{mc}}_{n-1}(s).\, T' \subseteq T\} & \text{if } n > 0 \text{ and } s \text{ has outgoing transitions} \\ \{\{(\varepsilon, 1)\}\} & \text{otherwise} \end{cases}$$

■

For the leftmost NPLTS in Fig. 6.3 we have $TD^{\mathrm{mc}}_n(r_2) = TD^{\mathrm{c}}_n(r_2)$ for all $n \in \mathbb{N}$ and $TD^{\mathrm{mc}}_n(r_1) = TD^{\mathrm{c}}_n(r_1)$ for $n \leq 2$, while $TD^{\mathrm{mc}}_3(r_1) \neq TD^{\mathrm{c}}_3(r_1)$ because the following two weighted trace sets of $TD^{\mathrm{c}}_3(r_1)$ do not include any weighted trace set of $TD^{\mathrm{mc}}_2(r_1)$ and hence cannot be part of $TD^{\mathrm{mc}}_3(r_1)$:

$$\{(\varepsilon, 1), (b, 1), \underline{(b\,c, p)}, \underline{(b\,d, 1-p)}, (b\,c\,e_1, p), (b\,d\,e_4, 1-p)\}$$
$$\{(\varepsilon, 1), (b, 1), \underline{(b\,d, p)}, \underline{(b\,c, 1-p)}, (b\,d\,e_2, p), (b\,c\,e_3, 1-p)\}$$

It holds that $TD^{\mathrm{mc}}_3(r_1) = TD^{\mathrm{mc}}_3(r_2) = TD^{\mathrm{c}}_3(r_2)$ so that overall $TD^{\mathrm{mc}}(r) = TD^{\mathrm{mc}}_0(r) \cup TD^{\mathrm{mc}}_1(r) \cup TD^{\mathrm{mc}}_2(r) \cup TD^{\mathrm{mc}}_3(r) \cup TD^{\mathrm{mc}}_4(r)$ where:

$$\begin{aligned}
TD^{\mathrm{mc}}_0(r) &= \{\{(\varepsilon, 1)\}\} \\
TD^{\mathrm{mc}}_1(r) &= \{\{(\varepsilon, 1), (a, 1)\}\} \\
TD^{\mathrm{mc}}_2(r) &= \{\{(\varepsilon, 1), (a, 1), (a\,b, 1)\}\} \\
TD^{\mathrm{mc}}_3(r) &= \{\{(\varepsilon, 1), (a, 1), (a\,b, 1), (a\,b\,c, 1)\}, \\
&\qquad \{(\varepsilon, 1), (a, 1), (a\,b, 1), (a\,b\,d, 1)\}\} \\
TD^{\mathrm{mc}}_4(r) &= \{\{(\varepsilon, 1), (a, 1), (a\,b, 1), (a\,b\,c, 1), \\
&\qquad\qquad (a\,b\,c\,e_1, p), (a\,b\,c\,e_3, 1-p)\}, \\
&\qquad \{(\varepsilon, 1), (a, 1), (a\,b, 1), (a\,b\,d, 1), \\
&\qquad\qquad (a\,b\,d\,e_2, p), (a\,b\,d\,e_4, 1-p)\}\}
\end{aligned}$$

with the various sets $TD^{\mathrm{mc}}_n(r)$ precisely capturing the trace distributions of the coherent resolutions of r and $TD^{\mathrm{mc}}(r) = TD^{\mathrm{mc}}(r')$.

Let us investigate the properties of the construction in Def. 6.3. Memoryfully coherent trace distributions $TD^{\mathrm{mc}}_n(s)$ coincide with the corresponding coherent ones $TD^{\mathrm{c}}_n(s)$ when $n \leq 2$ as a consequence of Def. 6.1. The example in Fig. 6.3 shows that, when $n \geq 3$, in general $TD^{\mathrm{mc}}_n(s)$ cannot be recursively characterized in a direct manner as

$TD^{\mathrm{c}}_n(s)$ in Def. 6.2, even though each element of a memoryfully coherent trace distribution can be expressed as a sum of elements of other memoryfully coherent trace distributions.

Proposition 6.3. Let $(S, A, \longrightarrow)$ be an NPLTS, $s \in S$, and $n \in \mathbb{N}$. If $n \leq 2$ or s has no outgoing transitions, then $TD^{\mathrm{mc}}_n(s) = TD^{\mathrm{c}}_n(s)$, otherwise each element of $TD^{\mathrm{mc}}_n(s)$ is obtained by summing up a suitable element of $TD^{\mathrm{mc}}_{n-1}(s')$ for every s' in the support of the target distribution of a transition of s.

Proof. We proceed by case analysis:

- If $n = 0$ or s has no outgoing transitions, then $TD^{\mathrm{mc}}_n(s) = \{\{(\varepsilon, 1)\}\} = TD^{\mathrm{c}}_n(s)$. Henceforth we suppose that s has outgoing transitions.

- If $n = 1$ then $TD^{\mathrm{mc}}_n(s) = TD^{\mathrm{c}}_n(s)$ because $TD^{\mathrm{mc}}_{n-1}(s) = \{\{(\varepsilon, 1)\}\}$ and every $T \in TD^{\mathrm{c}}_n(s)$ is of the form $\{(\varepsilon, 1), (a, 1)\}$ for some action a labeling an outgoing transition of s, thus satisfying $\{(\varepsilon, 1)\} \subseteq T$.

- If $n = 2$ then $TD^{\mathrm{mc}}_n(s) = TD^{\mathrm{c}}_n(s)$ as $TD^{\mathrm{mc}}_{n-1}(s) = \{\{(\varepsilon, 1), (a, 1)\} \mid s \xrightarrow{a} \Delta\}$ and for every transition $s \xrightarrow{a} \Delta$ it holds that $TD^{\mathrm{c}}_n(s)$ includes as a subset:
$$(\varepsilon, 1) \dagger a \,.\, \left(\sum_{\Theta \in tr(\Delta, n-1)} \sum_{s' \in supp(\Delta)}^{tr(TD^{\mathrm{c}}_{n-1}(s')) = \Theta} \Delta(s') \cdot TD^{\mathrm{c}}_{n-1}(s') \right)$$
each element T of which certainly satisfies $\{(\varepsilon, 1), (a, 1)\} \subseteq T$. Note that the presence of $(a, 1)$ in T stems from the summation of all pairs $(\varepsilon, \Delta(s'))$ occurring in the various summands $\Delta(s') \cdot TD^{\mathrm{c}}_{n-1}(s')$ due to the fact that the probabilities of identical traces are always added up according to Def. 6.1.

- If $n \geq 3$ then each element T of $TD^{\mathrm{mc}}_n(s)$, whose nonempty traces all start with some action a labeling an outgoing transition of s, say $s \xrightarrow{a} \Delta$, is obtained by summing up a suitable element $T_{s'}$ of $TD^{\mathrm{mc}}_{n-1}(s')$ for every $s' \in supp(\Delta)$. The reason is that no element $T'_{s'}$ of $TD^{\mathrm{c}}_{n-1}(s') \setminus TD^{\mathrm{mc}}_{n-1}(s')$ includes as a subset an element of $TD^{\mathrm{mc}}_{n-2}(s')$ and hence its weighted traces of length at most $n - 2$ cannot contribute to the construction of an element of

$TD^{\mathrm{mc}}_{n-1}(s)$ included as a subset of an element of $TD^{\mathrm{mc}}_{n}(s)$. Indeed, with respect to any element of $TD^{\mathrm{mc}}_{n-2}(s')$ of which it contains all traces of length at most $n-2$, each such $T'_{s'}$ has a different probability associated with at least one of those traces, with the probabilities of the corresponding left-extended traces of length $m \geq n$ executable by an upstream state (i.e., s or one of its predecessors) remaining different from each other because they are obtained by multiplying the probabilities of the original traces by the same value, which is the probability of reaching s' from the upstream state in $m-n+1$ steps. □

By virtue of Prop. 6.1, the equality $TD^{\mathrm{mc}}_{n}(s) = TD^{\mathrm{c}}_{n}(s)$ extends to all $n \in \mathbb{N}$, i.e., memoryfully coherent trace distributions boil down to the corresponding coherent ones, in the case of a fully probabilistic NPLTS. This holds in particular for resolutions.

Proposition 6.4. Let $(S, A, \longrightarrow)$ be a fully probabilistic NPLTS, $s \in S$, and $n \in \mathbb{N}$. Then $TD^{\mathrm{mc}}_{n}(s) = TD^{\mathrm{c}}_{n}(s)$.

Proof. We proceed by induction on $n \in \mathbb{N}$:

- For $n = 0$ we have that $TD^{\mathrm{mc}}_{n}(s) = \{\{(\varepsilon, 1)\}\} = TD^{\mathrm{c}}_{n}(s)$.
- Let $n = m + 1$ for some $m \in \mathbb{N}$ and suppose that the result holds when considering traces of length at most m. By virtue of Prop. 6.1 we have that:
  $$TD^{\mathrm{c}}_{n}(s) = \{\{(\alpha, p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s, \alpha)) = p\}\}$$
  $$TD^{\mathrm{c}}_{m}(s) = \{\{(\alpha, p) \in A^{\leq m} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s, \alpha)) = p\}\}$$
  where from $m < n$ it follows that $TD^{\mathrm{c}}_{m}(s) \subseteq TD^{\mathrm{c}}_{n}(s)$.
  Since $TD^{\mathrm{mc}}_{m}(s) = TD^{\mathrm{c}}_{m}(s)$ by the induction hypothesis, we have proved that the only weighted trace set in $TD^{\mathrm{c}}_{n}(s)$ includes as a subset the only weighted trace set in $TD^{\mathrm{mc}}_{m}(s)$, hence $TD^{\mathrm{mc}}_{n}(s) = TD^{\mathrm{c}}_{n}(s)$ too. □

6.4 Coherency Constraints for Resolutions: $\sim^{\mathbf{post,c}}_{\mathbf{PTr}}$ and $\sim^{\mathbf{pre,c}}_{\mathbf{PTr}}$

We are finally in a position of formalizing the coherency constraints based on the comparison of $tr(TD^{\mathrm{mc}}_{n})$ sets between models and resolutions and

on the completeness of traces up to a certain length within resolutions. As in Sect. 4, in the following $Res(\_)$ denotes any of the sets of resolutions introduced in Defs. 3.1 to 3.3.

Definition 6.4. Let $\mathcal{L} = (S, A, \longrightarrow_{\mathcal{L}})$ be an NPLTS, $s \in S$, and $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}}) \in Res(s)$ with correspondence function $corr_{\mathcal{Z}} : Z \to S$. We say that $\mathcal{Z}$ is a *coherent resolution* of s, written $\mathcal{Z} \in Res^{\mathrm{c}}(s)$, iff for all $z \in Z$, whenever $z \xrightarrow{a}_{\mathcal{Z}} \Delta$, then for all $n \in \mathbb{N}$ the following two constraints are met:

1. For all $z', z'' \in supp(\Delta)$, if:
$$tr(TD_n^{\mathrm{mc}}(corr_{\mathcal{Z}}(z'))) \;=\; tr(TD_n^{\mathrm{mc}}(corr_{\mathcal{Z}}(z'')))$$
then:
$$tr(TD_n^{\mathrm{mc}}(z')) \;=\; tr(TD_n^{\mathrm{mc}}(z''))$$

2. Either every $z' \in supp(\Delta)$ satisfies "for the only $T \in TD_n^{\mathrm{mc}}(z')$ there exists $\overline{T} \in TD_n^{\mathrm{mc}}(corr_{\mathcal{Z}}(z'))$ such that $tr(T) = tr(\overline{T})$", or at least one $z' \in supp(\Delta)$ does not satisfy the aforementioned property and every $z'' \in supp(\Delta)$ satisfying it is such that the longest trace in the corresponding $tr(\overline{T})$ has length less than n. ■

The first constraint requires that, whenever two states in the support of a transition in the original model have memoryfully coherent trace distributions with the same sets of traces up to an arbitrary length n, then so have the states to which they correspond in the resolution.

The second constraint requires what follows: either all the states in the support of a transition of the resolution have all the traces up to an arbitrary length n that occur in weighted trace sets of the states to which they are mapped in the original model, or at least from one of them the execution stops in less than n steps with respect to the state to which it is mapped in the original model and all the states satisfying the considered property exhibit traces of length less than n that are maximal in the states to which they are mapped in the original model. For instance, in Fig. 6.3 consider the resolution of r' featuring only the a-transition and take $n \geq 1$ for the state without transitions in the support of that transition, which is mapped to a state in the

original model with a b-transition instead. This resolution can be deemed coherent thanks to the second part of the second constraint.

It is worth noting that any complete submodel rooted at a state z of a coherent resolution turns out to be coherent too, where complete means that no state reachable from z in the resolution is cut off in the resolution submodel. Submodel completeness is important for satisfying in particular the second coherency constraint of Def. 6.4, i.e., trace completeness up to a certain length within resolutions.

Proposition 6.5. Let $\mathcal{L} = (S, A, \longrightarrow_{\mathcal{L}})$ be an NPLTS, $s \in S$, and $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}}) \in Res^{c}(s)$ with correspondence function $corr_{\mathcal{Z}} : Z \to S$. Let $\mathcal{Z}'_z = (Z', A, \longrightarrow_{\mathcal{Z}'})$ be the complete submodel of $\mathcal{Z}$ rooted at $z \in Z$. Then $\mathcal{Z}'_z \in Res^{c}(corr_{\mathcal{Z}}(z))$.

Proof. It immediately derives from the fact that every state of $\mathcal{Z}$ having an outgoing transition fulfills both constraints of Def. 6.4, because every state of $\mathcal{Z}'_z$ is a state of $\mathcal{Z}$ too (being $\mathcal{Z}'_z$ a submodel of $\mathcal{Z}$) and every state reachable from z in $\mathcal{Z}$ is reachable from z in $\mathcal{Z}'_z$ too (due to the completeness of $\mathcal{Z}'_z$). □

As far as the counterexamples in Figs. 5.1 to 6.2 are concerned, we observe that the resolution in Fig. 5.2 does not belong to $Res^{c}_{sp}(s_3)$ because it violates only the first constraint of Def. 6.4, while the resolution in Fig. 6.2 does not belong to $Res^{c}_{sp}(s_{10})$ because it violates only the second constraint of Def. 6.4 (consider the a-transition of z_{10}, take $n = 2$ for z'_{10} to z'''_{10}, and note that z''_{10} and z'''_{10} stop the execution earlier than s''_{10} and s'''_{10} but z'_{10} exhibits the maximal trace $b\,c_1$ of s'_{10} whose length is not less than 2). The resolutions in Figs. 5.1, 5.3, 6.1 do not respectively belong to $Res^{c}_{sp}(s_2)$, $Res^{c}_{sp}(s_6)$, $Res^{c}_{sp}(s_7)$ because they violate both constraints.

As a consequence, none of the resolutions above would be considered by the following coherency-based variants of $\sim^{post}_{PTr}$ and $\sim^{pre}_{PTr}$.

Definition 6.5. Let $(S, A, \longrightarrow)$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim^{post,c}_{PTr} s_2$ iff for each $\mathcal{Z}_1 \in Res^{c}(s_1)$ there exists $\mathcal{Z}_2 \in Res^{c}(s_2)$ such that for all $\alpha \in A^*$:

$$prob(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; prob(\mathcal{CC}(z_{s_2}, \alpha))$$

and the same condition holds when exchanging $\mathcal{Z}_1$ with $\mathcal{Z}_2$. ■

Definition 6.6. Let $(S, A, \longrightarrow)$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim_{\mathrm{PTr}}^{\mathrm{pre,c}} s_2$ iff, for all $\alpha \in A^*$, for each $\mathcal{Z}_1 \in \mathit{Res}^{\mathrm{c}}(s_1)$ there exists $\mathcal{Z}_2 \in \mathit{Res}^{\mathrm{c}}(s_2)$ such that:

$$\mathit{prob}(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; \mathit{prob}(\mathcal{CC}(z_{s_2}, \alpha))$$

and the same condition holds when exchanging $\mathcal{Z}_1$ with $\mathcal{Z}_2$. ■

We conclude by proving that the anomalies illustrated in the aforementioned figures disappear when using coherent resolutions. In the following, we lift a probabilistic behavioral equivalence $\sim$ from states to distributions over states by letting $\Delta_1 \sim \Delta_2$ iff $\Delta_1(C) = \Delta_2(C)$ for all equivalence classes C of $\sim$. Moreover, the action prefix construction $a\,.\,\Delta$ stands for an a-transition whose target distribution is Δ. Finally, $\sim_{\mathrm{PTr}}^{\mathrm{fp}}$ denotes the probabilistic trace equivalence for fully probabilistic processes defined in (Jou and Smolka, 1990), i.e., $s_1 \sim_{\mathrm{PTr}}^{\mathrm{fp}} s_2$ iff $\mathit{prob}(\mathcal{CC}(s_1, \alpha)) = \mathit{prob}(\mathcal{CC}(s_2, \alpha))$ for all $\alpha \in A^*$.

Theorem 6.1. Let $\mathcal{L} = (S, A, \longrightarrow)$ be an NPLTS, $s_1, s_2 \in S$, $\Delta_1, \Delta_2 \in \mathit{Distr}(S)$, and $\sim_{\mathrm{PTr}}^{\mathrm{c}} \in \{\sim_{\mathrm{PTr}}^{\mathrm{post,c}}, \sim_{\mathrm{PTr}}^{\mathrm{pre,c}}\}$. Under coherent resolutions induced by deterministic schedulers, it holds that:

1. $s_1 \sim_{\mathrm{PB}} s_2 \implies s_1 \sim_{\mathrm{PTr}}^{\mathrm{c}} s_2$.
2. For all $a \in A$, $\Delta_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} \Delta_2 \implies a\,.\,\Delta_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} a\,.\,\Delta_2$.
3. If $\mathcal{L}$ is fully probabilistic, then $s_1 \sim_{\mathrm{PTr}}^{\mathrm{c}} s_2 \iff s_1 \sim_{\mathrm{PTr}}^{\mathrm{fp}} s_2$.

Proof. Given an NPLTS $\mathcal{L} = (S, A, \longrightarrow)$, $s_1, s_2 \in S$, and $\Delta_1, \Delta_2 \in \mathit{Distr}(S)$, we proceed as follows:

1. We show that, from $(s_1, s_2) \in \mathcal{B}$ for some probabilistic bisimulation $\mathcal{B}$, it follows that $(\star)$ for each $\mathcal{Z}_1 = (Z_1, A, \longrightarrow_{\mathcal{Z}_1}) \in \mathit{Res}_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – resp. $\mathcal{Z}_2 = (Z_2, A, \longrightarrow_{\mathcal{Z}_2}) \in \mathit{Res}_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – there exists $\mathcal{Z}_2 = (Z_2, A, \longrightarrow_{\mathcal{Z}_2}) \in \mathit{Res}_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – resp. $\mathcal{Z}_1 = (Z_1, A, \longrightarrow_{\mathcal{Z}_1}) \in \mathit{Res}_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – such that for all $\alpha \in A^*$ it holds that:
   $$\mathit{prob}(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; \mathit{prob}(\mathcal{CC}(z_{s_2}, \alpha))$$
   where z_{s_i} denotes both the initial state of $\mathcal{Z}_i$ and the state to which s_i corresponds. From this it will follow that $s_1 \sim_{\mathrm{PB}} s_2 \implies s_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} s_2$ and also $s_1 \sim_{\mathrm{PB}} s_2 \implies s_1 \sim_{\mathrm{PTr}}^{\mathrm{pre,c}} s_2$ because

$s_1 \sim^{\mathrm{post,c}}_{\mathrm{PTr}} s_2 \implies s_1 \sim^{\mathrm{pre,c}}_{\mathrm{PTr}} s_2$.
Starting from s_1, we focus on an arbitrary $\mathcal{Z}_1 = (Z_1, A, \longrightarrow_{\mathcal{Z}_1}) \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_1)$, which we assume not to consist of a single state without transitions so as to avoid trivial cases. Let $z_{s_1} \xrightarrow{a}_{\mathcal{Z}_1} \Delta_1$ be the initial transition of $\mathcal{Z}_1$, which we assume to derive from $s_1 \xrightarrow{a} \Gamma_1$. Since $(s_1, s_2) \in \mathcal{B}$ and $\mathcal{B}$ is a probabilistic bisimulation, there must exist $\mathcal{Z}_2 = (Z_2, A, \longrightarrow_{\mathcal{Z}_2}) \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_2)$ with initial transition $z_{s_2} \xrightarrow{a}_{\mathcal{Z}_2} \Delta_2$, which we assume to derive from $s_2 \xrightarrow{a} \Gamma_2$, such that, in particular, for each $C \subseteq Z_1 \cup Z_2$, whose image via $corr_{\mathcal{Z}_1} \cup corr_{\mathcal{Z}_2}$ is an equivalence class in $S/\mathcal{B}$, it holds that:

$$\Gamma_1(corr_{\mathcal{Z}_1}(C \cap Z_1)) \;=\; \Gamma_2(corr_{\mathcal{Z}_2}(C \cap Z_2))$$

and hence by definition of deterministic scheduler we have that $(*)$:

$$\begin{aligned}\Delta_1(C \cap Z_1) \;&=\; \Gamma_1(corr_{\mathcal{Z}_1}(C \cap Z_1)) \;= \\ &=\; \Gamma_2(corr_{\mathcal{Z}_2}(C \cap Z_2)) \;=\; \Delta_2(C \cap Z_2)\end{aligned}$$

Among all the resolutions $\mathcal{Z}_2 \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_2)$ satisfying the equality above, we choose the one that can execute all the traces of $\mathcal{Z}_1$ (which must exist otherwise s_1 could execute a trace not executable by s_2 and hence $s_1 \sim_{\mathrm{PB}} s_2$ would be contradicted) and only those traces (i.e., longer traces are excluded).
Given an arbitrary $\alpha \in A^*$, we prove property $(\star)$ by proceeding by induction on $|\alpha| \in \mathbb{N}$:

- If $|\alpha| = 0$, i.e., $\alpha = \varepsilon$, then it trivially holds that:
$$prob(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; 1 \;=\; prob(\mathcal{CC}(z_{s_2}, \alpha))$$
- Let $|\alpha| = n + 1$ for some $n \in \mathbb{N}$, with $\alpha = a'\,\alpha'$ and $|\alpha'| = n$, and suppose that property $(\star)$ holds for each trace of length n when starting from two probabilistic bisimilar states. There are two cases:
  - If $a' \neq a$, since both $\mathcal{Z}_1$ and $\mathcal{Z}_2$ start with an a-transition, it trivially holds that:
  $$prob(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; 0 \;=\; prob(\mathcal{CC}(z_{s_2}, \alpha))$$
  - If $a' = a$, due to $(*)$ we observe that an arbitrary $C \subseteq Z_1 \cup Z_2$, whose image via $corr_{\mathcal{Z}_1} \cup corr_{\mathcal{Z}_2}$ is an equivalence class in $S/\mathcal{B}$, is either reachable via both

a-transitions, or via neither. Moreover, thanks to the coherency of $\mathcal{Z}_1$, the coherency of $\mathcal{Z}_2$, and the choice of $\mathcal{Z}_2$ with respect to the capability of executing all the traces of $\mathcal{Z}_1$ and only those traces, either α' is executable in all the states of C, or in none of them (for a counterexample in the absence of coherency, see Fig. 5.1).

Let $\mathcal{G}$ be the set of subsets of $Z_1 \cup Z_2$, whose images via $corr_{\mathcal{Z}_1} \cup corr_{\mathcal{Z}_2}$ are equivalence classes in $S/\mathcal{B}$, that are reachable via both a-transitions and in which α' is executable. Note that the other subsets do not contribute to $prob(\mathcal{CC}(z_{s_1}, \alpha))$ and $prob(\mathcal{CC}(z_{s_2}, \alpha))$. For each $C \in \mathcal{G}$, given an arbitrary $z_{C,1} \in C \cap supp(\Delta_1)$ and an arbitrary $z_{C,2} \in C \cap supp(\Delta_2)$ whose corresponding states in S are $s_{C,1}$ and $s_{C,2}$, since $s_{C,1} \sim_{\mathrm{PB}} s_{C,2}$ and $|\alpha'| = n$ by the induction hypothesis and the coherency of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ we have that:

$$prob(\mathcal{CC}(z_{C,1}, \alpha')) \;=\; prob(\mathcal{CC}(z_{C,2}, \alpha'))$$

with the same reasoning being applicable to any pair of states in $C \cap supp(\Delta_1)$ and any pair of states in $C \cap supp(\Delta_2)$. As a consequence, by virtue of $(*)$ it follows that:

$$\begin{aligned} prob(\mathcal{CC}(z_{s_1}, \alpha)) &= \sum_{C \in \mathcal{G}} \Delta_1(C \cap Z_1) \cdot prob(\mathcal{CC}(z_{C,1}, \alpha')) \\ &= \sum_{C \in \mathcal{G}} \Delta_2(C \cap Z_2) \cdot prob(\mathcal{CC}(z_{C,2}, \alpha')) \\ &= prob(\mathcal{CC}(z_{s_2}, \alpha)) \end{aligned}$$

2. Let $\Delta_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} \Delta_2$, i.e., $\Delta_1(K) = \Delta_2(K)$ for all equivalence classes $K \in S/\!\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$. Then, in particular, for each $s_1 \in supp(\Delta_1)$ there must exist $s_2 \in supp(\Delta_2)$ such that $s_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} s_2$, and vice versa. Given $a \in A$, the only interesting case in the proof of $a \,.\, \Delta_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} a \,.\, \Delta_2$ is the one in which we consider a trace of the form $a\,\alpha'$ and, for $j \in \{1, 2\}$, a resolution $\mathcal{Z}_j = (Z_j, A, \longrightarrow_{\mathcal{Z}_j}) \in Res_{\mathrm{sp}}^{\mathrm{c}}(a \,.\, \Delta_j)$, whose initial state we denote by $z_{a \,.\, \Delta_j}$, not consisting of a single state without transitions. By definition of deterministic scheduler, the initial a-transition of $\mathcal{Z}_j$ reaches with probability $p_K = \Delta_j(K)$ the set of states in $\mathcal{Z}_j$ whose corresponding states in $\mathcal{L}$ via $corr_{\mathcal{Z}_j}$

are in the same equivalence class $K \in S/{\sim}^{\text{post,c}}_{\text{PTr}}$. It is correct to consider the sum p_K of the probabilities of those states belonging to the support of the target distribution of the initial a-transition of $\mathcal{Z}_j$ because, thanks to the coherency of $\mathcal{Z}_j$, two states in that support have to possess the same traces if so do their corresponding states in $supp(\Delta_j)$, as is the case with the states of K (for a counterexample in the absence of coherency, see Fig. 5.2).
Given $K \in S/{\sim}^{\text{post,c}}_{\text{PTr}}$ and $s_{j,K} \in supp(\Delta_j) \cap K$, consider the complete submodel $\mathcal{Z}_{j,K} = (Z_{j,K}, A, \longrightarrow_{\mathcal{Z}_{j,K}}) \in Res^{\text{c}}_{\text{sp}}(s_{j,K})$ of $\mathcal{Z}_j$, whose initial state we denote by $z_{s_{j,K}}$. Then, for any other $s'_{j,K} \in supp(\Delta_j) \cap K$, the complete submodel $\mathcal{Z}'_{j,K} = (Z'_{j,K}, A, \longrightarrow_{\mathcal{Z}'_{j,K}}) \in Res^{\text{c}}_{\text{sp}}(s'_{j,K})$ of $\mathcal{Z}_j$ must match $\mathcal{Z}_{j,K}$ according to ${\sim}^{\text{post,c}}_{\text{PTr}}$, i.e., $z_{s_{j,K}} \sim^{\text{post,c}}_{\text{PTr}} z_{s'_{j,K}}$, because $s_{j,K} \sim^{\text{post,c}}_{\text{PTr}} s'_{j,K}$, resolution $\mathcal{Z}_j$ is coherent, and $\mathcal{Z}_{j,K}$ and $\mathcal{Z}'_{j,K}$ are complete submodels of $\mathcal{Z}_j$.
Starting from $a \,.\, \Delta_1$ and $\mathcal{Z}_1$, for any $\alpha = a\,\alpha' \in A^*$ we have that:

$$\begin{aligned} prob(\mathcal{CC}(z_{a\,.\,\Delta_1}, \alpha)) &= \sum_{K \in S/{\sim}^{\text{post,c}}_{\text{PTr}}}^{K \cap supp(\Delta_1) \neq \emptyset} p_K \cdot prob(\mathcal{CC}(z_{s_{1,K}}, \alpha')) \\ &= \sum_{K \in S/{\sim}^{\text{post,c}}_{\text{PTr}}}^{K \cap supp(\Delta_2) \neq \emptyset} p_K \cdot prob(\mathcal{CC}(z_{s_{2,K}}, \alpha')) \\ &= prob(\mathcal{CC}(z_{a\,.\,\Delta_2}, \alpha)) \end{aligned}$$

where the existence of $\mathcal{Z}_{2,K} = (Z_{2,K}, A, \longrightarrow_{\mathcal{Z}_{2,K}}) \in Res^{\text{c}}_{\text{sp}}(s_{2,K})$ matching $\mathcal{Z}_{1,K}$ according to $\sim^{\text{post,c}}_{\text{PTr}}$ stems from the existence – mentioned at the beginning of the proof – of $s_{2,K} \in supp(\Delta_2)$ such that $s_{1,K} \sim^{\text{post,c}}_{\text{PTr}} s_{2,K}$. Therefore $\mathcal{Z}_2 = (Z_2, A, \longrightarrow_{\mathcal{Z}_2}) \in Res^{\text{c}}_{\text{sp}}(s_2)$, which starts with an a-transition and continues as $\mathcal{Z}_{2,K}$ for $s_{2,K} \in supp(\Delta_2) \cap K$, matches $\mathcal{Z}_1$ according to $\sim^{\text{post,c}}_{\text{PTr}}$.

3. If $\mathcal{L}$ is fully probabilistic, then it has a single maximal resolution, which (coincides with $\mathcal{L}$ itself if $\mathcal{L}$ is acyclic and) is the one on which the probabilities of all the traces are computed when verifying $\sim^{\text{fp}}_{\text{PTr}}$. Any of the other coherent resolutions, which is considered only when verifying $\sim^{\text{c}}_{\text{PTr}}$, is obtained by stopping in advance the execution of $\mathcal{L}$. This is accomplished in a way that is coherent along all branches of the maximal resolution, not only in terms of transition selection due to the first constraint of Def. 6.4,

but also in terms of complete presence of computations up to a certain length by virtue of the second constraint of Def. 6.4. □

Note that property 2, i.e., congruence with respect to action prefix under deterministic schedulers, can be restored in its most liberal form only for $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$. In the case of $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$, a more limited form should be considered like the one in Thm. 4.2 of (Bernardo, 2018) where, instead of admitting two equivalent distributions whose supports may have different cardinalities, pairs of equivalent states are embedded into two copies of the same distribution skeleton.

To understand why resolution coherency is not enough for $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$, look at Fig. 4.1 and consider a distribution Δ_1 such that $supp(\Delta_1) = \{s', s''\}$, with $\Delta_1(s') = p$ and $\Delta_1(s'') = 1 - p$ for $p \in \mathbb{R}_{]0,1[}$, and a distribution Δ_2 such that $supp(\Delta_2) = \{s'''\}$. Note that $|supp(\Delta_1)| \neq |supp(\Delta_2)|$ and recall that s', s'', s''' are all related by $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$. Then $\Delta_1 \sim_{\mathrm{PTr}}^{\mathrm{pre,c}} \Delta_2$ because $\Delta_1(C) = \Delta_2(C) = 1$ for the equivalence class C of $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ containing the three states in the support of the two distributions. However $a' \,.\, \Delta_1 \not\sim_{\mathrm{PTr}}^{\mathrm{pre,c}} a' \,.\, \Delta_2$ because $a' \,.\, \Delta_1$ can execute trace $a'\, a\, b_1$ with probabilities $p \cdot 0.5$, $(1-p) \cdot 0.5$, 0.5 depending on the considered coherent resolution, while $a' \,.\, \Delta_2$ can execute that trace only with probability 0.5, which cannot match $p \cdot 0.5$ and $(1-p) \cdot 0.5$.

The variety of those execution probabilities arises not only from the $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$-equivalent states s' and s'' being in the support of the target distribution of the same transition, but also from the fact that the states reachable from s' and s'' contain probabilistic choices. This is not the case in Fig. 5.2, where the most liberal form of congruence with respect to action prefix applies because the states reachable from the $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$-equivalent states s'_3 and s''_3 do not contain probabilistic choices.

Resolution coherency under deterministic schedulers was unfortunately neglected in (Bernardo *et al.*, 2014a; Bernardo *et al.*, 2014b), so that property 1 above is the rectified version of a chain of results in (Bernardo *et al.*, 2014b) consisting of Thms. 6.5(2), 5.9(3), 4.5(2), while property 3 above is the rectified version of Thm. 3.4(2) of (Bernardo *et al.*, 2014a; Bernardo *et al.*, 2014b).

As a final remark, we note that property 3 holds also in the case of randomized/interpolating schedulers thanks to constraint 2 of Def. 6.4.

7

Alternative Characterizations of Trace Semantics

Trace equivalences can in general be characterized through some form of trace set equality akin to language equivalence. For instance, two fully nondeterministic processes turn out to be trace equivalent iff they possess the same trace set (Brookes *et al.*, 1984). As another example, two fully probabilistic processes turn out to be trace equivalent iff they possess the same weighted trace set (Jou and Smolka, 1990).

Alternative characterizations are not as straightforward over nondeterministic and probabilistic processes, because traces can have different execution probabilities in different coherent resolutions. However, they can be obtained under centralized, memoryless schedulers of deterministic nature: $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$ coincides with memoryfully coherent trace distribution equality (Sect. 7.1) whereas $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ coincides with coherent weighted trace set equality (Sect. 7.2). We use the latter to express some remarks about congruence with respect to parallel composition (Sect. 7.3).

7.1 Alternative Characterization of $\sim_{\mathbf{PTr}}^{\mathbf{post,c}}$

The definition of $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$ essentially requires that two states have the same memoryfully coherent trace distributions. Therefore, it is natural to expect an alternative characterization of $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$ based on the con-

struction of Defs. 6.2 and 6.3. Incidentally, this would fully justify the construction itself, given that the probabilities contained in the TD_n^{mc} sets have not been exploited in the coherency constraints of Def. 6.4.

The following lemma, where Prop. 6.1 is exploited again together with Props. 6.5, 6.3, and 6.4, lays the basis for a characterization of $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$ in terms of memoryfully coherent trace distribution equality. It establishes that a weighted trace set T is in the memoryfully coherent trace distribution of a state s iff it can be exhibited by a coherent resolution $\mathcal{Z}$ of that state. In the lemma, z_s denotes both the initial state of $\mathcal{Z}$ and the state to which s corresponds.

Lemma 7.1. Let $(S, A, \longrightarrow)$ be an NPLTS, $s \in S$, $n \in \mathbb{N}$, and $T \subseteq A^* \times \mathbb{R}_{]0,1]}$. Then $T \in TD_n^{\mathrm{mc}}(s)$ iff there exists $\mathcal{Z} \in Res_{\mathrm{sp}}^{\mathrm{c}}(s)$ such that $TD_n^{\mathrm{mc}}(z_s) = \{T\}$.

Proof. We proceed by induction on $n \in \mathbb{N}$:

- For $n = 0$ we have that $TD_n^{\mathrm{mc}}(s) = \{\{(\varepsilon, 1)\}\} = TD_n^{\mathrm{mc}}(z_s)$ with z_s being the initial state of any $\mathcal{Z} \in Res_{\mathrm{sp}}^{\mathrm{c}}(s)$, hence the result trivially follows.

- Let $n = m + 1$ for some $m \in \mathbb{N}$ and suppose that the result holds for each weighted trace set, of any state, whose traces have length at most m. To avoid trivial cases, we assume that $T \neq \{(\varepsilon, 1)\}$ and that s has outgoing transitions. The proof is divided into two parts:

  - Let $T \in TD_n^{\mathrm{mc}}(s)$, with all of its nonempty traces starting with some action a labeling an outgoing transition of s, say $s \xrightarrow{a} \Delta$. From $T \in TD_n^{\mathrm{mc}}(s)$ we derive that $T \in TD_n^{\mathrm{c}}(s)$ and hence $[\star]$:
    $$T \in (\varepsilon, 1) \dagger a \,.\, \left(\sum_{\Theta \in tr(\Delta, m)} \sum_{s' \in supp(\Delta)}^{tr(TD_m^{\mathrm{c}}(s')) = \Theta} \Delta(s') \cdot TD_m^{\mathrm{c}}(s') \right)$$
    From Prop. 6.3 we further derive that T is obtained by adding up – according to Def. 6.1 applied to the double summation of $[\star]$ – a suitable element $T_{s'}$ of $TD_m^{\mathrm{mc}}(s')$ for every $s' \in supp(\Delta)$.

From the induction hypothesis, for each such $T_{s'}$ it follows that there exists $\mathcal{Z}_{s'} = (Z_{s'}, A, \longrightarrow_{\mathcal{Z}_{s'}}) \in Res^{c}_{sp}(s')$ such that $TD^{mc}_{m}(z_{s'}) = \{T_{s'}\}$. Without loss of generality, we can assume that each $\mathcal{Z}_{s'}$ has computations of length at most m and, subject to this, all of its computations are maximal (in the sense that no further steps can be added to reach length m) with respect to the corresponding computations from s'.
Consider now the resolution $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}})$ of s such that (i) the image via $corr_{\mathcal{Z}}$ of its initial transition $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$ is $s \xrightarrow{a} \Delta$ and (ii) from each $z_{s'} \in supp(\Delta')$ the resolution behaves as $\mathcal{Z}_{s'}$. Then $\{T\} = TD^{c}_{n}(z_s) = TD^{mc}_{n}(z_s)$, due to Prop. 6.4, and $\mathcal{Z} \in Res^{c}_{sp}(s)$, because each $\mathcal{Z}_{s'}$ is coherent and $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$ satisfies both constraints of Def. 6.4 (for an example, start from the b-transition of r_1 in Fig. 6.3).
The satisfaction of the first constraint stems from T being an element of $TD^{mc}_{n}(s)$ and $\mathcal{Z}$ having computations of length at most n. Specifically, for $z'_1, z'_2 \in supp(\Delta')$, if $tr(TD^{mc}_{m}(corr_{\mathcal{Z}}(z'_1))) = tr(TD^{mc}_{m}(corr_{\mathcal{Z}}(z'_2)))$ then we have that $TD^{c}_{m}(corr_{\mathcal{Z}}(z'_1))$ and $TD^{c}_{m}(corr_{\mathcal{Z}}(z'_2))$ are added up in the innermost summation of $[\star]$ according to Def. 6.1. This requires in the construction of T based on $[\star]$ that the two sets $T_{corr_{\mathcal{Z}}(z'_1)} \in TD^{mc}_{m}(corr_{\mathcal{Z}}(z'_1))$ and $T_{corr_{\mathcal{Z}}(z'_2)} \in TD^{mc}_{m}(corr_{\mathcal{Z}}(z'_2))$ are coherent with each other, thus guaranteeing that $tr(TD^{mc}_{m}(z'_1)) = tr(TD^{mc}_{m}(z'_2))$. The constraint holds for any $m' < m$ too, because T contains as a subset $T' \in TD^{mc}_{m'}(s)$ and hence coherency constraints on shorter traces are not forgotten.
The satisfaction of the second constraint stems from all computations of each $\mathcal{Z}_{s'}$ being maximal (i.e., no further steps can be added to reach length m) with respect to the corresponding computations from s'.

– Let $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}}) \in Res^{c}_{sp}(s)$ be such that $TD^{mc}_{n}(z_s) = \{T\}$, with all nonempty traces of T starting with some action a labeling an outgoing transition of z_s, say $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$. By virtue of Prop. 6.4, from $\{T\} = TD^{mc}_{n}(z_s)$ we derive that

$\{T\}$ is equal to:

$$(\varepsilon,1)\dagger a\,.\left(\sum_{\Theta\in tr(\Delta',m)}\ \sum_{z'\in supp(\Delta')}^{tr(TD_m^{\mathrm{mc}}(z'))=\Theta}\Delta'(z')\cdot TD_m^{\mathrm{mc}}(z')\right)$$

For all $z'\in supp(\Delta')$, since the complete submodel of $\mathcal{Z}$ rooted at z' is still coherent, thanks to Prop. 6.5, and satisfies $TD_m^{\mathrm{mc}}(z')=\{T_{z'}\}$ for some $T_{z'}\subseteq A^{\leq m}\times\mathbb{R}_{]0,1]}$, thanks to Props. 6.4 and 6.1, from the induction hypothesis it follows that $T_{z'}\in TD_m^{\mathrm{mc}}(corr_{\mathcal{Z}}(z'))$.

Since T is obtained by summing up $T_{z'}\in TD_m^{\mathrm{mc}}(corr_{\mathcal{Z}}(z'))$ for every $z'\in supp(\Delta')$, for $s\xrightarrow{a}\Delta$ corresponding via $corr_{\mathcal{Z}}$ to $z_s\xrightarrow{a}_{\mathcal{Z}}\Delta'$ it holds that T belongs to:

$$(\varepsilon,1)\dagger a\,.\left(\sum_{\Theta\in tr(\Delta,m)}\ \sum_{corr_{\mathcal{Z}}(z')\in supp(\Delta)}^{tr(T_{z'})=\Theta}\Delta(corr_{\mathcal{Z}}(z'))\cdot T_{z'}\right)$$

and hence $T\in TD_n^{\mathrm{mc}}(s)$ because $\mathcal{Z}\in Res_{\mathrm{sp}}^{\mathrm{c}}(s)$. □

Theorem 7.2. Let $(S,A,\longrightarrow)$ be an NPLTS and $s_1,s_2\in S$. Then $s_1\sim_{\mathrm{PTr}}^{\mathrm{post,c}}s_2$ iff $TD^{\mathrm{mc}}(s_1)=TD^{\mathrm{mc}}(s_2)$.

Proof. By definition, $s_1\sim_{\mathrm{PTr}}^{\mathrm{post,c}}s_2$ iff for each $\mathcal{Z}_1\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – resp. $\mathcal{Z}_2\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – there exists $\mathcal{Z}_2\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – resp. $\mathcal{Z}_1\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – such that for all $\alpha\in A^*$:

$$prob(\mathcal{CC}(z_{s_1},\alpha))\ =\ prob(\mathcal{CC}(z_{s_2},\alpha))$$

Let $A^{\leq n}=\{\alpha\in A^*\mid|\alpha|\leq n\}$ for $n\in\mathbb{N}$. Then $s_1\sim_{\mathrm{PTr}}^{\mathrm{post,c}}s_2$ iff for each $\mathcal{Z}_1\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – resp. $\mathcal{Z}_2\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – there exists $\mathcal{Z}_2\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – resp. $\mathcal{Z}_1\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – such that for all $n\in\mathbb{N}$ and $\alpha\in A^{\leq n}$:

$$prob(\mathcal{CC}(z_{s_1},\alpha))\ =\ prob(\mathcal{CC}(z_{s_2},\alpha))$$

Thanks to Props. 6.4 and 6.1, for $j\in\{1,2\}$ we have that for all $n\in\mathbb{N}$:

$$TD_n^{\mathrm{mc}}(z_{s_j})\ =\ \{\{(\alpha,p)\in A^{\leq n}\times\mathbb{R}_{]0,1]}\mid prob(\mathcal{CC}(z_{s_j},\alpha))=p\}\}$$

Therefore $s_1\sim_{\mathrm{PTr}}^{\mathrm{post,c}}s_2$ iff for each $\mathcal{Z}_1\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – resp. $\mathcal{Z}_2\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – there exists $\mathcal{Z}_2\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2)$ – resp. $\mathcal{Z}_1\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1)$ – such that for all $n\in\mathbb{N}$:

$$TD_n^{\mathrm{mc}}(z_{s_1})\ =\ TD_n^{\mathrm{mc}}(z_{s_2})$$

This is the same as for all $n\in\mathbb{N}$ and $T\subseteq A^*\times\mathbb{R}_{]0,1]}$:

$$\exists\mathcal{Z}\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_1).\ TD_n^{\mathrm{mc}}(z_{s_1})=\{T\}$$
$$\Longleftrightarrow$$
$$\exists\mathcal{Z}\in Res_{\mathrm{sp}}^{\mathrm{c}}(s_2).\ TD_n^{\mathrm{mc}}(z_{s_2})=\{T\}$$

which by virtue of Lemma 7.1 amounts to for all $n \in \mathbb{N}$ and $T \subseteq A^* \times \mathbb{R}_{]0,1]}$:

$$T \in TD_n^{\mathrm{mc}}(s_1) \iff T \in TD_n^{\mathrm{mc}}(s_2)$$

which in turn is equivalent to $TD^{\mathrm{mc}}(s_1) = TD^{\mathrm{mc}}(s_2)$. □

7.2 Alternative Characterization of $\sim_{\mathbf{PTr}}^{\mathbf{pre,c}}$

As far as $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ is concerned, similar to (Bernardo *et al.*, 2014a) we can provide an alternative characterization based on trace sets built by considering all weighted traces executable from state s at once, i.e., without keeping track of the resolutions of s in which they are feasible. This is consistent with the focus of $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ on individual traces rather than on trace distributions. In the definition below, which has the same structure as Def. 6.2, there is no need of a double summation in the case that $n > 0$ and s has outgoing transitions thanks to the commutativity and associativity of weighted trace set addition deriving from Def. 6.1.

Definition 7.1. Let $(S, A, \longrightarrow)$ be an NPLTS and $s \in S$. The *coherent weighted trace set* of s is the subset of $A^* \times \mathbb{R}_{]0,1]}$ defined as follows:

$$T^{\mathrm{c}}(s) = \bigcup_{n \in \mathbb{N}} T_n^{\mathrm{c}}(s)$$

with $T_n^{\mathrm{c}}(s)$, the coherent weighted trace set of s whose traces have length at most n, being defined as:

$$\begin{cases} \{(\varepsilon,1)\} \cup \bigcup\limits_{s \xrightarrow{a} \Delta} a \,.\, \left(\sum\limits_{s' \in supp(\Delta)} \Delta(s') \cdot T_{n-1}^{\mathrm{c}}(s') \right) & \\ & \text{if } n > 0 \text{ and } s \text{ has outgoing transitions} \\ \{(\varepsilon,1)\} & \\ & \text{otherwise} \end{cases}$$ ■

For the NPLTS in Fig. 6.3 we have that $T^{\mathrm{c}}(r) = T_0^{\mathrm{c}}(r) \cup T_1^{\mathrm{c}}(r) \cup T_2^{\mathrm{c}}(r) \cup T_3^{\mathrm{c}}(r) \cup T_4^{\mathrm{c}}(r)$ where:

$$\begin{aligned} T_0^{\mathrm{c}}(r) &= \{(\varepsilon,1)\} \\ T_1^{\mathrm{c}}(r) &= \{(\varepsilon,1),(a,1)\} \\ T_2^{\mathrm{c}}(r) &= \{(\varepsilon,1),(a,1),(a\,b,1)\} \\ T_3^{\mathrm{c}}(r) &= \{(\varepsilon,1),(a,1),(a\,b,1),(a\,b\,c,1),(a\,b\,d,1)\} \\ T_4^{\mathrm{c}}(r) &= \{(\varepsilon,1),(a,1),(a\,b,1),(a\,b\,c,1),(a\,b\,d,1), \\ &\quad (a\,b\,c\,e_1,p),(a\,b\,c\,e_3,1-p),(a\,b\,d\,e_2,p),(a\,b\,d\,e_4,1-p)\} \end{aligned}$$

with the various sets $T_n^{\mathrm{c}}(r)$ precisely capturing the weighted traces of the coherent resolutions of r.

Let us investigate the properties of coherent weighted trace sets. Firstly, $T^{\mathrm{c}}(\_)$ is the flattened version of $TD^{\mathrm{mc}}(\_)$, as witnessed by the example above, but not of $TD^{\mathrm{c}}(\_)$. This can be seen by looking in Fig. 6.3 at the flattening of $TD^{\mathrm{c}}(r_1)$, which contains the additional pairs $(b\,c, p)$, $(b\,c, 1-p)$, $(b\,d, p)$, $(b\,d, 1-p)$ not occurring in $T^{\mathrm{c}}(r_1)$.

Proposition 7.1. Let $(S, A, \longrightarrow)$ be an NPLTS and $s \in S$. Then $T^{\mathrm{c}}(s) = \bigcup_{T \in TD^{\mathrm{mc}}(s)} T$.

Proof. The result will follow by proving that $T_n^{\mathrm{c}}(s) = \bigcup_{T \in TD_n^{\mathrm{mc}}(s)} T$ for all $n \in \mathbb{N}$. We proceed by induction on $n \in \mathbb{N}$:

- For $n = 0$ we have that $T_n^{\mathrm{c}}(s) = \{(\varepsilon, 1)\} = \bigcup_{T \in TD_n^{\mathrm{mc}}(s)} T$ because $TD_n^{\mathrm{mc}}(s) = \{\{(\varepsilon, 1)\}\}$.
- Let $n = m + 1$ for some $m \in \mathbb{N}$ and suppose that the result holds for the coherent weighted trace set $T_m^{\mathrm{c}}(s')$ of any state $s' \in S$. To avoid trivial cases, we assume that s has at least one outgoing transition. Then:

$$\begin{aligned} T_n^{\mathrm{c}}(s) &= \{(\varepsilon, 1)\} \cup \bigcup_{s \xrightarrow{a} \Delta} a \,.\, \left(\sum_{s' \in supp(\Delta)} \Delta(s') \cdot T_m^{\mathrm{c}}(s') \right) \\ &= \{(\varepsilon, 1)\} \cup \bigcup_{s \xrightarrow{a} \Delta} a \,.\, \left(\sum_{s' \in supp(\Delta)} \Delta(s') \cdot \bigcup_{T' \in TD_m^{\mathrm{mc}}(s')} T' \right) \\ &= \bigcup_{T \in TD_n^{\mathrm{mc}}(s)} T \end{aligned}$$

  due to the induction hypothesis, Prop. 6.3, and the fact that the probabilities of identical traces in the various sets T' of different states s' are always added up as established by Def. 6.1. □

Secondly, it is easy to characterize $T_n^{\mathrm{c}}(s)$ in the case of a fully probabilistic NPLTS. This holds in particular for resolutions.

Proposition 7.2. Let $(S, A, \longrightarrow)$ be a fully probabilistic NPLTS, $s \in S$, and $n \in \mathbb{N}$. Let $A^{\leq n} = \{\alpha \in A^* \mid |\alpha| \leq n\}$. Then $T_n^{\mathrm{c}}(s) = \{(\alpha, p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s, \alpha)) = p\}$.

Proof. We proceed by induction on $n \in \mathbb{N}$:

- For $n = 0$ we have that $T^{\mathrm{c}}_n(s) = \{(\varepsilon, 1)\} = \{(\alpha, p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s, \alpha)) = p\}$.

- Let $n = m + 1$ for some $m \in \mathbb{N}$ and suppose that the result holds for the coherent weighted trace set $T^{\mathrm{c}}_m(s')$ of any state $s' \in S$. To avoid trivial cases, we assume that s has an outgoing transition $s \xrightarrow{a} \Delta$, which is unique as the considered NPLTS is fully probabilistic. Then:
$$T^{\mathrm{c}}_n(s) \;=\; \{(\varepsilon, 1)\} \cup a \,.\, \left(\sum_{s' \in supp(\Delta)} \Delta(s') \cdot T^{\mathrm{c}}_m(s') \right)$$
From the induction hypothesis, for each $s' \in supp(\Delta)$ it follows that:
$$T^{\mathrm{c}}_m(s') \;=\; \{(\alpha', p') \in A^{\leq m} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s', \alpha')) = p'\}$$
and hence:
$$\begin{aligned} T^{\mathrm{c}}_n(s) \;&=\; \{(\varepsilon, 1)\} \cup \sum_{s' \in supp(\Delta)} \{(a\,\alpha', \Delta(s') \cdot p') \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid \\ &\qquad\qquad prob(\mathcal{CC}(s', \alpha')) = p'\} \\ &=\; \{(\alpha, p) \in A^{\leq n} \times \mathbb{R}_{]0,1]} \mid prob(\mathcal{CC}(s, \alpha)) = p\} \end{aligned}$$
□

Thirdly, the construction in Def. 7.1 turns out to be monotonic, in the sense that $T^{\mathrm{c}}_n(s)$ includes as a subset $T^{\mathrm{c}}_{n-1}(s)$.

Proposition 7.3. Let $(S, A, \longrightarrow)$ be an NPLTS, $s \in S$, $(\alpha, p) \in A^* \times \mathbb{R}_{]0,1]}$, and $n \in \mathbb{N}_{\geq |\alpha|}$. Then $(\alpha, p) \in T^{\mathrm{c}}_n(s)$ implies $(\alpha, p) \in T^{\mathrm{c}}_{n+1}(s)$.

Proof. We proceed by induction on $|\alpha| \in \mathbb{N}$:

- If $|\alpha| = 0$, i.e., $\alpha = \varepsilon$, then for all $n \in \mathbb{N}$ we have that $(\varepsilon, p) \in T^{\mathrm{c}}_n(s)$ iff $p = 1$, from which the result trivially follows because $(\varepsilon, 1) \in T^{\mathrm{c}}_n(s)$ for all $n \in \mathbb{N}$.

- Let $|\alpha| = m + 1$ for some $m \in \mathbb{N}$, with $\alpha = a\,\alpha'$ and $|\alpha'| = m$, and suppose that the result holds for each trace of length m.
If $(\alpha, p) \in T^{\mathrm{c}}_n(s)$ for some $n \geq |\alpha|$, then there exists a transition $s \xrightarrow{a} \Delta$ such that:
$$(\alpha', p) \;\in \sum_{s' \in supp(\Delta)} \Delta(s') \cdot T^{\mathrm{c}}_{n-1}(s')$$
For each $s' \in supp(\Delta)$, either α' does not occur in $T^{\mathrm{c}}_{n-1}(s')$, or α' occurs in $T^{\mathrm{c}}_{n-1}(s')$ with some probability $p_{s'} \in \mathbb{R}_{]0,1]}$ (if α'

occurs several times with different probabilities due to internal nondeterminism, $p_{s'}$ is the probability of the only occurrence that contributes to p). We denote with S' the set of states $s' \in supp(\Delta)$ such that α' occurs in $T^{\mathrm{c}}_{n-1}(s')$, where $S' \neq \emptyset$, because $p > 0$, and $\sum_{s' \in S'} \Delta(s') \cdot p_{s'} = p$, because according to the weighted trace set addition of Def. 6.1 the probabilities of weighted traces sharing the same trace – α' in our case – are always summed up.
For each $s' \in S'$, since $(\alpha', p_{s'}) \in T^{\mathrm{c}}_{n-1}(s')$ and $|\alpha'| = m$, from the induction hypothesis it follows that $(\alpha', p_{s'}) \in T^{\mathrm{c}}_{n}(s')$ too. As a consequence, it also holds that:

$$(\alpha', p) \in \sum_{s' \in supp(\Delta)} \Delta(s') \cdot T^{\mathrm{c}}_{n}(s')$$

and hence $(\alpha, p) \in T^{\mathrm{c}}_{n+1}(s)$ too. □

The following lemma, which exploits Props. 7.2 and 7.3, provides the basis for a characterization of $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ in terms of coherent weighted trace set equality. It establishes that a weighted trace (α, p) is in the coherent weighted trace set of a state s iff it can be exhibited by a coherent resolution $\mathcal{Z}$ of that state. In the lemma, z_s denotes both the initial state of $\mathcal{Z}$ and the state to which s corresponds.

Lemma 7.3. Let $(S, A, \longrightarrow)$ be an NPLTS, $s \in S$, and $(\alpha, p) \in A^* \times \mathbb{R}_{]0,1]}$. Then $(\alpha, p) \in T^{\mathrm{c}}(s)$ iff there exists $\mathcal{Z} \in Res^{\mathrm{c}}_{\mathrm{sp}}(s)$ such that $prob(\mathcal{CC}(z_s, \alpha)) = p$.

Proof. We proceed by induction on $|\alpha| \in \mathbb{N}$:

- Let $|\alpha| = 0$, i.e., $\alpha = \varepsilon$. On the one hand, we have that $(\varepsilon, p) \in T^{\mathrm{c}}(s)$ iff $p = 1$. On the other hand, for each $\mathcal{Z} \in Res^{\mathrm{c}}_{\mathrm{sp}}(s)$ it holds that $prob(\mathcal{CC}(z_s, \varepsilon)) = 1$. Therefore, the result trivially follows.

- Let $|\alpha| = m + 1$ for some $m \in \mathbb{N}$, with $\alpha = a\,\alpha'$ and $|\alpha'| = m$, and suppose that the result holds for each trace of length m. The proof is divided into two parts:

  - Let $(\alpha, p) \in T^{\mathrm{c}}(s)$. Then $(\alpha, p) \in T^{\mathrm{c}}_{n}(s)$ for $n = |\alpha|$ and hence there exists a transition $s \xrightarrow{a} \Delta$ such that [⋆]:

$$(\alpha', p) \in \sum_{s' \in supp(\Delta)} \Delta(s') \cdot T^{\mathrm{c}}_{m}(s')$$

For each $s' \in supp(\Delta)$, either α' does not occur in $T^{\mathrm{c}}_m(s')$, or α' occurs in $T^{\mathrm{c}}_m(s')$ with some probability $p_{s'} \in \mathbb{R}_{]0,1]}$ (if α' occurs several times with different probabilities due to internal nondeterminism, $p_{s'}$ is the probability of the only occurrence that contributes to p). We denote with S' the set of states $s' \in supp(\Delta)$ such that α' occurs in $T^{\mathrm{c}}_m(s')$, where $S' \neq \emptyset$, because $p > 0$, and $\sum_{s' \in S'} \Delta(s') \cdot p_{s'} = p$, because according to the weighted trace set addition of Def. 6.1 the probabilities of weighted traces sharing the same trace – α' in our case – are always summed up.

For each $s' \in S'$, since $(\alpha', p_{s'}) \in T^{\mathrm{c}}_m(s') \subseteq T^{\mathrm{c}}(s')$ and $|\alpha'| = m$, from the induction hypothesis it follows that there exists $\mathcal{Z}_{s'} = (Z_{s'}, A, \longrightarrow_{\mathcal{Z}_{s'}}) \in Res^{\mathrm{c}}_{\mathrm{sp}}(s')$ such that $prob(\mathcal{CC}(z_{s'}, \alpha')) = p_{s'}$. Without loss of generality, we can assume that each $\mathcal{Z}_{s'}$ has computations of length at most m and, subject to this, all of its computations are maximal (in the sense that no further steps can be added to reach length m) with respect to the corresponding computations from s'.

We proceed in a similar way for each $s' \in supp(\Delta) \setminus S'$, i.e., we take an arbitrary $\mathcal{Z}_{s'} = (Z_{s'}, A, \longrightarrow_{\mathcal{Z}_{s'}}) \in Res^{\mathrm{c}}_{\mathrm{sp}}(s')$ satisfying the aforementioned assumption about the length and the maximality of its computations. Moreover, for $s'_1, s'_2 \in supp(\Delta) \setminus S'$, we select $\mathcal{Z}_{s'_1}$ and $\mathcal{Z}_{s'_2}$ in such a way that $tr(TD^{\mathrm{mc}}_{m'}(s'_1)) = tr(TD^{\mathrm{mc}}_{m'}(s'_2))$ implies $tr(TD^{\mathrm{mc}}_{m'}(z_{s'_1})) = tr(TD^{\mathrm{mc}}_{m'}(z_{s'_2}))$ for all $m' \leq m$.

Consider now the resolution $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}})$ of s such that (i) the image via $corr_{\mathcal{Z}}$ of its initial transition $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$ is $s \xrightarrow{a} \Delta$ and (ii) from each $z_{s'} \in supp(\Delta')$ the resolution behaves as $\mathcal{Z}_{s'}$. Then $prob(\mathcal{CC}(z_s, \alpha)) = p$, by construction, and $\mathcal{Z} \in Res^{\mathrm{c}}_{\mathrm{sp}}(s)$, because each $\mathcal{Z}_{s'}$ is coherent and $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$ satisfies both constraints of Def. 6.4 (for an example, start from the b-transition of r_1 in Fig. 6.3).

The satisfaction of the first constraint stems from (α, p) being an element of $T^{\mathrm{c}}_n(s)$ and $\mathcal{Z}$ having computations of length at most n. Specifically, for $z'_1, z'_2 \in supp(\Delta')$ with $corr_{\mathcal{Z}}(z'_1), corr_{\mathcal{Z}}(z'_2) \in S'$ to avoid trivial cases, if

$tr(TD^{\mathrm{mc}}_m(corr_{\mathcal{Z}}(z'_1))) = tr(TD^{\mathrm{mc}}_m(corr_{\mathcal{Z}}(z'_2)))$ then we have that $tr(T^{\mathrm{c}}_m(corr_{\mathcal{Z}}(z'_1))) = tr(T^{\mathrm{c}}_m(corr_{\mathcal{Z}}(z'_2)))$. When adding up $T^{\mathrm{c}}_m(corr_{\mathcal{Z}}(z'_1))$ and $T^{\mathrm{c}}_m(corr_{\mathcal{Z}}(z'_2))$ in the summation of $[\star]$ according to Def. 6.1, every probability occurring in the former is thus summed with some probability occurring in the latter and viceversa, which guarantees by the construction of $\mathcal{Z}$ that $tr(TD^{\mathrm{mc}}_m(z'_1)) = tr(TD^{\mathrm{mc}}_m(z'_2))$. The constraint holds for any $m' < m$ too, because due to Prop. 7.3 $T^{\mathrm{c}}_n(s)$ contains as a subset $T^{\mathrm{c}}_{m'}(s)$ and hence coherency constraints on shorter traces are not forgotten.

The satisfaction of the second constraint stems from all computations of each $\mathcal{Z}_{s'}$ being maximal (i.e., no further steps can be added to reach length m) with respect to the corresponding computations from s'.

– Let $\mathcal{Z} = (Z, A, \longrightarrow_{\mathcal{Z}}) \in Res^{\mathrm{c}}_{\mathrm{sp}}(s)$ be such that $prob(\mathcal{CC}(z_s, \alpha)) = p$ with $p \in \mathbb{R}_{]0,1]}$. Then there exists a transition $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$ such that, by virtue of Prop. 7.2, it holds that:

$$p = \sum_{z' \in supp(\Delta')} \Delta'(z') \cdot prob(\mathcal{CC}(z', \alpha'))$$

For each $z' \in supp(\Delta')$, either α' is not executable from z', or there exists $p_{z'} \in \mathbb{R}_{]0,1]}$ such that $prob(\mathcal{CC}(z', \alpha')) = p_{z'}$. We denote with Z' the set of states $z' \in supp(\Delta')$ for which there exists $p_{z'} \in \mathbb{R}_{]0,1]}$ such that $prob(\mathcal{CC}(z', \alpha')) = p_{z'}$, where $Z' \neq \emptyset$, because $p > 0$, and $p = \sum_{z' \in Z'} \Delta(z') \cdot p_{z'}$.

For all $z' \in Z'$, since the complete submodel of $\mathcal{Z}$ rooted at z' is still coherent due to Prop. 6.5 and satisfies $prob(\mathcal{CC}(z', \alpha')) = p_{z'}$ with $|\alpha'| = m$, from the induction hypothesis it follows that $(\alpha', p_{z'}) \in T^{\mathrm{c}}(corr_{\mathcal{Z}}(z'))$, hence $(\alpha', p_{z'}) \in T^{\mathrm{c}}_m(corr_{\mathcal{Z}}(z'))$. For $s \xrightarrow{a} \Delta$ corresponding via $corr_{\mathcal{Z}}$ to $z_s \xrightarrow{a}_{\mathcal{Z}} \Delta'$, we thus have that:

$$(\alpha', p) \in \sum_{corr_{\mathcal{Z}}(z') \in supp(\Delta)} \Delta(corr_{\mathcal{Z}}(z')) \cdot T^{\mathrm{c}}_m(corr_{\mathcal{Z}}(z'))$$

and hence $(\alpha, p) \in T^{\mathrm{c}}_{m+1}(corr_{\mathcal{Z}}(z_s)) \subseteq T^{\mathrm{c}}(corr_{\mathcal{Z}}(z_s)) = T^{\mathrm{c}}(s)$. □

Theorem 7.4. Let $(S, A, \longrightarrow)$ be an NPLTS and $s_1, s_2 \in S$. Then $s_1 \sim^{\mathrm{pre,c}}_{\mathrm{PTr}} s_2$ iff $T^{\mathrm{c}}(s_1) = T^{\mathrm{c}}(s_2)$.

Proof. By definition, $s_1 \sim^{\mathrm{pre,c}}_{\mathrm{PTr}} s_2$ iff for all $\alpha \in A^*$ it holds that for each $\mathcal{Z}_1 \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_1)$ – resp. $\mathcal{Z}_2 \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_2)$ – there exists $\mathcal{Z}_2 \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_2)$ – resp. $\mathcal{Z}_1 \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_1)$ – such that:

$$prob(\mathcal{CC}(z_{s_1}, \alpha)) \;=\; p \;=\; prob(\mathcal{CC}(z_{s_2}, \alpha))$$

The case $p = 0$, which implies $\alpha \neq \varepsilon$, is not important because, if the considered resolution of the challenger yields $p = 0$ for α, then it can trivially be matched by the resolution of the defender containing only the initial state without any outgoing transition, as the latter cannot perform α thus yielding $p = 0$ too.

Therefore $s_1 \sim^{\mathrm{pre,c}}_{\mathrm{PTr}} s_2$ iff for all $(\alpha, p) \in A^* \times \mathbb{R}_{]0,1]}$:

$$\exists \mathcal{Z} \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_1).\, prob(\mathcal{CC}(z_{s_1}, \alpha)) = p$$
$$\iff$$
$$\exists \mathcal{Z} \in Res^{\mathrm{c}}_{\mathrm{sp}}(s_2).\, prob(\mathcal{CC}(z_{s_2}, \alpha)) = p$$

which by virtue of Lemma 7.3 amounts to for all $(\alpha, p) \in A^* \times \mathbb{R}_{]0,1]}$:

$$(\alpha, p) \in T^{\mathrm{c}}(s_1) \iff (\alpha, p) \in T^{\mathrm{c}}(s_2)$$

which in turn is equivalent to $T^{\mathrm{c}}(s_1) = T^{\mathrm{c}}(s_2)$. □

We conclude with two remarks about coherent weighted trace sets. The first is that the construction in Def. 7.1 is identical to the one in Def. 3.5 of (Bernardo *et al.*, 2014a), but this should not be the case as coherency was neglected in that paper. Indeed, before Def. 3.5 of (Bernardo *et al.*, 2014a), the definition of $X + Y$ – i.e., $T_1 + T_2$ using the notation of this monograph as of Def. 6.1 – should have included also the pairs $(\alpha, q_1) \in X$ and $(\alpha, q_2) \in Y$ without summing them up, otherwise the right-to-left implication in Lemma 3.7 of (Bernardo *et al.*, 2014a) does not hold as can be seen from trace $a\,b$ of the (incoherent) resolution in Fig. 5.1 of this monograph. That definition of $X + Y$ works here instead because the focus on coherency requires to always sum up the probabilities of weighted traces sharing the same trace.

The second remark is that looser coherency constraints based on coherent weighted trace sets – rather than on memoryfully coherent trace distributions as in Def. 6.4 – would not work. If we used T^{c}_n sets instead of TD^{mc}_n sets, then probabilistic trace equivalent NPLTS models like the ones in Fig. 7.1 would be told apart. Indeed, we would have $tr(T^{\mathrm{c}}(s'_1)) = \{\epsilon, b, b\,c_1, b\,c_2, b\,c\} = tr(T^{\mathrm{c}}(s'_2))$ – whereas $tr(TD^{\mathrm{mc}}(s'_1)) \neq tr(TD^{\mathrm{mc}}(s'_2))$ thus making the first coherency constraint trivially satisfied – hence

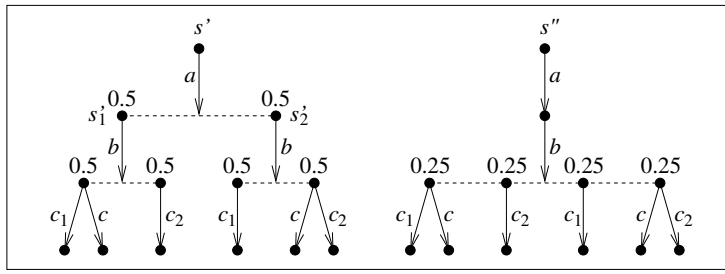

Figure 7.1: Weighted trace sets are not adequate for coherency constraints

the only coherent resolution of s' with traces of length 3 could not include c-transitions and would execute traces $a\,b\,c_1$ and $a\,b\,c_2$ only with probability 0.5, while s'' also admits coherent resolutions in which c-transitions are present and traces $a\,b\,c_1$ and $a\,b\,c_2$ have execution probability 0.25.

7.3 Parallel Composition

Alternative characterizations can be useful to investigate the properties of behavioral equivalences, in particular congruence with respect to parallel composition. For consistency with Def. 4.5, we introduce parallel composition in the style of (Brookes *et al.*, 1984):

$$\frac{s_1 \xrightarrow{a} \Delta_1 \quad a \notin L}{s_1 \parallel_L s_2 \xrightarrow{a} \Delta_1 \cdot \delta_{s_2}} \qquad \frac{s_2 \xrightarrow{a} \Delta_2 \quad a \notin L}{s_1 \parallel_L s_2 \xrightarrow{a} \delta_{s_1} \cdot \Delta_2}$$

$$\frac{s_1 \xrightarrow{a} \Delta_1 \quad s_2 \xrightarrow{a} \Delta_2 \quad a \in L}{s_1 \parallel_L s_2 \xrightarrow{a} \Delta_1 \cdot \Delta_2}$$

where $L \subseteq A$ is the set of synchronizing actions, the distribution δ_s is such that $\delta_s(s) = 1$, and $(\Delta' \cdot \Delta'')(s' \parallel_L s'') = \Delta'(s') \cdot \Delta''(s'')$.

While $\sim^{\text{post}}_{\text{PTr}}$ is a congruence with respect to parallel composition under distributed scheduling (de Alfaro *et al.*, 2001; Cheung *et al.*, 2006), this is not the case under centralized scheduling (Segala, 1995b), with the coarsest congruence contained in it turning out to be a variant of the simulation equivalence of (Segala and Lynch, 1994) as shown in (Lynch *et al.*, 2003).

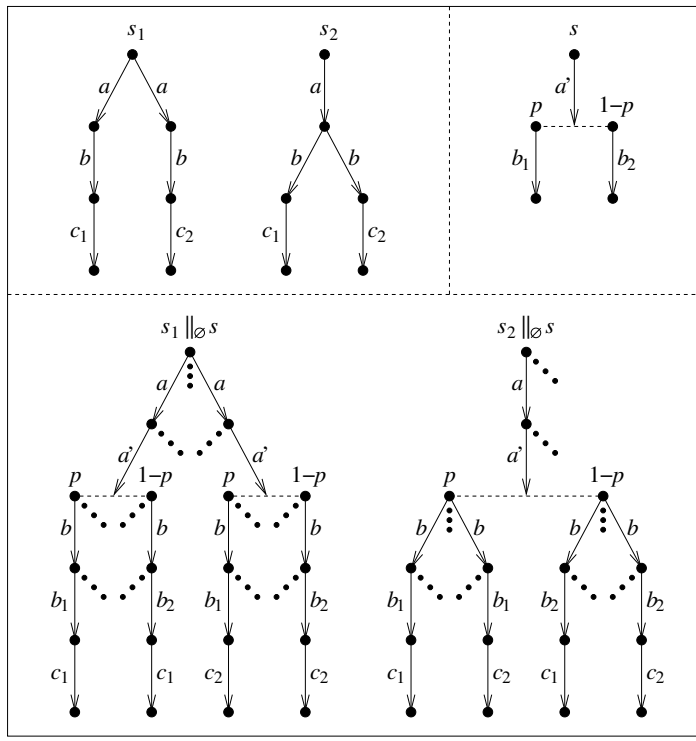

Figure 7.2: $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$ is not a congruence with respect to parallel composition

Not even $\sim_{\mathrm{PTr}}^{\mathrm{post,c}}$ is a congruence with respect to parallel composition, as can be seen in Fig. 7.2 (dots stands for transitions that are not shown). It holds that $s_1 \sim_{\mathrm{PTr}}^{\mathrm{post,c}} s_2$, but $s_1 \|_\emptyset s \not\sim_{\mathrm{PTr}}^{\mathrm{post,c}} s_2 \|_\emptyset s$ as witnessed by the maximal resolutions of $s_1 \|_\emptyset s$ and $s_2 \|_\emptyset s$ that start with trace $a\,a'$ and then continue with some of the traces in $\{b\,b_1\,c_1, b\,b_1\,c_2, b\,b_2\,c_1, b\,b_2\,c_2\}$. For instance, the resolution of $s_2 \|_\emptyset s$ whose maximal traces are $a\,a'\,b\,b_1\,c_1$ and $a\,a'\,b\,b_2\,c_2$ is not matched by any resolution of $s_1 \|_\emptyset s$. Apart from ensuring trace completeness within resolutions, coherency plays no role, in the sense that the counterexample applies to $\sim_{\mathrm{PTr}}^{\mathrm{post}}$ too.

On the other hand, $\sim_{\mathrm{PTr}}^{\mathrm{pre}}$ is a congruence with respect to parallel composition under centralized scheduling (Bernardo *et al.*, 2014a). However, $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ is not, as can be seen in Fig. 7.3. It holds that $s_1 \sim_{\mathrm{PTr}}^{\mathrm{pre,c}} s_2$, but $s_1 \|_A s \not\sim_{\mathrm{PTr}}^{\mathrm{pre,c}} s_2 \|_A s$ as witnessed by the maximal resolution of

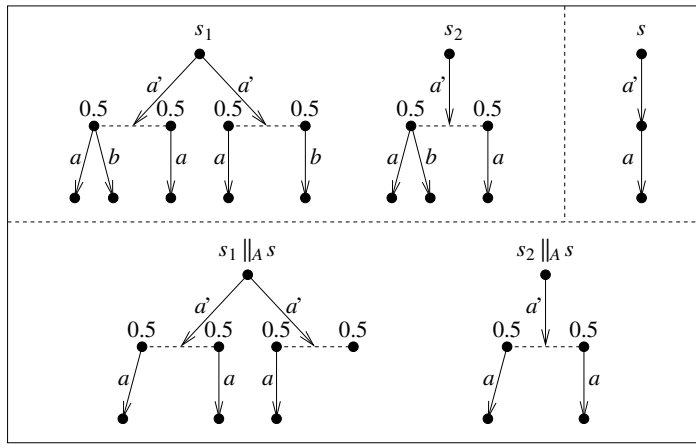

Figure 7.3: $\sim^{\text{pre,c}}_{\text{PTr}}$ is not a congruence with respect to parallel composition

$s_1 \|_A s$ starting with the rightmost a'-transition – where trace $a'\,a$ can be executed with probability 0.5 – which cannot be matched by any resolution of $s_2 \|_A s$, not even by the maximal one as trace $a'\,a$ can be executed there only with probability 1 due to coherency.

It is worth investigating this counterexample by making use of coherent weighted trace sets, where as expected $T^{\text{c}}(s_1) = T^{\text{c}}(s_2)$ and $T^{\text{c}}(s_1 \|_A s) \neq T^{\text{c}}(s_2 \|_A s)$. We have that:

- $T^{\text{c}}(s_1) = \{(\varepsilon, 1), (a', 1), (a'\,a, 0.5), (a'\,b, 0.5), (a'\,a, 1)\} = T^{\text{c}}(s_2)$.
- $T^{\text{c}}(s) = \{(\varepsilon, 1), (a', 1), (a'\,a, 1)\}$.
- $T^{\text{c}}(s_1 \|_A s) = \{(\varepsilon, 1), (a', 1), (a'\,a, 0.5), (a'\,a, 1)\}$.
- $T^{\text{c}}(s_2 \|_A s) = \{(\varepsilon, 1), (a', 1), (a'\,a, 1)\}$.

If $\sim^{\text{pre,c}}_{\text{PTr}}$ were a congruence with respect to parallel composition, thanks to Thm. 7.4 we could extend $\|_L$ to coherent weighted trace sets in such a way that $T^{\text{c}}(s' \|_L s'') = T^{\text{c}}(s') \|_L T^{\text{c}}(s'')$ with the latter being defined as:

$$\{(\alpha, p' \cdot p'') \mid (\alpha', p') \in T^{\text{c}}(s') \wedge (\alpha'', p'') \in T^{\text{c}}(s'') \wedge \alpha' \|_L \alpha'' \vdash \alpha\}$$

where parallel composition is lifted from actions to traces as follows:

$$\frac{}{\varepsilon \parallel_L \varepsilon \vdash \varepsilon} \qquad \frac{\alpha_1 \parallel_L \alpha_2 \vdash \alpha}{a\,\alpha_1 \parallel_L a\,\alpha_2 \vdash a\,\alpha}\ a \in L$$

$$\frac{\alpha_1 \parallel_L \alpha_2 \vdash \alpha}{a\,\alpha_1 \parallel_L \alpha_2 \vdash a\,\alpha}\ a \notin L \qquad \frac{\alpha_1 \parallel_L \alpha_2 \vdash \alpha}{\alpha_1 \parallel_L a\,\alpha_2 \vdash a\,\alpha}\ a \notin L$$

We note that $(a'\,a, 0.5) \in T^{\mathrm{c}}(s_2)$ and $(a'\,a, 1) \in T^{\mathrm{c}}(s)$ do not yield $(a'\,a, 0.5) \in T^{\mathrm{c}}(s_2 \parallel_A s)$ due to the impossibility of synchronizing on b that in turn triggers coherency, whilst on the side of $s_1 \parallel_A s$ it originates the maximal resolution in which a single a-transition occurs.

This example clearly indicates that $\sim_{\mathrm{PTr}}^{\mathrm{pre,c}}$ is a congruence with respect to parallel composition as long as we restrict to NPLTS models whose initial state s is such that $T^{\mathrm{c}}(s)$ contains every trace at most once, i.e., no trace can occur in $T^{\mathrm{c}}(s)$ with different probabilities.

8

Anomalies of Probabilistic Testing Equivalence

The probabilistic testing equivalence $\sim_{\text{PTe-}\sqcup\sqcap}$ of Def. 4.6 suffers from the anomaly of being only partially backward compatible. More precisely, it is compatible with the testing equivalence for fully nondeterministic processes of (De Nicola and Hennessy, 1984) – in the sense that the two equivalences coincide on those processes – only if tests are restricted to be fully nondeterministic in $\sim_{\text{PTe-}\sqcup\sqcap}$. Likewise, it is compatible with the testing equivalence for fully probabilistic processes of (Cleaveland *et al.*, 1999) only if tests are restricted to be fully probabilistic in $\sim_{\text{PTe-}\sqcup\sqcap}$.

We recall below the definition of the two testing equivalences for the two aforementioned restricted classes of processes.

Definition 8.1. Let $\mathcal{L} = (S, A, \longrightarrow_{\mathcal{L}})$ be a fully nondeterministic NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim_{\text{FNDTe}} s_2$ iff for every fully nondeterministic NPT $\mathcal{T} = (O, A, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$ it holds that:

- There exists a successful computation from (s_1, o) iff there exists a successful computation from (s_2, o) – *may testing.*
- All maximal computations from (s_1, o) are successful iff all maximal computations from (s_2, o) are successful – *must testing.* ■

Definition 8.2. Let $\mathcal{L} = (S, A, \longrightarrow_{\mathcal{L}})$ be a fully probabilistic NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim_{\text{FPTe}} s_2$ iff for every fully probabilistic NPT $\mathcal{T} = (O, A, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$ it holds that:

$$prob(\mathcal{SC}(s_1, o)) \;=\; prob(\mathcal{SC}(s_2, o))$$

■

The reason for the incomplete backward compatibility has to do with the sensitivity to the *moment of occurrence* of nondeterministic or probabilistic choices that are *internal*, i.e., among identical actions, as we now illustrate with two well known examples.

For the two fully nondeterministic NPLTS models in Fig. 8.1(A) whose initial states are s_1 and s_2, it holds that $s_1 \sim_{\text{FNDTe}} s_2$ but $s_1 \not\sim_{\text{PTe-}\sqcup\sqcap} s_2$ because the fully probabilistic NPT with initial state o tells them apart. Assuming $p \geq 1 - p$, the interaction system with initial state (s_1, o) has two maximal resolutions yielding $\sqcup\{p, 1-p\} = p$ and $\sqcap\{p, 1-p\} = 1-p$, while the interaction system with initial state (s_2, o) has four maximal resolutions yielding $\sqcup\{p, 1, 0, 1-p\} = 1$ and $\sqcap\{p, 1, 0, 1-p\} = 0$.

The synchronization of the nondeterministic choice between the two b-transitions reachable from s_2 with the probabilistic choice between the two b-transitions reachable from o creates two *copies* of state s_2' in the interaction system with initial state (s_2, o). The same internal nondeterministic choice is enabled in either copy, thereby giving the scheduler the opportunity of performing the incoherent selections that lead to the two maximal resolutions of that interaction system with initial states $z''_{s_2,o}$ and $z'''_{s_2,o}$, which respectively yield the extremal success probabilities 1 and 0.

The situation is similar in Fig. 8.1(B) with the two $\sim_{\text{FPTe}}$-equivalent fully probabilistic NPLTS models whose initial states are r_1 and r_2, which are distinguished with respect to $\sim_{\text{PTe-}\sqcup\sqcap}$ by the fully nondeterministic NPT with initial state u. This is due to the two copies of u' in the interaction system with initial state (r_2, u), in each of which the same internal nondeterministic choice is enabled. The choice is solved differently in the two maximal resolutions of that interaction system with initial states $z''_{r_2,u}$ and $z'''_{r_2,u}$, which respectively yield the extremal success probabilities 1 and 0, whereas in the interaction system with initial state (r_1, u) the extremal success probabilities are p and $1 - p$.

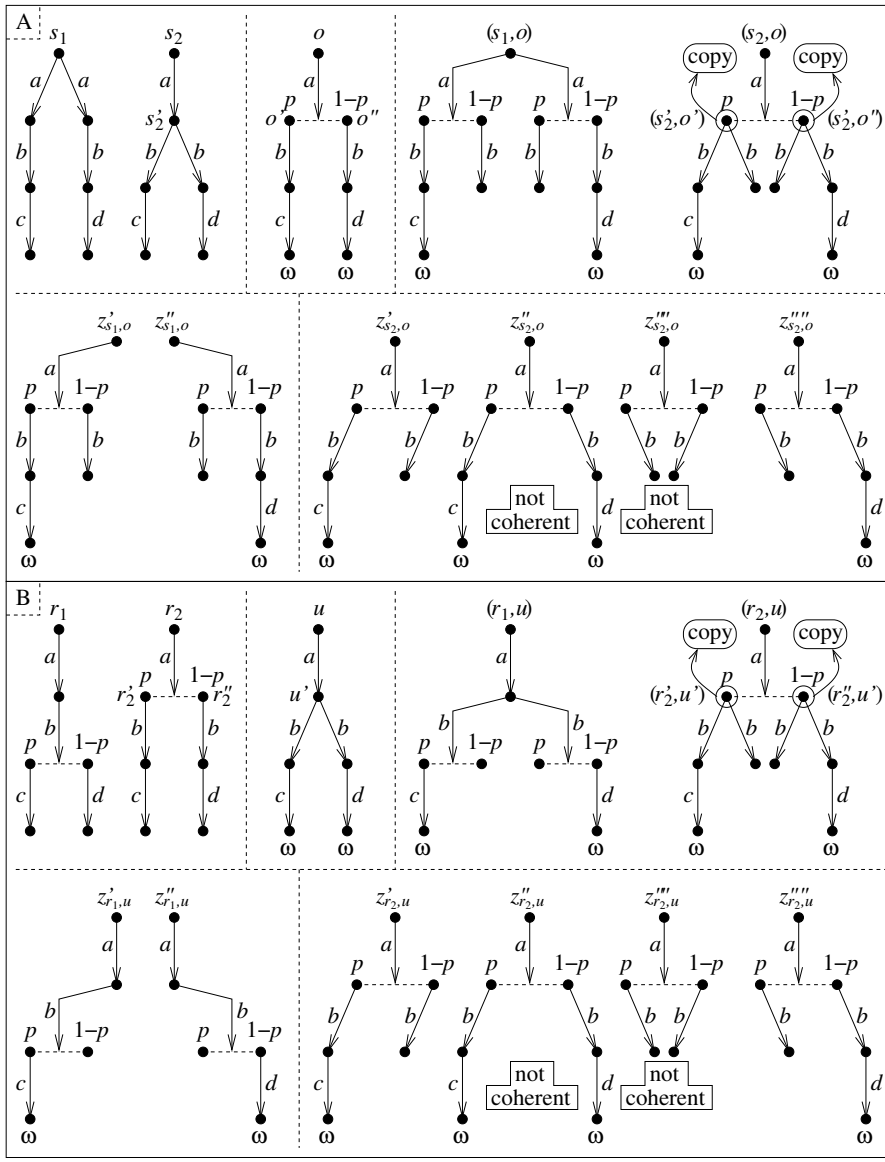

Figure 8.1: (A) Two $\sim_{\text{FNDTe}}$-equivalent fully nondeterministic NPLTS models that are $\sim_{\text{PTe-}\sqcup\sqcap}$-distinguished by a fully probabilistic test. (B) Two $\sim_{\text{FPTe}}$-equivalent fully probabilistic NPLTS models that are $\sim_{\text{PTe-}\sqcup\sqcap}$-distinguished by a fully nondeterministic test. In both cases, an internal nondeterministic choice on b synchronizes with an internal probabilistic choice on b. This originates copies of the same state in the corresponding interaction systems, as well as sensitivity to the moment of occurrence of the internal choice in the original systems under test.

9

Anomaly Avoidance via Transition Decorations

The anomalies illustrated in Fig. 8.1 are due again to the freedom of schedulers of making different decisions in states enabling the same actions. Although developed for trace semantics, we now show that the notion of coherent resolution effectively applies to testing semantics as well. Similar to (Georgievska and Andova, 2012), in addition to coherency, within maximal resolutions of interaction systems we need suitable decorations to differentiate identically labeled transitions that depart from states deriving from copies of a state of the process under test or of the test. In this way we obtain a variant $\sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap}$ of $\sim_{\text{PTe-}\sqcup\sqcap}$ possessing a higher degree of compatibility with $\sim_{\mathrm{FNDTe}}$ and $\sim_{\mathrm{FPTe}}$.

For instance, in Fig. 8.1 both states (s'_2, o') and (s'_2, o'') embody a copy of s'_2. Therefore, with respect to a scheduler, in those two states only the choice of their two left b-transitions or right b-transitions should be considered coherent, which can be achieved by decorating in the same way corresponding transitions departing from the two considered states. The situation is similar for (r'_2, u') and (r''_2, u'), with the difference that the state being copied comes from the test.

In the following we present our decoration procedure and consequently adapt all coherency-related definitions (Sect. 9.1). Then we

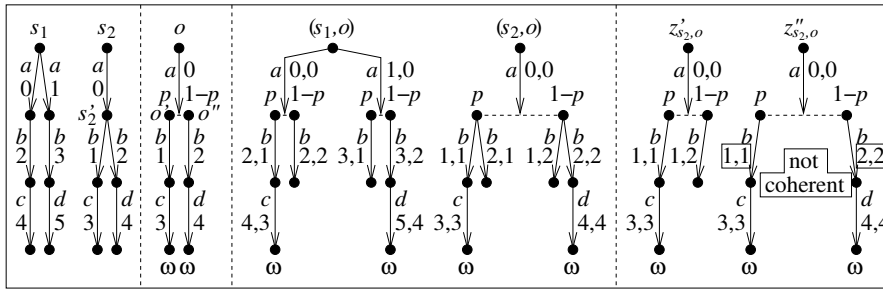

Figure 9.1: Decoration procedure based on serial numbers for transitions

illustrate to what extent the backward compatibility of probabilistic testing equivalence turns out to be enhanced (Sect. 9.2).

9.1 Decoration Procedure and Coherency

Unlike (Georgievska and Andova, 2012), our decoration procedure is very simple. The decoration of each transition of the process under test and of the test is just a serial number, then each transition of the interaction system inherits the serial numbers of the two transitions from which it is originated.

This is illustrated in Fig. 9.1. Note that in the maximal resolution whose initial state is $z'_{s_2,o}$ the two b-transitions are coherent with each other, because they both derive from the b-transition of s'_2 decorated with 1, while this is not the case in the maximal resolution whose initial state is $z''_{s_2,o}$, because its two b-transitions respectively stem from the two b-transitions of s'_2, whose decorations are 1 and 2.

To take decorations into account, which we assume to be unique within any NPLTS, we replace the action set A with $A \times \mathbb{N} \times \mathbb{N}$, where the double decoration arising from the double occurrence of $\mathbb{N}$ refers to interaction systems, and adapt the operations in Def. 6.1 accordingly. Moreover, we replace trace equality with decorated trace equivalence $\equiv$, which relates decorated traces of the same length and exhibiting the same trace whenever at each step they differ at most in one decoration. The relation $\equiv$ is then lifted to weighted trace sets and trace distributions in the expected way.

Definition 9.1. Let $B = A \times \mathbb{N} \times \mathbb{N}$. For $b \in B$, $p \in \mathbb{R}$, $TD \subseteq 2^{B^* \times \mathbb{R}}$, and $T \subseteq B^* \times \mathbb{R}$ we define:

$$\begin{aligned} b\,.\,TD &= \{b\,.\,T \mid T \in TD\} \\ b\,.\,T &= \{(b\,\beta, p') \mid (\beta, p') \in T\} \\ p \cdot TD &= \{p \cdot T \mid T \in TD\} \\ p \cdot T &= \{(\beta, p \cdot p') \mid (\beta, p') \in T\} \\ dtr(TD) &= \{dtr(T) \mid T \in TD\} \\ dtr(T) &= \{\beta \in B^* \mid \exists p' \in \mathbb{R}.\,(\beta, p') \in T\} \end{aligned}$$

while for $TD_1, TD_2 \subseteq 2^{B^* \times \mathbb{R}}$ we define:

$$TD_1 + TD_2 = \begin{cases} \{T_1 + T_2 \mid T_1 \in TD_1 \wedge T_2 \in TD_2 \wedge dtr(T_1) \equiv dtr(T_2)\} \\ \qquad\qquad \text{if } dtr(TD_1) \equiv dtr(TD_2) \\ \{T_1 + T_2 \mid T_1 \in TD_1 \wedge T_2 \in TD_2\} \\ \qquad\qquad \text{otherwise} \end{cases}$$

where for $T_1, T_2 \subseteq B^* \times \mathbb{R}$ we define:

$$\begin{aligned} T_1 + T_2 = {} & \{(\beta_1, p_1 + p_2) \mid (\beta_1, p_1) \in T_1 \wedge (\beta_2, p_2) \in T_2 \wedge \beta_1 \equiv \beta_2\} \cup \\ & \{(\beta, p) \in T_1 \mid \text{there is no } (\xi, q) \in T_2 \text{ such that } \beta \equiv \xi\} \cup \\ & \{(\beta, p) \in T_2 \mid \text{there is no } (\xi, q) \in T_1 \text{ such that } \beta \equiv \xi\} \end{aligned}$$

with:

- $\beta_1 \equiv \beta_2$ iff either $\beta_1 = \beta_2 = \varepsilon$, or $\beta_1 = \langle a, h_1, k_1 \rangle \beta_1'$, $\beta_2 = \langle a, h_2, k_2 \rangle \beta_2'$, $h_1 = h_2 \vee k_1 = k_2$, and $\beta_1' \equiv \beta_2'$.
- $dtr(T_1) \equiv dtr(T_2)$ iff for each $\beta_1 \in dtr(T_1)$ there exists $\beta_2 \in dtr(T_2)$ such that $\beta_1 \equiv \beta_2$, and vice versa.
- $dtr(TD_1) \equiv dtr(TD_2)$ iff for each $T_1 \in TD_1$ there exists $T_2 \in TD_2$ such that $dtr(T_1) \equiv dtr(T_2)$, and vice versa. ■

We then adapt Defs. 6.2 and 6.3 as follows. Note that in the adaptation of the former the first summation has to range over decorated trace distributions that are not related by $\equiv$, so as not to count equivalent distributions more than once.

Definition 9.2. Let $(S, B, \longrightarrow)$ be an NPLTS and $s \in S$. The *coherent decorated trace distribution* of s is the subset of $2^{B^* \times \mathbb{R}_{]0,1]}}$ defined as follows:

$$TD^c(s) = \bigcup_{n \in \mathbb{N}} TD^c_n(s)$$

with $TD^c_n(s)$, the coherent decorated trace distribution of s whose traces

have length at most n, being defined as:

$$\begin{cases} (\varepsilon,1) \dagger \bigcup\limits_{s \xrightarrow{a,h,k} \Delta} \langle a,h,k \rangle . \left(\sum\limits_{\Theta \in dtr(\Delta,n-1)} \sum\limits_{s' \in supp(\Delta)}^{dtr(TD^{c}_{n-1}(s')) \equiv \Theta} \Delta(s') \cdot TD^{c}_{n-1}(s') \right) \\ \qquad\qquad \text{if } n > 0 \text{ and } s \text{ has outgoing transitions} \\ \{\{(\varepsilon,1)\}\} \\ \qquad\qquad \text{otherwise} \end{cases}$$

where $dtr(\Delta, n-1)$ is the maximum subset of $\{dtr(TD^{c}_{n-1}(s')) \mid s' \in supp(\Delta)\}$ satisfying $\Theta_1 \not\equiv \Theta_2$ for all $\Theta_1, \Theta_2 \in dtr(\Delta, n-1)$ such that $\Theta_1 \neq \Theta_2$, and the operator $(\varepsilon,1) \dagger \_$ is extended in such a way that $(\varepsilon,1) \dagger TD = \{\{(\varepsilon,1)\} \cup T \mid T \in TD\}$. ■

Definition 9.3. Let $(S, B, \longrightarrow)$ be an NPLTS and $s \in S$. The *memoryfully coherent decorated trace distribution* of s is the subset of $2^{B^* \times \mathbb{R}_{]0,1]}}$ defined as follows:

$$TD^{mc}(s) = \bigcup_{n \in \mathbb{N}} TD^{mc}_{n}(s)$$

with $TD^{mc}_{n}(s)$, the memoryfully coherent decorated trace distribution of s whose traces have length at most n, being the subset of $TD^{c}_{n}(s)$ defined as:

$$\begin{cases} \{T \in TD^{c}_{n}(s) \mid \exists T' \in TD^{mc}_{n-1}(s).\, T' \subseteq T\} \\ \qquad\qquad \text{if } n > 0 \text{ and } s \text{ has outgoing transitions} \\ \{\{(\varepsilon,1)\}\} \\ \qquad\qquad \text{otherwise} \end{cases}$$ ■

Since testing semantics makes use of maximal resolutions only, the following adaptation of the first coherency constraint of Def. 6.4 suffices.

Definition 9.4. Let $\mathcal{L} = (S, B, \longrightarrow_{\mathcal{L}})$ be an NPLTS, $s \in S$, and $\mathcal{Z} = (Z, B, \longrightarrow_{\mathcal{Z}}) \in Res_{\max}(s)$ with correspondence function $corr_{\mathcal{Z}} : Z \to S$. We say that $\mathcal{Z}$ is a *coherent maximal resolution* of s, written $\mathcal{Z} \in Res^{c}_{\max}(s)$, iff for all $z \in Z$, whenever $z \xrightarrow{a,h,k}_{\mathcal{Z}} \Delta$, then for all $n \in \mathbb{N}$ and $z', z'' \in supp(\Delta)$, if $dtr(TD^{mc}_{n}(corr_{\mathcal{Z}}(z'))) \equiv dtr(TD^{mc}_{n}(corr_{\mathcal{Z}}(z'')))$ then $dtr(TD^{mc}_{n}(z')) \equiv dtr(TD^{mc}_{n}(z''))$. ■

We finally adapt Defs. 4.5 and 4.6 thereby obtaining $\sim^{c}_{\text{PTe-}\sqcup\sqcap}$.

Definition 9.5. Let $\mathcal{L} = (S, A \times \mathbb{N}, \longrightarrow_{\mathcal{L}})$ be an NPLTS and $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ be an NPT. The *decorated interaction system* of $\mathcal{L}$

and $\mathcal{T}$ is the NPLTS $\mathcal{I}(\mathcal{L},\mathcal{T}) = (S \times O, B, \longrightarrow)$ where $(s,o) \xrightarrow{a,h,k} \Delta$ iff $s \xrightarrow{a,h}_{\mathcal{L}} \Delta_1$ and $o \xrightarrow{a,k}_{\mathcal{T}} \Delta_2$ with $\Delta(s',o') = \Delta_1(s') \cdot \Delta_2(o')$ for all $(s',o') \in S \times O$. ■

Definition 9.6. Let $\mathcal{L} = (S, A \times \mathbb{N}, \longrightarrow_{\mathcal{L}})$ be an NPLTS and $s_1, s_2 \in S$. We write $s_1 \sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap} s_2$ iff for every NPT $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$ it holds that:

$$\bigsqcup_{\mathcal{Z}_1 \in Res^{\mathrm{c}}_{\max}(s_1,o)} prob(\mathcal{SC}(z_{s_1,o})) = \bigsqcup_{\mathcal{Z}_2 \in Res^{\mathrm{c}}_{\max}(s_2,o)} prob(\mathcal{SC}(z_{s_2,o}))$$

$$\bigsqcap_{\mathcal{Z}_1 \in Res^{\mathrm{c}}_{\max}(s_1,o)} prob(\mathcal{SC}(z_{s_1,o})) = \bigsqcap_{\mathcal{Z}_2 \in Res^{\mathrm{c}}_{\max}(s_2,o)} prob(\mathcal{SC}(z_{s_2,o}))$$

■

In Fig. 9.1 $dtr(TD^{\mathrm{mc}}_1(s'_2,o')) = \{\{\varepsilon, \langle b,1,1\rangle\}, \{\varepsilon, \langle b,2,1\rangle\}\}$ is identified via $\equiv$ with $dtr(TD^{\mathrm{mc}}_1(s'_2,o'')) = \{\{\varepsilon, \langle b,1,2\rangle\}, \{\varepsilon, \langle b,2,2\rangle\}\}$, hence the states to which they correspond in any coherent maximal resolution of (s_2,o) must result in an analogous identification. In contrast, $\langle b,1,1\rangle$ cannot be identified with $\langle b,2,2\rangle$ because $1 = h_1 \neq h_2 = 2$ and $1 = k_1 \neq k_2 = 2$. Likewise, $\langle b,2,1\rangle$ cannot be identified with $\langle b,1,2\rangle$ because $2 = h_1 \neq h_2 = 1$ and $1 = k_1 \neq k_2 = 2$. As a consequence, the two maximal resolutions of (s_2,o) in Fig. 8.1 respectively having initial states $z''_{s_2,o}$ and $z'''_{s_2,o}$ and success probabilities 1 and 0, with the former appearing also in Fig. 9.1 together with its decorations, are not coherent. It thus turns out that $s_1 \sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap} s_2$; for similar reasons, $r_1 \sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap} r_2$.

9.2 Limits to the Backward Compatibility of $\sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap}$

We finally prove that the joint use of coherency and decorations makes $\sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap}$ insensitive to the moment of occurrence of internal nondeterministic or probabilistic choices. Before that, we show that *full* backward compatibility with $\sim_{\text{FNDTe}}$ and $\sim_{\text{FPTe}}$ cannot be achieved, though.

Consider the two fully nondeterministic NPLTS models with initial states t_1 and t_2 in Fig. 9.2. They are known to be failure equivalent and hence identified by $\sim_{\text{FNDTe}}$ (De Nicola, 1987). We may thus expect them to be identified by $\sim^{\mathrm{c}}_{\text{PTe-}\sqcup\sqcap}$ too, but it is not the case.

This is witnessed by the fully probabilistic NPT with initial state w because of the maximal resolution of (t_2,w) with success probability 1, i.e., the one starting with the transition labeled with $\langle a,1,0\rangle$, which

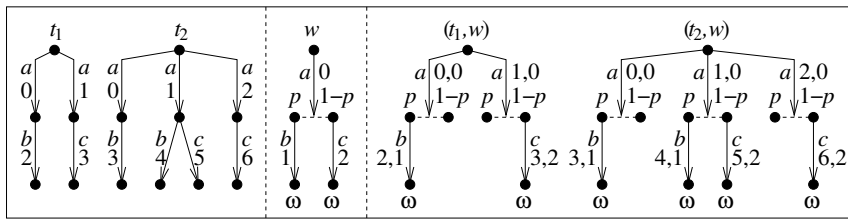

Figure 9.2: Impact of external choices on the backward compatibility of $\sim^{c}_{\text{PTe-}\sqcup\sqcap}$

is not matched by any maximal resolution of (t_1, w). In that maximal resolution the *external* nondeterministic choice between the b-transition and the c-transition after the transition of t_2 labeled with $\langle a, 1\rangle$ synchronizes with the *external* probabilistic choice between the b-transition and the c-transition departing from the two states in the support of the transition of w labeled with $\langle a, 0\rangle$. The two copies of the nondeterministic choice have been differentiated in the interaction system (t_2, w), in the sense that they enable two different sets of actions.

The backward compatibility of $\sim^{c}_{\text{PTe-}\sqcup\sqcap}$ extends till the point in which the success probability can be increased in the interaction system. This happens when the copies of an external nondeterministic choice – whose states are in the support of the same transition, like the one labeled with $\langle a, 1, 0\rangle$ in (t_2, w) – enable different sets of actions. In that case, the success probabilities of the computations starting from those states can be summed up, as the coherency constraint based on additional decorations is trivially satisfied.

This can be formalized through the following property $\mathcal{S}_{\text{ext}}$ about the syncronization of *external* choices: whenever an external nondeterministic choice of the process (resp. test) synchronizes with an external probabilistic choice of the test (resp. process), then all the states in the support of the target distribution of the resulting transition in the interaction system enable the same set of actions. Such a property mimics what naturally happens when an internal nondeterministic choice synchronizes with an external probabilistic choice.

Theorem 9.1. Let $\mathcal{L} = (S, A \times \mathbb{N}, \longrightarrow_{\mathcal{L}})$ be an NPLTS and $s_1, s_2 \in S$ and consider only NPTs $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ such that $\mathcal{I}(\mathcal{L}, \mathcal{T}) = (S \times O, B, \longrightarrow)$ meets $\mathcal{S}_{\text{ext}}$. Then:

1. $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2 \iff s_1 \sim_{\text{FNDTe}} s_2$ when $\mathcal{L}$ is fully nondeterministic.
2. $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2 \iff s_1 \sim_{\text{FPTe}} s_2$ when $\mathcal{L}$ is fully probabilistic.

Proof. Given an NPLTS $\mathcal{L} = (S, A \times \mathbb{N}, \longrightarrow_{\mathcal{L}})$ and two states $s_1, s_2 \in S$, we proceed as follows:

1. Suppose that $\mathcal{L}$ is fully nondeterministic.
   The implication $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2 \implies s_1 \sim_{\text{FNDTe}} s_2$ is straightforward. When restricting ourselves to fully nondeterministic tests, which are the only ones admitted by $\sim_{\text{FNDTe}}$, each interaction system involving $\mathcal{L}$ turns out to be fully nondeterministic too, and trivially meets $\mathcal{S}_{\text{ext}}$. As a consequence, given a fully nondeterministic NPT $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$, the maximal resolutions of $\mathcal{I}(\mathcal{L}, \mathcal{T})$ coincide with the maximal computations of $\mathcal{I}(\mathcal{L}, \mathcal{T})$ itself, hence the probability of performing a successful computation within a maximal resolution of $\mathcal{I}(\mathcal{L}, \mathcal{T})$ can only be 1 or 0. Therefore, for all $s \in S$ it holds that s may pass $\mathcal{T}$ – i.e., there exists at least one successful computation from (s, o) – iff $\bigsqcup_{\mathcal{Z} \in Res^{\text{c}}_{\max}(s,o)} prob(\mathcal{SC}(z_{s,o})) = 1$ and s must pass $\mathcal{T}$ – i.e., all maximal computations from (s, o) are successful – iff $\bigsqcap_{\mathcal{Z} \in Res^{\text{c}}_{\max}(s,o)} prob(\mathcal{SC}(z_{s,o})) = 1$. From $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2$ it thus follows that the $\sqcup$-equality constraint implies the may-part of $\sim_{\text{FNDTe}}$ and the $\sqcap$-equality constraint implies the must-part of $\sim_{\text{FNDTe}}$, hence $s_1 \sim_{\text{FNDTe}} s_2$.
   We now assume that $s_1 \sim_{\text{FNDTe}} s_2$ and, to avoid falling back into the previous case, consider an NPT $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$ that is *not* fully nondeterministic, so that it features at least one transition whose target distribution contains *several* states in its support. Suppose that $\mathcal{I}(\mathcal{L}, \mathcal{T})$ meets $\mathcal{S}_{\text{ext}}$. Thanks to the coherency constraint based on additional decorations, it holds that copies in $\mathcal{I}(\mathcal{L}, \mathcal{T})$ of internal and external nondeterministic choices in $\mathcal{L}$ are dealt with consistently in any $\mathcal{Z} \in Res^{\text{c}}_{\max}(s, o)$ for all $s \in S$:

- Let us address internal nondeterministic choices first. Distinct computations of $\mathcal{L}$ with a common initial part up to a state with an internal nondeterministic choice on some action b cannot be all involved in the generation of computations in the same resolution $\mathcal{Z}$, even in the presence of a transition in $\mathcal{T}$ whose target distribution contains in its support several states with outgoing b-transitions that can synchronize with those of the aforementioned state in $\mathcal{L}$. Due to the coherency constraint based on additional decorations, only one of the considered computations of $\mathcal{L}$ can be involved, and the continuations of those computations in $\mathcal{Z}$ (each starting with b) are all based on the continuation (starting with b as well) of the only computation of $\mathcal{L}$ involved, thereby exercising the same resolution of $\mathcal{T}$.
- This holds true also in the case of an external nondeterministic choice of $\mathcal{L}$ that, in the synchronization with a probabilistic choice of $\mathcal{T}$, yields in $\mathcal{I}(\mathcal{L},\mathcal{T})$ copies in each of which the same actions are enabled.

In conclusion, the coherency constraint based on additional decorations ensures that, as long as the fully nondeterministic NPLTS $\mathcal{L}$ features no nondeterministic choices or only nondeterministic choices each of which:

- does not synchronize with any probabilistic choice of $\mathcal{T}$;
- is internal and synchronizes with probabilistic choices of $\mathcal{T}$;
- is external (possibly with several transitions labeled with the same actions) and synchronizes with probabilistic choices of $\mathcal{T}$ in such a way that, for each synchronization, the same actions are enabled in all the copies arising from that synchronization;

every resolution $\mathcal{Z} \in \mathit{Res}^{\mathrm{c}}_{\max}(s,o)$ stems from the synchronization of a *single* computation of s labeled with some action sequence $\alpha \in A^*$ and a resolution $\mathcal{Z}' \in \mathit{Res}^{\mathrm{c}}_{\max}(o)$ in which α is executable. Thus, observing that $\mathit{prob}(\mathcal{SC}(z_{s,o})) = \sum_{\alpha \in A^*} \mathit{prob}(\mathcal{SCC}(z_{s,o},\alpha))$,

where $\mathcal{SCC}(z_{s,o}, \alpha)$ is the set of successful computations from $z_{s,o}$ compatible with α, since from the may-part of $s_1 \sim_{\text{FNDTe}} s_2$ it follows that s_1 and s_2 are trace equivalent (De Nicola and Hennessy, 1984) and probabilistic choices can only be inside $\mathcal{T}$ which is the same for both s_1 and s_2, we derive that $s_1 \sim_{\text{FNDTe}} s_2$ implies $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2$.

2. Suppose that $\mathcal{L}$ is fully probabilistic.
The implication $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2 \implies s_1 \sim_{\text{FPTe}} s_2$ is straightforward. When restricting ourselves to fully probabilistic tests, which are the only ones admitted by $\sim_{\text{FPTe}}$, each interaction system involving $\mathcal{L}$ turns out to be fully probabilistic too and trivially meets $\mathcal{S}_{\text{ext}}$. As a consequence, given a fully probabilistic NPT $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$, $\mathcal{I}(\mathcal{L}, \mathcal{T})$ has a single maximal resolution, which coincides with $\mathcal{I}(\mathcal{L}, \mathcal{T})$ itself. Therefore, for all $s \in S$ it holds that $\bigsqcup_{\mathcal{Z} \in Res^{\text{c}}_{\max}(s,o)} prob(\mathcal{SC}(z_{s,o})) = \bigsqcap_{\mathcal{Z} \in Res^{\text{c}}_{\max}(s,o)} prob(\mathcal{SC}(z_{s,o})) = prob(\mathcal{SC}(s,o))$. From $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2$ it thus follows that $s_1 \sim_{\text{FPTe}} s_2$.
We now assume that $s_1 \sim_{\text{FPTe}} s_2$ and, to avoid falling back into the previous case, consider an NPT $\mathcal{T} = (O, A \times \mathbb{N}, \longrightarrow_{\mathcal{T}})$ with initial state $o \in O$ that is *not* fully probabilistic, so that it features at least one state that has *several* outgoing transitions. Suppose that $\mathcal{I}(\mathcal{L}, \mathcal{T})$ meets $\mathcal{S}_{\text{ext}}$. The proof that from $s_1 \sim_{\text{FPTe}} s_2$ we derive $s_1 \sim^{\text{c}}_{\text{PTe-}\sqcup\sqcap} s_2$ is similar to the one of property 1 in which we started from $s_1 \sim_{\text{FNDTe}} s_2$, with the following differences:

   - The various cases related to internal/external nondeterministic choices apply to $\mathcal{T}$ instead of $\mathcal{L}$.
   - In those cases, every resolution $\mathcal{Z} \in Res^{\text{c}}_{\max}(s,o)$ stems from the synchronization of the complete submodel of $\mathcal{L}$ rooted at s and a resolution $\mathcal{Z}' \in Res^{\text{c}}_{\max}(o)$, which are both fully probabilistic.
   - We exploit the fact that from $s_1 \sim_{\text{FPTe}} s_2$ it follows that, for all $\alpha \in A^*$, s_1 and s_2 perform the action sequence α with the same probability. □

10

Conclusions

The presence of a multitude of behavioral equivalences in concurrency theory gives us the opportunity of applying the one that we consider to be the most appropriate in any specific context. In general, for fully nondeterministic processes, testing equivalence is deemed to have a balanced discriminating power, while the use of bisimulation equivalence is preferred for its proof technique, even when one is interested in trace equivalence checking. In the case of processes featuring nondeterminism and probabilities, the discriminating power of behavioral equivalences, which depends on the class of schedulers used to resolve nondeterminism, turns out to be excessive. In many cases this hampers the attainment of desirable properties such as inclusion between equivalences, compositionality with respect to typical process operators, and backward compatibility with corresponding equivalences over less expressive processes.

In this monograph, after surveying various approaches against almighty schedulers appearing in the literature as well as providing a uniform definition of structure-preserving and structure-modifying resolutions of nondeterminism, we have addressed trace and testing semantics for nondeterministic and probabilistic processes represented

as simple probabilistic automata. We have shown that the overwhelming power of centralized, memoryless schedulers can be suitably reduced by restricting ourselves to consider only coherent resolutions of nondeterminism. These have been formalized through constraints based on memoryfully coherent trace distributions and trace completeness up to a certain length, together with additional transition decorations in the case of probabilistic testing semantics.

The highlighted anomalies of probabilistic trace semantics mostly have to do with structure-preserving resolutions induced by deterministic schedulers, so one may wonder why not to avoid those schedulers altogether. The first reason is that, as shown in the spectrum of (Bernardo *et al.*, 2014b), the use of a specific family of schedulers has an impact on the discriminating power of behavioral equivalences, so there might be situations in which considering deterministic schedulers is more appropriate. The second reason is that, as witnessed by Figs. 5.3 to 6.2, some of the examined anomalies affect also equivalences defined on structure-modifying resolutions generated by randomized or interpolating schedulers. The third reason is that in more general frameworks, like the ULTraS metamodel (Bernardo, 2019b) of which simple probabilistic automata are an instance, the applicability of deterministic schedulers is always possible, while this might not be the case for other families of schedulers.

References

Abramsky, S. (1987). "Observational Equivalence as a Testing Equivalence". *Theoretical Computer Science*. 53: 225–241.

Baier, C., J.-P. Katoen, H. Hermanns, and V. Wolf. (2005). "Comparative Branching-Time Semantics for Markov Chains". *Information and Computation*. 200: 149–214.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton University Press.

Bernardo, M. (2007). "A Survey of Markovian Behavioral Equivalences". In: *Formal Methods for Performance Evaluation*. Vol. 4486. *LNCS*. Springer. 180–219.

Bernardo, M. (2018). "ULTraS at Work: Compositionality Metaresults for Bisimulation and Trace Semantics". *Journal of Logical and Algebraic Methods in Programming*. 94: 150–182.

Bernardo, M. (2019a). "Coherent Resolutions of Nondeterminism". In: *Proc. of the 16th European Performance Engineering Workshop (EPEW 2019)*. Vol. 12039. *LNCS*. Springer. 16–32.

Bernardo, M. (2019b). "Genesis and Evolution of ULTraS: Metamodel, Metaequivalences, Metaresults". In: *Models, Languages, and Tools for Concurrent and Distributed Programming*. Vol. 11665. *LNCS*. Springer. 92–111.

Bernardo, M. (2020a). “Alternative Characterizations of Probabilistic Trace Equivalences on Coherent Resolutions of Nondeterminism”. In: *Proc. of the 17th Int. Conf. on the Quantitative Evaluation of Systems (QEST 2020)*. Vol. 12289. *LNCS*. Springer. 35–53.

Bernardo, M. (2020b). “Extending Backward Compatibility of Probabilistic Testing via Coherent Resolutions”. In: *Proc. of the 21st Italian Conf. on Theoretical Computer Science (ICTCS 2020)*. Vol. 2756. CEUR-WS. 208–222.

Bernardo, M., R. De Nicola, and M. Loreti. (2014a). “Revisiting Trace and Testing Equivalences for Nondeterministic and Probabilistic Processes”. *Logical Methods in Computer Science*. 10(1:16): 1–42.

Bernardo, M., R. De Nicola, and M. Loreti. (2014b). “Relating Strong Behavioral Equivalences for Processes with Nondeterminism and Probabilities”. *Theoretical Computer Science*. 546: 63–92.

Bernardo, M., R. De Nicola, and M. Loreti. (2015). “Revisiting Bisimilarity and its Modal Logic for Nondeterministic and Probabilistic Processes”. *Acta Informatica*. 52: 61–106.

Bernardo, M., D. Sangiorgi, and V. Vignudelli. (2014c). “On the Discriminating Power of Testing Equivalences for Reactive Probabilistic Systems: Results and Open Problems”. In: *Proc. of the 11th Int. Conf. on the Quantitative Evaluation of Systems (QEST 2014)*. Vol. 8657. *LNCS*. Springer. 281–296.

Bianco, A. and L. de Alfaro. (1995). “Model Checking of Probabilistic and Nondeterministic Systems”. In: *Proc. of the 15th Int. Conf. on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 1995)*. Vol. 1026. *LNCS*. Springer. 499–513.

Bonchi, F., A. Sokolova, and V. Vignudelli. (2019). “The Theory of Traces for Systems with Nondeterminism and Probability”. In: *Proc. of the 34th ACM/IEEE Symp. on Logic in Computer Science (LICS 2019)*. IEEE-CS Press. (19:62)1–14.

Brookes, S. D., C. A. R. Hoare, and A. W. Roscoe. (1984). “A Theory of Communicating Sequential Processes”. *Journal of the ACM*. 31: 560–599.

Cheung, L., N. A. Lynch, R. Segala, and F. Vaandrager. (2006). “Switched PIOA: Parallel Composition via Distributed Scheduling”. *Theoretical Computer Science*. 365: 83–108.

Cleaveland, R., Z. Dayar, S. A. Smolka, and S. Yuen. (1999). "Testing Preorders for Probabilistic Processes". *Information and Computation*. 154: 93–148.

de Alfaro, L. (1999). "The Verification of Probabilistic Systems under Memoryless Partial-Information Policies is Hard". In: *Proc. of the 2nd Int. Workshop on Probabilistic Methods in Verification (PROBMIV 1999)*. University of Birmingham, Technical Report CSR-99-9. 19–32.

de Alfaro, L., T. A. Henzinger, and R. Jhala. (2001). "Compositional Methods for Probabilistic Systems". In: *Proc. of the 12th Int. Conf. on Concurrency Theory (CONCUR 2001)*. Vol. 2154. *LNCS*. Springer. 351–365.

De Nicola, R. (1987). "Extensional Equivalences for Transition Systems". *Acta Informatica*. 24: 211–237.

De Nicola, R. and M. Hennessy. (1984). "Testing Equivalences for Processes". *Theoretical Computer Science*. 34: 83–133.

Deng, Y., R. J. van Glabbeek, M. Hennessy, and C. Morgan. (2008). "Characterising Testing Preorders for Finite Probabilistic Processes". *Logical Methods in Computer Science*. 4(4:4): 1–33.

Deng, Y., R. J. van Glabbeek, C. Morgan, and C. Zhang. (2007). "Scalar Outcomes Suffice for Finitary Probabilistic Testing". In: *Proc. of the 16th European Symp. on Programming (ESOP 2007)*. Vol. 4421. *LNCS*. Springer. 363–378.

Derman, C. (1970). *Finite State Markovian Decision Processes*. Academic Press.

Georgievska, S. and S. Andova. (2012). "Probabilistic May/Must Testing: Retaining Probabilities by Restricted Schedulers". *Formal Aspects of Computing*. 24: 727–748.

Giro, S. and P. R. D'Argenio. (2007). "Quantitative Model Checking Revisited: Neither Decidable nor Approximable". In: *Proc. of the 5th Int. Conf. on Formal Modeling and Analysis of Timed Systems (FORMATS 2007)*. Vol. 4763. *LNCS*. Springer. 179–194.

Giro, S. and P. R. D'Argenio. (2009). "On the Expressive Power of Schedulers in Distributed Probabilistic Systems". In: *Proc. of the 7th Int. Workshop on Quantitative Aspects of Programming Languages (QAPL 2009)*. Vol. 253(3). *ENTCS*. Elsevier. 45–71.

Hansson, H. and B. Jonsson. (1990). "A Calculus for Communicating Systems with Time and Probabilities". In: *Proc. of the 11th IEEE Real-Time Systems Symp. (RTSS 1990)*. IEEE-CS Press. 278–287.

Huynh, D. T. and L. Tian. (1992). "On Some Equivalence Relations for Probabilistic Processes". *Fundamenta Informaticae*. 17: 211–234.

Jifeng, H., K. Seidel, and A. McIver. (1997). "Probabilistic Models for the Guarded Command Language". *Science of Computer Programming*. 28: 171–192.

Jonsson, B., C. Ho-Stuart, and W. Yi. (1994). "Testing and Refinement for Nondeterministic and Probabilistic Processes". In: *Proc. of the 3rd Int. Symp. on Formal Techniques in Real Time and Fault Tolerant Systems (FTRTFT 1994)*. Vol. 863. *LNCS*. Springer. 418–430.

Jonsson, B. and W. Yi. (1995). "Compositional Testing Preorders for Probabilistic Processes". In: *Proc. of the 10th IEEE Symp. on Logic in Computer Science (LICS 1995)*. IEEE-CS Press. 431–441.

Jonsson, B. and W. Yi. (2002). "Testing Preorders for Probabilistic Processes Can Be Characterized by Simulations". *Theoretical Computer Science*. 282: 33–51.

Jou, C.-C. and S. A. Smolka. (1990). "Equivalences, Congruences, and Complete Axiomatizations for Probabilistic Processes". In: *Proc. of the 1st Int. Conf. on Concurrency Theory (CONCUR 1990)*. Vol. 458. *LNCS*. Springer. 367–383.

Keller, R. M. (1976). "Formal Verification of Parallel Programs". *Communications of the ACM*. 19: 371–384.

Kemeny, J. G. and J. L. Snell. (1960). *Finite Markov Chains*. Van Nostrand.

Larsen, K. G. and A. Skou. (1991). "Bisimulation Through Probabilistic Testing". *Information and Computation*. 94: 1–28.

Lynch, N. A., R. Segala, and F. Vaandrager. (2003). "Compositionality for Probabilistic Automata". In: *Proc. of the 14th Int. Conf. on Concurrency Theory (CONCUR 2003)*. Vol. 2761. *LNCS*. Springer. 208–221.

Milner, R. (1989). *Communication and Concurrency*. Prentice Hall.

Park, D. (1981). "Concurrency and Automata on Infinite Sequences". In: *Proc. of the 5th GI Conf. on Theoretical Computer Science*. Vol. 104. *LNCS*. Springer. 167–183.

Rabin, M. O. (1963). "Probabilistic Automata". *Information and Control*. 6: 230–245.

Segala, R. (1995a). *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD Thesis.

Segala, R. (1995b). "A Compositional Trace-Based Semantics for Probabilistic Automata". In: *Proc. of the 6th Int. Conf. on Concurrency Theory (CONCUR 1995)*. Vol. 962. *LNCS*. Springer. 234–248.

Segala, R. (1996). "Testing Probabilistic Automata". In: *Proc. of the 7th Int. Conf. on Concurrency Theory (CONCUR 1996)*. Vol. 1119. *LNCS*. Springer. 299–314.

Segala, R. and N. A. Lynch. (1994). "Probabilistic Simulations for Probabilistic Processes". In: *Proc. of the 5th Int. Conf. on Concurrency Theory (CONCUR 1994)*. Vol. 836. *LNCS*. Springer. 481–496.

Sokolova, A. and E. P. de Vink. (2004). "Probabilistic Automata: System Types, Parallel Composition and Comparison". In: *Validation of Stochastic Systems*. Vol. 2925. *LNCS*. Springer. 1–43.

Song, L., L. Zhang, J. C. Godskesen, and F. Nielson. (2013). "Bisimulations Meet PCTL Equivalences for Probabilistic Automata". *Logical Methods in Computer Science*. 9(2:7): 1–34.

Tracol, M., J. Desharnais, and A. Zhioua. (2011). "Computing Distances Between Probabilistic Automata". In: *Proc. of the 9th Int. Workshop on Quantitative Aspects of Programming Languages (QAPL 2011)*. Vol. 57. *EPTCS*. 148–162.

van Glabbeek, R. J. (2001). "The Linear Time – Branching Time Spectrum I". In: *Handbook of Process Algebra*. Elsevier. 3–99.

van Glabbeek, R. J., S. A. Smolka, and B. Steffen. (1995). "Reactive, Generative and Stratified Models of Probabilistic Processes". *Information and Computation*. 121: 59–80.

Vardi, M. Y. (1985). "Automatic Verification of Probabilistic Concurrent Finite-State Programs". In: *Proc. of the 26th IEEE Symp. on Foundations of Computer Science (FOCS 1985)*. IEEE-CS Press. 327–338.

Yi, W. and K. G. Larsen. (1992). "Testing Probabilistic and Nondeterministic Processes". In: *Proc. of the 12th Int. Symp. on Protocol Specification, Testing and Verification (PSTV 1992)*. North-Holland. 47–61.