Constructive Logical Characterizations of Bisimilarity for Reactive Probabilistic Systems

Marco Bernardo[a], Marino Miculan[b]

[a]Dipartimento di Scienze Pure e Applicate, Università di Urbino, Italy
[b]Dipartimento di Scienze Matematiche, Informatiche e Fisiche, Università di Udine, Italy

Abstract

Larsen and Skou characterized bisimilarity over reactive probabilistic systems with a logic including negation, conjunction, and a diamond modality decorated with a probabilistic lower bound. Later on, working on continuous state spaces, Desharnais, Edalat, and Panangaden proved that negation is not necessary to characterize the same equivalence. In this paper, we redemonstrate the former result with a simpler proof and the latter result directly on discrete state spaces without resorting to measure-theoretic arguments. Moreover, we show that conjunction can be replaced by disjunction in both logics, still characterizing the same bisimilarity. To these ends, we introduce *reactive probabilistic trees*, a fully abstract model for reactive probabilistic systems that allows us to uniformly prove expressiveness of the four probabilistic modal logics by means of a compactness argument. Our proofs are constructive, as they induce for each considered logic an algorithm that builds a distinguishing formula in case of two inequivalent reactive probabilistic systems.

1. Introduction

The automata-based representation of *probabilistic systems* uses probability distributions to describe which action is executed at a certain state and/or which state is reached after executing that action. There are several families of probabilistic system models depending on the degree of nondeterminism allowed by the models themselves. In the model of *generative probabilistic systems* [15], which is an action-based extension of Markov chains [21], there is no nondeterminism, as the choice among the actions enabled by a state is fully probabilistic. In the model of *reactive probabilistic systems* [15], also known as Rabin probabilistic automata [26] or Markov decision processes [10], only external nondeterminism is allowed; every state has at most one outgoing transition labeled with a certain action, with the choice among transitions being nondeterministic and each transition leading to a probability distribution over states. In the model of *nondeterministic and probabilistic systems*, also known as Vardi concurrent Markov chains [29] or Hansson alternating models [16] or Segala probabilistic automata [28], internal nondeterminism is allowed too, which means that a state may have several outgoing transitions labeled with the same action.

A semantic notion central to all the afore-mentioned system families is that of *probabilistic bisimilarity.* This was originally introduced by Larsen and Skou on reactive probabilistic systems [22] and corresponds to Milner strong bisimilarity for fully nondeterministic systems [24] as well as ordinary lumpability for Markov chains [21]. In essence, two probabilistically bisimilar states are able to mimic each other behavior by stepwise performing the same actions with the same probabilities. In their seminal paper, Larsen and Skou provided a logical characterization of probabilistic bisimilarity based on a propositional modal logic similar to Hennessy-Milner one [17]. The syntax of this probabilistic modal logic as subsequently redefined in [23] features the usual constructs $\top$, $\neg$, and $\wedge$ together with a diamond modality $\langle a \rangle_p \phi$, which is satisfied by a state if, after performing action a, the probability of being in a state satisfying ϕ is at least p. This result is important both from a foundational viewpoint, as it allows us to understand the logical complexity of probabilistic bisimilarity, and from an applicative viewpoint, as it can be exploited to provide diagnostic information in case of inequivalence.

Larsen and Skou proved their logical characterization result under the *minimal deviation assumption*, i.e., for reactive probabilistic systems such that the probability associated with any state in the support of the target distribution of a transition is a multiple of some value, which implies that the support contains finitely many states. The result was subsequently extended by Desharnais, Edalat, and Panangaden to encompass discrete and continuous state spaces [11]. They generalized the definition of probabilistic bisimilarity to reactive probabilistic systems whose state space is an *analytic space*, in terms of spans of zig-zag morphisms of the category of labeled Markov processes. Then they showed that neither finite branching assumptions on the model nor infinitary conjunctions in the logic are needed and, most surprisingly, that negation is *not* necessary to characterize probabilistic bisimilarity on reactive probabilistic systems. The elimination of negation has no counterpart in the nonprobabilistic setting, where Hennessy-Milner logic without negation characterizes *simulation equivalence*, which is strictly coarser than bisimilarity [14]. However, it is consistent with the fact that bisimilarity and simulation equivalence coincide on reactive probabilistic systems [2].

Desharnais, Edalat, and Panangaden also provided in [11] a proof of the logical characterization result without negation for finite-state reactive probabilistic systems, in which the use of measure theory is kept to a minimum in the sense that only the π-λ theorem of Dynkin [5] is employed. A further simplification of the proof for discrete-state reactive probabilistic systems, which however makes use again of the measure-theoretic result mentioned before, was later developed by Deng and Wu [9]. More recently, in [13] Fijalkow, Klin, and Panangaden simplified the proof of the logical characterization result and further extended its validity to the case of uncountably many labels under the assumption that the transition structure be continuous instead of just measurable, thanks to a new game-theoretic way of understanding bisimilarity. On the other hand, Danos, Desharnais, Laviolette, and Panangaden extended the result to general measure spaces by introducing event bisimilarity in terms of cospans of morphisms [8]. The logical characterization result without negation on general measure spaces was redemonstrated – together with a number of other results – by Jacobs and Sokolova in the coalgebraic framework of dual adjunctions between spaces and logics [18].

In this paper, we provide several contributions about the logical characterization of bisimilarity over reactive probabilistic systems with discrete states spaces. Firstly, we redemonstrate the result of Larsen and Skou with a simpler proof that relaxes the minimal deviation assumption to image finiteness. Secondly, we redemonstrate the result of Desharnais, Edalat, and Panangaden directly on discrete state spaces without resorting to measure-theoretic arguments; this was shown to be possible for the first time by Worrell in an unpublished note cited in [25]. Thirdly, we prove that conjunction can be replaced by *disjunction* in both logics of the previous results, still characterizing the same bisimilarity. While this is obvious in the presence of negation, it is surprising that the constructs $\top$, $\vee$, and $\langle a \rangle_p$ suffice to characterize probabilistic bisimilarity on reactive probabilistic systems. The intuition is that from a conjunctive distinguishing formula we can often derive a disjunctive one by suitably increasing some probabilistic lower bounds. Not even this result has a counterpart in the nonprobabilistic setting, where replacing conjunction with disjunction in the absence of negation yields trace equivalence, which by the way is different from bisimilarity also on reactive probabilistic systems.

Our proofs of the four logical characterization results are *uniform*, in the sense that they share the same technique. First of all, using a simple categorical construction, we show that each reactive probabilistic system can be given a semantics in a precise canonical form, which we call *reactive probabilistic tree*. These trees can be seen as the probabilistic counterpart of Winskel synchronization trees used for nondeterministic systems. The semantics is *fully abstract*, i.e., two states are probabilistically bisimilar if and only if they are mapped to the same reactive probabilistic tree. Moreover, the semantics is *compact*, in the sense that two (possibly infinite) trees are equal if and only if all of their finite approximations are equal. Hence, in order to prove that one of the four logics characterizes probabilistic bisimilarity, it suffices to prove that it allows to discriminate *finite* reactive probabilistic trees. Indeed, given two different finite trees, we show how to construct (by induction on the height of one of the trees) a *distinguishing formula* of the logic that tells the two trees apart and has a depth not exceeding the height of the two trees. As a consequence, our proofs of the four logical characterization results are *constructive* too, thereby naturally leading to the definition of algorithms in the style of [6] that explain the inequivalence of two reactive probabilistic systems. We point out that our technique can be used in any computational model that has a compact, fully abstract semantics.

This paper, which is an extended and revised version of [3], is organized as follows. In Sect. 2, we recall the basic definitions about reactive probabilistic systems, bisimilarity, and logics. In Sect. 3, we characterize probabilistic bisimilarity in terms of finite reactive probabilistic trees. In Sect. 4, we show that the four probabilistic modal logics considered in the paper can discriminate these finite trees – and hence characterize probabilistic bisimilarity – together with their related algorithms for building distinguishing formulas. Conclusions and directions for future work are in Sect. 5.

2. Transition Systems, Probabilistic Bisimilarity, Modal Logics

In this section, we recall some background notions related to the labeled transition system representation of reactive probabilistic systems, the definition of strong probabilistic bisimilarity for those systems, and syntax and semantics of the four probabilistic modal logics considered in the paper.

2.1. Reactive Probabilistic Labeled Transition Systems

Probabilistic processes can be represented as labeled transitions systems [20] enriched with probabilistic information used to determine which action is executed or which state is reached. Following the terminology of [15], we focus on *reactive* probabilistic processes, where every state has for each action at most one outgoing distribution over states; the choice among these arbitrarily many, differently labeled distributions is nondeterministic. For a countable (i.e., finite or countably infinite) set X, the set of finitely supported (a.k.a. simple) probability distributions over X is given by:

$$D(X) \triangleq \{\Delta : X \to \mathbb{R}_{[0,1]} \mid |\mathit{supp}(\Delta)| < \omega, \sum_{x \in X} \Delta(x) = 1\} \tag{1}$$

where the *support* of a distribution Δ is defined as $\mathit{supp}(\Delta) \triangleq \{x \in X \mid \Delta(x) > 0\}$. We restrict ourselves to finitely supported distributions, i.e., image-finite transition systems, in order to guarantee the existence of the final coalgebra in the construction that will be developed in Sect. 3.

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

- S is a countable set of *states*;
- A is a countable set of *actions*;
- $\longrightarrow \subseteq S \times A \times D(S)$ is a *transition relation* such that $(s, a, \Delta_1), (s, a, \Delta_2) \in \longrightarrow$ implies $\Delta_1 = \Delta_2$. ■

An RPLTS can be seen as a directed graph whose edges are labeled by pairs $(a, p) \in A \times \mathbb{R}_{]0,1]}$. For every $s \in S$ and $a \in A$, if there are a-labeled edges outgoing from s, then these are finitely many (image finiteness) because the considered distributions are finitely supported and the numbers on them add up to 1. As usual, we denote $(s, a, \Delta) \in \longrightarrow$ as $s \xrightarrow{a} \Delta$, where the set of reachable states coincides with $\mathit{supp}(\Delta)$. We also define cumulative reachability as $\Delta(S') = \sum_{s' \in S'} \Delta(s')$ for all $S' \subseteq S$.

2.2. Strong Probabilistic Bisimilarity

Probabilistic bisimilarity for the RPLTS model was introduced by Larsen and Skou [22] in terms of equivalence relations that capture the ability of two states to mimic each other behavior by stepwise performing the same actions with the same probabilities.

Definition 2.2. Let $(S, A, \longrightarrow)$ be an RPLTS. An equivalence relation $\mathcal{B}$ over S is a *probabilistic bisimulation* iff, whenever $(s_1, s_2) \in \mathcal{B}$, then for all actions $a \in A$:

- if $s_1 \xrightarrow{a} \Delta_1$, then $s_2 \xrightarrow{a} \Delta_2$ and $\Delta_1(C) = \Delta_2(C)$ for all equivalence classes $C \in S/\mathcal{B}$;
- if $s_2 \xrightarrow{a} \Delta_2$, then $s_1 \xrightarrow{a} \Delta_1$ and $\Delta_1(C) = \Delta_2(C)$ for all equivalence classes $C \in S/\mathcal{B}$.

We say that $s_1, s_2 \in S$ are *probabilistically bisimilar*, written $s_1 \sim_{\mathrm{PB}} s_2$, iff there exists a probabilistic bisimulation including the pair (s_1, s_2). ■

2.3. Probabilistic Modal Logics

In our setting, a probabilistic modal logic is a pair formed by a set $\mathcal{L}$ of *formulas* and an RPLTS-indexed family of *satisfaction relations* $\models \subseteq S \times \mathcal{L}$. The *logical equivalence* induced by $\mathcal{L}$ over S is defined by letting $s_1 \cong_{\mathcal{L}} s_2$, where $s_1, s_2 \in S$, iff $s_1 \models \phi \iff s_2 \models \phi$ for all $\phi \in \mathcal{L}$. We say that $\mathcal{L}$ *characterizes* a binary relation $\mathcal{R}$ over S when $\mathcal{R} = \cong_{\mathcal{L}}$.

We are especially interested in probabilistic modal logics characterizing $\sim_{\mathrm{PB}}$. The logics considered in this paper are similar to Hennessy-Milner logic [17], but the diamond modality is decorated with a probabilistic lower bound as follows:

$$\begin{array}{ll}
\mathrm{PML}_{\neg\wedge}: & \phi ::= \top \mid \neg\phi \mid \phi \wedge \phi \mid \langle a \rangle_p \phi \\
\mathrm{PML}_{\neg\vee}: & \phi ::= \top \mid \neg\phi \mid \phi \vee \phi \mid \langle a \rangle_p \phi \\
\mathrm{PML}_{\wedge}: & \phi ::= \top \mid \phi \wedge \phi \mid \langle a \rangle_p \phi \\
\mathrm{PML}_{\vee}: & \phi ::= \top \mid \phi \vee \phi \mid \langle a \rangle_p \phi
\end{array}$$

where $p \in \mathbb{R}_{[0,1]}$; trailing $\top$'s will be omitted for sake of readability. Their semantics with respect to an RPLTS state s is defined as follows:

$$\begin{array}{rcl}
s \models \top & \iff & \text{true} \\
s \models \neg\phi & \iff & s \not\models \phi \\
s \models \phi_1 \wedge \phi_2 & \iff & s \models \phi_1 \text{ and } s \models \phi_2 \\
s \models \phi_1 \vee \phi_2 & \iff & s \models \phi_1 \text{ or } s \models \phi_2 \\
s \models \langle a \rangle_p \phi & \iff & s \xrightarrow{a} \Delta \text{ and } \Delta(\{s' \in S \mid s' \models \phi\}) \geq p
\end{array}$$

Larsen and Skou [22] proved that $\mathrm{PML}_{\neg\wedge}$ characterizes $\sim_{\mathrm{PB}}$. This holds also for $\mathrm{PML}_{\neg\vee}$ because $\mathrm{PML}_{\neg\vee}$ is equivalent to $\mathrm{PML}_{\neg\wedge}$. Desharnais, Edalat, and Panangaden [11] then proved that $\mathrm{PML}_{\wedge}$ characterizes $\sim_{\mathrm{PB}}$ too, meaning that negation is not necessary for discrimination purposes. We will see in Sect. 4 that $\mathrm{PML}_{\vee}$ suffices as well, i.e., conjunction can be replaced by disjunction even in the absence of negation.

3. Compact Characterization of Probabilistic Bisimilarity

In this section, we provide a characterization of probabilistic bisimilarity by means of *finite* structures in a canonical form. To this end, we introduce *reactive probabilistic trees*, a concrete representation of probabilistic behaviors that we will exploit in our logical characterization proofs.

3.1. Coalgebras for Probabilistic Systems

We begin by recalling the coalgebraic setting for probabilistic systems; see, e.g., [30]. The function D defined in (1) extends to a functor $D : \mathsf{Set} \to \mathsf{Set}$ whose action on morphisms is, for $f : X \to Y$:

$$D(f) : D(X) \to D(Y) \qquad D(f)(\Delta) \; = \; \lambda y.\Delta(f^{-1}(y))$$

Then, it is easy to see that every RPLTS corresponds to a coalgebra of the following functor:

$$B_{RP} : \mathsf{Set} \to \mathsf{Set} \qquad B_{RP}(X) \; = \; (D(X) + 1)^A$$

Indeed, given $S = (S, A, \longrightarrow)$, we define the corresponding coalgebra (S, σ) as

$$\sigma : S \to B_{RP}(S) \qquad \sigma(s) \triangleq \lambda a. \begin{cases} \Delta & \text{if } s \xrightarrow{a} \Delta \\ * & \text{otherwise} \end{cases}$$

A *homomorphism* $h : (S, \sigma) \to (T, \tau)$ is a function $h : S \to T$ which respects the coalgebraic structures, i.e., $\tau \circ h = (B_{RP} h) \circ \sigma$. We denote by *Coalg*$(B_{RP})$ the category of B_{RP}-coalgebras and their homomorphisms.

Aczel and Mendler [1] introduced a general notion of bisimulation for coalgebras, which in our setting instantiates as follows:

Definition 3.1. Let (S_1, σ_1) and (S_2, σ_2) be B_{RP}-coalgebras. A relation $\mathcal{R} \subseteq S_1 \times S_2$ is a *B_{RP}-bisimulation* iff there exists a coalgebra structure $\rho : \mathcal{R} \to B_{RP}\mathcal{R}$ such that the projections $\pi_1 : \mathcal{R} \to S_1$ and $\pi_2 : \mathcal{R} \to S_2$ are homomorphisms, i.e., $\sigma_i \circ \pi_i = B_{RP}\pi_i \circ \rho$ for $i = 1, 2$.

We say that $s_1 \in S_1$ and $s_2 \in S_2$ are *B_{RP}-bisimilar*, written $s_1 \sim s_2$, iff there exists a B_{RP}-bisimulation including (s_1, s_2). ■

The following result shows that probabilistic bisimilarity corresponds to B_{RP}-bisimilarity, and it is an immediate consequence of [30, Lemma 4.4 and Thm. 4.5].

Proposition 3.2. Probabilistic bisimilarity over an RPLTS $(S, A, \longrightarrow)$ coincides with B_{RP}-bisimilarity over the corresponding coalgebra (S, σ).

The next step is to associate each state of a given RPLTS with its *behavior*, i.e., a structure in some canonical form which we can reason about. These structures can be seen as the elements of the final coalgebra of B_{RP}, which exists because we consider only finitely supported distributions, as proved in [30, Thm. 4.6]:

Proposition 3.3. The functor B_{RP} admits final coalgebra.

Proof The functor D is bounded because it is restricted to distributions with finite support. Hence also B_{RP} is bounded ([30, Thm. 4.6]); then the final coalgebra exists by the general result [27, Thm. 10.4]. ■

Let (Z, ζ) be the final B_{RP}-coalgebra (which is unique up-to isomorphism). This coalgebra can be seen as the RPLTS which subsumes all possible behaviors of any RPLTS. Moreover, elements of Z can be seen as "canonical" representatives of behaviors, because different states of Z are never bisimilar as established by the following extensionality result:

Proposition 3.4. For all $z_1, z_2 \in Z$: $z_1 \sim z_2$ iff $z_1 = z_2$.

Proof The "if" direction is trivial. For the "only if" direction, let us suppose that there are $z_1, z_2 \in Z$ such that $z_1 \sim z_2$ but $z_1 \neq z_2$. Then we can consider the coalgebra $(Z \setminus \{z_1\}, \{z_1 \mapsto z_2\} \circ \zeta)$: this coalgebra has the universal property of final coalgebras (any coalgebra has a unique homomorphism to it), but it is not isomorphic to (Z, ζ), which is absurd. ■

3.2. Reactive Probabilistic Trees

Although Prop. 3.3 guarantees the existence of the final coalgebra, it does not provide us with a concrete representation of its elements. In this subsection, we introduce *reactive probabilistic trees*, a representation of the final B_{RP}-coalgebra which can be seen as the natural extension to the probabilistic setting of *strongly extensional trees* used to represent the final $\mathcal{P}_f$-coalgebra [31].

Reactive probabilistic trees are unordered trees where each node has for each action either no successor, or a finite set of successors labeled with positive real numbers adding up to 1; moreover, subtrees rooted at these successors are all different.

Definition 3.5. An *(A-labeled) reactive probabilistic tree* is a pair (X, succ) where $X \in \mathsf{Set}$ and $\mathit{succ} : X \times A \to \mathcal{P}_f(X \times \mathbb{R}_{(0,1]})$ are such that the relation $\leq$ over X defined by:

$$\frac{}{x \leq x} \qquad \frac{x \leq y \quad z \in \mathit{succ}(y, a)}{x \leq z}$$

is a partial order with a least element, called *root*, and for all $x \in X$ and $a \in A$ it holds that:

1. the set $\{y \in X \mid y \leq x\}$ is finite and well ordered;
2. for all $(x_1, p_1), (x_2, p_2) \in \mathit{succ}(x, a)$, if $x_1 = x_2$ then $p_1 = p_2$;
3. for all $(x_1, p_1), (x_2, p_2) \in \mathit{succ}(x, a)$, if the subtrees rooted at x_1 and x_2 are isomorphic then $x_1 = x_2$;
4. if $\mathit{succ}(x, a) \neq \emptyset$ then $\sum_{(y,p) \in \mathit{succ}(x,a)} p = 1$.

We denote by RPT, ranged over by $t, t_1, t_2, \ldots$, the set of all reactive probabilistic trees (possibly of infinite height), up-to isomorphism. ■

The trivial tree is $nil \triangleq (\{\bot\}, \lambda x, a.\emptyset)$. For $t = (X, succ)$, we denote its root by $\bot_t$, its a-successors by $t(a) \triangleq succ(\bot_t, a)$, and the subtree rooted at $x \in X$ by $t[x] \triangleq (\{y \in X \mid x \leq y\}, \lambda y, a.succ(y, a))$; thus, $\bot_{t[x]} = x$. See the forthcoming Fig. 1 for some examples.

We define $height : RPT \to \mathbb{N} \cup \{\omega\}$ in the obvious way:

$$height(t) \triangleq \sup\{1 + height(t') \mid (t', p) \in t(a), a \in A\}$$

where $\sup \emptyset = 0$; hence, $height(nil) = 0$. In particular, we denote by $RPT_f \triangleq \{t \in RPT \mid height(t) < \omega\}$ the set of reactive probabilistic trees of finite height.

A (possibly infinite) tree can be truncated at any height n, yielding a finite tree where the missing subtrees are replaced by *nil*. In order to obtain a tree in RPT_f, we need to collapse isomorphic subtrees resulting from the truncation. More formally, we define first the *truncation* function tr_n by induction on n:

$$\begin{aligned}
tr_0(t) &\triangleq nil \\
tr_{n+1}(t) &\triangleq \Big(\{\bot_t\} \cup \bigcup\{X' \mid ((X', succ'), p') \in tr_n(t(a)), a \in A\}, succ_Y\Big) \\
\text{where } succ_Y(\bot_t, a) &\triangleq \{(\bot_{t'}, p') \mid (t', p') \in q(t(a))\}
\end{aligned}$$

The tree returned by tr_n is always finite, but possibly not extensional. Hence we have to collapse its isomorphic subtrees adding up their weights by means of the *coll* function as follows:

$$\begin{aligned}
coll(t) &\triangleq \Big(\{\bot_t\} \cup \bigcup\{X' \mid ((X', succ'), p') \in U_a, a \in A\}, succ_c\Big) \\
\text{where } W_a &\triangleq \{(coll(t'), p) \mid (t', p) \in succ(\bot_t, a)\} \\
U_a &\triangleq \{(s, \textstyle\sum_{(s,p) \in W_a} p) \mid s \in \pi_1(W_a)\} \\
succ_c &\triangleq (\textstyle\bigcup_{a \in A}\{(\bot_t, a) \mapsto \{(\bot_s, p) \mid (s, p) \in U_a\}\}) \cup \bigcup_{(s,p) \in U_a} succ_s
\end{aligned}$$

Finally we can define the *pruning* of t as $t|_n \triangleq coll(tr_n(t))$.

We have now to show that RPT is (the carrier of) the final B_{RP}-coalgebra (up-to isomorphism). In order to simplify the proof, we reformulate B_{RP} in a slightly more "relational" format. We define a functor $D' : \mathsf{Set} \to \mathsf{Set}$ by letting for any set X:

$$D'X = \Big\{U \in \mathcal{P}_f(X \times \mathbb{R}_{(0,1]}) \mid \text{if } U \neq \emptyset \text{ then } \textstyle\sum_{(x,p) \in U} p = 1 \text{ and } \forall (x, p), (x', p') \in U : x = x' \Rightarrow p = p'\Big\}$$

and for any $f : X \to Y$, the function $D'f : D'X \to D'Y$ maps $U \in D'X$ to $\{(f(x), \sum_{(x,p) \in U} p) \mid x \in \pi_1(U)\}$. Then:

Proposition 3.6. The following hold true:

1. $D' \cong D + 1$.

2. $D'^A \cong B_{RP}$.

3. $Coalg(D'^A) \cong Coalg(B_{RP})$.

4. The supports of the final D'^A-coalgebra and of the final B_{RP}-coalgebra are isomorphic.

Proof 1: For $X \in \mathsf{Set}$, define $\phi_X : D'X \to DX + 1$ as $\phi_X(\emptyset) = *$, and for $U \neq \emptyset$, $\phi_X(U) : X \to \mathbb{R}_{[0,1]}$ maps x to p if $(x, p) \in U$, to 0 otherwise. It is easy to check that the ϕ_X's are invertible and form a natural isomorphism $\phi : D' \xrightarrow{\sim} D + 1$.

2: Trivial by 1; let $\psi : D'^A \xrightarrow{\sim} B_{RP}$ be the underlying natural isomorphism.

3: A D'^A-coalgebra $(X, \sigma : X \to D'(X)^A)$ is mapped to $(X, \psi_X \circ \sigma : X \to B_{RP}(X))$; the vice versa is similar, using ψ_X^{-1}. It is easy to check that these maps are inverse to each other.

4: Trivial by 3. ■

We can now prove that RPT is the carrier of the final B_{RP}-coalgebra. First, we observe that the set RPT can be endowed with a D'^A-coalgebra structure $\rho : RPT \to (D'(RPT))^A$ defined as follows, for $t = (X, succ)$:

$$\rho(t)(a) \triangleq \{(t[x], p) \mid (x, p) \in succ(\bot_t, a)\}$$

Theorem 3.7. (RPT, ρ) is a final B_{RP}-coalgebra.

Proof By Prop. 3.6, it suffices to prove that (RPT, ρ) is the final D'^A-coalgebra. To this end, we follow the construction given by Worrell in [31, Thm. 11]. We define an ordinal-indexed *final sequence* of sets $(B_\alpha)_\alpha$ together with "projection functions" $(f^\beta_\gamma : B_\beta \to B_\gamma)_{\gamma\leq\beta}$:

$$\begin{array}{rclrcl} B_0 & = & \{nil\} \cong 1 & f^1_0 & = & ! \\ B_{\alpha+1} & = & D'(B_\alpha)^A & f^{\alpha+2}_{\alpha+1} & = & D'(f^{\alpha+1}_\alpha)^A \\ B_\lambda & = & \lim\limits_{\alpha<\lambda} B_\alpha & f^\lambda_\alpha & = & \pi_\alpha \quad \text{for } \lambda \text{ a limit ordinal} \end{array}$$

the remaining f^β_γ being given by suitable compositions. D' is ω-accessible (because we restrict to finitely supported distributions), thus by [31, Thm. 13] and Prop. 3.6 the final sequence converges in at most $\omega + \omega$ steps to the set $B_{\omega+\omega}$ which is the carrier of the final D'^A-coalgebra.

Now, we have to prove that $B_{\omega+\omega}$ is isomorphic to RPT. An element of $B_{\omega+\omega}$ is a sequence of finite trees $\vec{t} = (t_0, t_1, \dots)$ such that for each $k \in \omega$ there exists $N_k \in \mathbb{N}$ such that nodes at depth k of any tree t_i have at most N_k successors for each label $a \in A$. These sequences can be seen as compatible partial views of a single (possibly infinite) tree. Thus, given a sequence $\vec{t}$, the corresponding tree $u \in RPT$ is obtained by *amalgamating* $\vec{t}$: u at depth k is defined by the level k of a suitable tree t_i, where i is such that for all $j \geq i$, t_j is equal to t_i up to depth k. On the other hand, given $u \in RPT$, we can define the corresponding sequence $\vec{t} \in B_{\omega+\omega}$ as $t_i = u|_i$.

It can be checked that these two maps form an isomorphism between $B_{\omega+\omega}$ and RPT. Moreover, they respect the coalgebraic structures, where $\tau : B_{\omega+\omega} \to D'(B_{\omega+\omega})^A$ is given by $\tau(\vec{t})(a) = \{\vec{t'} \in B_{\omega+\omega} \mid \forall i \in \omega : t'_i \in succ(t_i, a)\}$. Therefore, $(B_{\omega+\omega}, \tau)$ and (RPT, ρ) are isomorphic D'^A-coalgebras, hence the thesis. ■

3.3. Full Abstraction and Compactness

By virtue of Thm. 3.7, given an RPLTS $S = (S, A, \longrightarrow)$ there exists a unique coalgebra homomorphism $[\![\cdot]\!] : S \to RPT$, called the *(final) semantics* of S, which associates each state in S with its behavior. This semantics is *fully abstract*:

Theorem 3.8. Let $(S, A, \longrightarrow)$ be an RPLTS. For all $s_1, s_2 \in S$: $s_1 \sim_{\mathrm{PB}} s_2$ iff $[\![s_1]\!] = [\![s_2]\!]$.

Proof It follows from Props. 3.2 and 3.4 and Thm. 3.7. ■

A key property of reactive probabilistic trees is that they are *compact*, i.e., two different trees can be distinguished by looking at their finite subtrees only:

Theorem 3.9. For all $t_1, t_2 \in RPT$: $t_1 = t_2$ iff for all $n \in \mathbb{N} : t_1|_n = t_2|_n$.

Proof The "only if" is trivial. For the "if" direction, let us assume that $t_1 \neq t_2$; we have to find n such that $t_1|_n \neq t_2|_n$. Given a tree u_0, a *finite path* in u_0 is a sequence $(a_1, p_1, a_2, p_2, \dots, a_n, p_n)$ such that for $i = 1, \dots, n : (x_i, p_i) \in u_{i-1}(a_i)$ and $u_i = u[x_i]$. If $t_1 \neq t_2$, then there is a path of length n in, say, t_1 which cannot be replayed in t_2: in t_2 we reach a tree t'_{n-1} which either has no descendants for the label a_n (i.e., $t'_{n-1}(a_n) = \emptyset$), or $t'_{n-1}(a_n)$ does not contain any node associated with p_n. Therefore $t_1|_n \neq t_2|_n$. ■

Corollary 3.10. Let $(S, A, \longrightarrow)$ be an RPLTS. For all $s_1, s_2 \in S$: $s_1 \sim_{\mathrm{PB}} s_2$ iff for all $n \in \mathbb{N} : [\![s_1]\!]|_n = [\![s_2]\!]|_n$. ■

4. Constructive Proofs of Characterization of $\sim_{\mathrm{PB}}$ for $\mathrm{PML}_{\neg\wedge}$, $\mathrm{PML}_{\neg\vee}$, $\mathrm{PML}_{\vee}$, $\mathrm{PML}_{\wedge}$

By virtue of the categorical construction leading to Cor. 3.10, in order to prove that a modal logic characterizes $\sim_{\mathrm{PB}}$ over reactive probabilistic processes, it is enough to show that it can discriminate all reactive probabilistic trees of *finite* height. Based on this consideration, we now develop a constructive proof of logical characterization of $\sim_{\mathrm{PB}}$ for each of the four modal logics in Sect. 2.3, which naturally leads to a variant of Cleaveland algorithm [6] that explains $\sim_{\mathrm{PB}}$-inequivalence.

This section is organized as follows. We start with discussing a condition on the depth of distinguishing formulas (Sect. 4.1) and preparing the basis for probabilistic variants of Cleaveland algorithm (Sect. 4.2). In Sect. 4.3, we redemonstrate Larsen and Skou result for $\mathrm{PML}_{\neg\wedge}$ in the RPT_f setting with a proof that, with respect to the one in [22], is simpler and does not require the minimal deviation assumption (i.e., that the probability associated with any state in the support of the target distribution of a transition be a multiple of some value). This provides a proof scheme for the subsequent steps. In Sect. 4.4, we demonstrate that, as expected, $\mathrm{PML}_{\neg\vee}$ characterizes $\sim_{\mathrm{PB}}$ too, by adapting the proof scheme to cope with the replacement of conjunction with disjunction. In Sect. 4.5, we demonstrate the new result according to which also $\mathrm{PML}_{\vee}$ characterizes $\sim_{\mathrm{PB}}$, by further adapting the proof scheme to cope with the absence of negation. Finally, in Sect. 4.6 we redemonstrate Desharnais, Edalat, and Panangaden result for $\mathrm{PML}_{\wedge}$ through yet another adaptation of the proof scheme that, unlike the proof in [11], not only is constructive, but works directly on discrete state spaces without making use of measure-theoretic arguments based on analytic spaces.

4.1. On the Depth of Distinguishing Formulas

In each of our proofs, we show how to build a distinguishing formula for two arbitrary reactive probabilistic trees of finite height. In order for this formula to be distinguishing also on RPLTS models of which the two trees are finite approximations, a specific condition on the depth of the formula has to be satisfied, where $\mathit{depth}(\phi)$ is defined inductively on the syntactical structure of ϕ as follows:

$$\begin{array}{rcl}
\mathit{depth}(\top) & = & 0 \\
\mathit{depth}(\neg\phi') & = & \mathit{depth}(\phi') \\
\mathit{depth}(\phi_1 \wedge \phi_2) & = & \max(\mathit{depth}(\phi_1), \mathit{depth}(\phi_2)) \\
\mathit{depth}(\phi_1 \vee \phi_2) & = & \max(\mathit{depth}(\phi_1), \mathit{depth}(\phi_2)) \\
\mathit{depth}(\langle a\rangle_p \phi') & = & 1 + \mathit{depth}(\phi')
\end{array}$$

Proposition 4.1. Let $\mathcal{L}$ be one of the modal logics in Sect. 2.3. If $\mathcal{L}$ characterizes $=$ over RPT_f and for any two nodes t_1 and t_2 of an arbitrary RPT_f model such that $t_1 \neq t_2$ there exists $\phi \in \mathcal{L}$ distinguishing t_1 from t_2 such that $\mathit{depth}(\phi) \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$, then $\mathcal{L}$ characterizes $\sim_{\mathrm{PB}}$ over RPLTS models.

Proof Given two states s_1 and s_2 of an RPLTS, we have to prove that $s_1 \sim_{\mathrm{PB}} s_2$ iff s_1 and s_2 satisfy the same formulas of $\mathcal{L}$. If $s_1 \sim_{\mathrm{PB}} s_2$, then for all $n \in \mathbb{N}$ it holds that $[\![s_1]\!]|_n = [\![s_2]\!]|_n$ thanks to Cor. 3.10, hence s_1 and s_2 satisfy the same formulas of $\mathcal{L}$ because $\mathcal{L}$ characterizes $=$ over RPT_f. Suppose now that $s_1 \not\sim_{\mathrm{PB}} s_2$, from which it follows that there exists $n \in \mathbb{N}_{\geq 1}$ such that $[\![s_1]\!]|_n \neq [\![s_2]\!]|_n$ due to Cor. 3.10. Then there exists $\phi \in \mathcal{L}$ distinguishing $[\![s_1]\!]|_n$ from $[\![s_2]\!]|_n$ such that $\mathit{depth}(\phi) \leq \max(\mathit{height}([\![s_1]\!]|_n), \mathit{height}([\![s_2]\!]|_n)) = n$, hence the same formula ϕ also distinguishes s_1 from s_2. ■

Notice that, in the proof above, if $\mathit{depth}(\phi)$ were greater than n then, in general, ϕ may not distinguish higher prunings of $[\![s_1]\!]$ and $[\![s_2]\!]$, nor may any formula of depth at most n and derivable from ϕ still distinguish $[\![s_1]\!]|_n$ from $[\![s_2]\!]|_n$.

Example 4.2. Consider a process whose initial state s_1 has only an a-transition to a state having only a c-transition to *nil*, and another process whose initial state s_2 has only a b-transition to a state having only a d-transition to *nil*. Their corresponding trees differ at height $n = 1$ because $[\![s_1]\!]|_1$ has an a-transition to *nil* while $[\![s_2]\!]|_1$ has a b-transition to *nil*.

The formula of depth 2 given by $\langle a\rangle_1 \neg\langle c\rangle_1$ distinguishes $[\![s_1]\!]|_1$ from $[\![s_2]\!]|_1$, but this is no longer the case with $[\![s_1]\!]|_2$ and $[\![s_2]\!]|_2$ as neither satisfies that formula.

The formula of depth 2 given by $\langle a\rangle_1 \vee \langle b\rangle_1\langle c\rangle_1$ distinguishes $[\![s_1]\!]|_1$ from $[\![s_2]\!]|_1$, but this is no longer the case with the derived formula $\langle a\rangle_1 \vee \langle b\rangle_1$ of depth 1 as both nodes satisfy it. ■

4.2. Algorithms for Verifying $\sim_{\rm PB}$ and Explaining $\not\sim_{\rm PB}$

Several algorithms have been developed to verify bisimilarity and its variants on finite-state systems. Each of them exploits the fact that an equivalence relation like bisimilarity can be viewed as a partition, i.e., a set of pairwise disjoint subsets of states, usually called blocks and broadly representing equivalence classes, whose union yields the entire set of states. Starting with the partition formed by a single block containing all states, each of these algorithms repeatedly refines the current partition by splitting blocks according to the actions executable by their states and the blocks that are consequently reached. When the partition becomes stable, i.e., does not change anymore, the associated equivalence relation turns out to be a bisimulation.

Kanellakis and Smolka algorithm [19] for purely nondeterministic systems splits a block B with respect to an action a and a block B' by separating the states in B having an a-transition to a state in B' from the states in B having no a-transition to B'. It is easy to adapt this scheme to finite-state RPLTS models by observing that the splitting of B with respect to a and B' originates as many new blocks as there are different probabilities with which states in B reach states in B' after performing a. We show in Table 1 the resulting algorithm for verifying whether $s_1 \sim_{\rm PB} s_2$ with $s_1, s_2 \in S$ for a finite-state RPLTS $(S, A, \longrightarrow)$. We assume that every set is implemented as a queue so that, when adding to the current partition the local partition resulting from the splitting of a block, the new blocks are inserted at the end of the queue and hence can be later considered in the same for loop on the current partition.

The algorithm is formulated in a style inspired by [7] and is extended with the construction of the tree of blocks required by Cleaveland algorithm [6] for generating a distinguishing modal logic formula in case of inequivalence. The root of this tree is created with *make_root* and is labeled with the block containing all the states in S. Whenever a block B is split with respect to action a and block B', a new node is created with *make_node* for each of its sub-blocks; a new branch is also created with *make_branch*, which goes from the node of B to the node of its sub-block and is labeled with a, B', and the probability p with which each of the states in the sub-block of B reaches B' after performing a. The definition of *disting_formula* depends on the specific modal logic characterizing $\sim_{\rm PB}$ and hence is deferred to the following four sections.

4.3. PML$_{\neg\wedge}$ Characterizes $\sim_{\rm PB}$: A New Proof

Following Prop. 4.1, we prove that PML$_{\neg\wedge}$ characterizes $=$ over RPT_f. The interesting part is to show that the logical equivalence induced by PML$_{\neg\wedge}$ implies node equality $=$. We reason on the contrapositive, so given two nodes t_1 and t_2 such that $t_1 \neq t_2$, we proceed by induction on the height of t_1 to find a distinguishing PML$_{\neg\wedge}$ formula whose depth is not greater than the heights of t_1 and t_2. The idea is to exploit negation, so to ensure that certain distinguishing formulas are *satisfied* by a specific derivative t' of t_1 rather than the derivatives of t_2 different from t' (a derivative being a node reachable in one step), then take the *conjunction* of those formulas preceded by a diamond decorated with the probability for t_1 of *reaching* t'.

The only non-trivial case is the one in which t_1 and t_2 enable the same actions. At least one of those actions, say a, is such that, after performing it, the two nodes reach two distributions $\Delta_{1,a}$ and $\Delta_{2,a}$ such that $\Delta_{1,a} \neq \Delta_{2,a}$. Given a node $t' \in supp(\Delta_{1,a})$ such that $\Delta_{1,a}(t') > \Delta_{2,a}(t')$, by the induction hypothesis there exists a PML$_{\neg\wedge}$ formula $\phi'_{2,j}$ that distinguishes t' from a specific $t'_{2,j} \in supp(\Delta_{2,a}) \setminus \{t'\}$. We can assume that $t' \models \phi'_{2,j} \not\models t'_{2,j}$ otherwise, thanks to the presence of negation in PML$_{\neg\wedge}$, it would suffice to consider $\neg\phi'_{2,j}$. As a consequence, $t_1 \models \langle a \rangle_{\Delta_{1,a}(t')} \bigwedge_j \phi'_{2,j} \not\models t_2$ because $\Delta_{1,a}(t') > \Delta_{2,a}(t')$ and $\Delta_{2,a}(t')$ is the maximum probabilistic lower bound for which t_2 satisfies a formula of that form. Notice that $\Delta_{1,a}(t')$ may not be the maximum probabilistic lower bound for which t_1 satisfies such a formula, because $\bigwedge_j \phi'_{2,j}$ might be satisfied by other a-derivatives of t_1 in $supp(\Delta_{1,a}) \setminus \{t'\}$.

Theorem 4.3. Let $(T, A, \longrightarrow)$ be in RPT_f and $t_1, t_2 \in T$. Then $t_1 = t_2$ iff $t_1 \models \phi \iff t_2 \models \phi$ for all $\phi \in$ PML$_{\neg\wedge}$. Moreover, if $t_1 \neq t_2$, then there exists $\phi \in$ PML$_{\neg\wedge}$ distinguishing t_1 from t_2 such that $depth(\phi) \leq \max(height(t_1), height(t_2))$.

Proof Given $t_1, t_2 \in T$, we proceed as follows:

- If $t_1 = t_2$, then obviously $t_1 \models \phi \iff t_2 \models \phi$ for all $\phi \in$ PML$_{\neg\wedge}$.

```
$prob\_bisim(s_1, s_2, S, A, \longrightarrow)$
  begin
    $curr\_partition := \{S\}$;
    $prev\_partition := \emptyset$;
    $block\_tree := make\_root(S)$;
    while $(curr\_partition \neq prev\_partition)$ do begin
      $prev\_partition := curr\_partition$;
      for each $B \in curr\_partition$ do
        for each $(a, B') \in A \times prev\_partition$ do begin
          $local\_partition := split(B, a, B', \longrightarrow, block\_tree)$;
          if $(local\_partition \neq \{B\})$ then begin
            $curr\_partition := curr\_partition \setminus \{B\} \cup local\_partition$;
            break;
          end
        end
    end
    if $(same\_block(s_1, s_2, curr\_partition))$ then
      return $yes$;
    else
      return $disting\_formula(s_1, s_2, \longrightarrow, block\_tree)$;
  end
```

```
$split(B, a, B', \longrightarrow, block\_tree)$
  begin
    $prob\_set := \emptyset$;
    for each $s \in B$ do
      if $(s \xrightarrow{a} \Delta)$ then
        $prob\_set := prob\_set \cup \{\Delta(B')\}$;
      else
        $prob\_set := prob\_set \cup \{0\}$;
    $local\_partition := \emptyset$;
    for each $p \in prob\_set$ do begin
      $sub\_block := \{s \in B \mid (s \overset{a}{\not\longrightarrow} \wedge p = 0) \vee (s \xrightarrow{a} \Delta \wedge \Delta(B') = p)\}$;
      $local\_partition := local\_partition \cup \{sub\_block\}$;
      $make\_node(block\_tree, sub\_block)$;
      $make\_branch(block\_tree, B, sub\_block, (a, B', p))$;
    end
    return $local\_partition$;
  end
```

```
$same\_block(s_1, s_2, partition)$
  begin
    $same := false$;
    for each $B \in partition$ do
      if $(s_1 \in B \wedge s_2 \in B)$ then begin
        $same := true$;
        break;
      end
    return $same$;
  end
```

Table 1: Variant of Kanellakis and Smolka algorithm for $\sim_{\mathrm{PB}}$ extended with the block tree construction

- Assuming that $t_1 \neq t_2$, we show that there is $\phi \in \mathrm{PML}_{\neg\wedge}$, with $\mathit{depth}(\phi) \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$, such that it is not the case that $t_1 \models \phi \iff t_2 \models \phi$ by proceeding by induction on $\mathit{height}(t_1) \in \mathbb{N}$:
  - If $\mathit{height}(t_1) = 0$, then $\mathit{height}(t_2) \geq 1$ because $t_1 \neq t_2$. As a consequence, t_2 has at least one outgoing transition, say labeled with a, hence $t_1 \not\models \langle a \rangle_1 \mathrel{=\!\mid} t_2$. Notice that $\mathit{depth}(\langle a \rangle_1) = 1 \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$.
  - Let $\mathit{height}(t_1) = n + 1$ for some $n \in \mathbb{N}$ and suppose that for all $t'_1, t'_2 \in T$ such that $t'_1 \neq t'_2$ and $\mathit{height}(t'_1) \leq n$ there exists $\phi' \in \mathrm{PML}_{\neg\wedge}$, with $\mathit{depth}(\phi') \leq \max(\mathit{height}(t'_1), \mathit{height}(t'_2))$, such that it is not the case that $t'_1 \models \phi' \iff t'_2 \models \phi'$. Let $\mathit{init}(t_h)$, $h \in \{1, 2\}$, be the set of actions in A labeling the transitions departing from t_h:
    * If $\mathit{init}(t_1) \neq \mathit{init}(t_2)$, then it holds that $t_1 \models \langle a \rangle_1 \mathrel{\not=\!\mid} t_2$ for some $a \in \mathit{init}(t_1) \setminus \mathit{init}(t_2)$ or $t_1 \not\models \langle a \rangle_1 \mathrel{=\!\mid} t_2$ for some $a \in \mathit{init}(t_2) \setminus \mathit{init}(t_1)$, where $\mathit{depth}(\langle a \rangle_1) = 1 \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$.
    * If $\mathit{init}(t_1) = \mathit{init}(t_2)$, then $\mathit{init}(t_1) \neq \emptyset \neq \mathit{init}(t_2)$ as $\mathit{height}(t_1) \geq 1$. Since $t_1 \neq t_2$, there must exist $a \in \mathit{init}(t_1)$ such that $t_1 \xrightarrow{a} \Delta_{1,a}$, $t_2 \xrightarrow{a} \Delta_{2,a}$, and $\Delta_{1,a} \neq \Delta_{2,a}$. From $\Delta_{1,a} \neq \Delta_{2,a}$, it follows that there exists $t' \in \mathit{supp}(\Delta_{1,a})$ such that $1 \geq \Delta_{1,a}(t') > \Delta_{2,a}(t') \geq 0$. Assuming that $\mathit{supp}(\Delta_{2,a}) \backslash \{t'\} = \{t'_{2,1}, t'_{2,2}, \ldots, t'_{2,k}\}$, which cannot be empty because there must also exist $t'_2 \in \mathit{supp}(\Delta_{2,a})$ such that $0 \leq \Delta_{1,a}(t'_2) < \Delta_{2,a}(t'_2) \leq 1$, by the induction hypothesis for each $j = 1, 2, \ldots, k$ there exists $\phi'_{2,j} \in \mathrm{PML}_{\neg\wedge}$, with $\mathit{depth}(\phi'_{2,j}) \leq \max(\mathit{height}(t'),$ $\mathit{height}(t'_{2,j}))$, such that it is not the case that $t' \models \phi'_{2,j} \iff t'_{2,j} \models \phi'_{2,j}$. Since $\mathrm{PML}_{\neg\wedge}$ includes negation, without loss of generality we can assume that $t' \models \phi'_{2,j} \mathrel{\not=\!\mid} t'_{2,j}$. Therefore, it holds that $t_1 \models \langle a \rangle_{\Delta_{1,a}(t')} \bigwedge_{1 \leq j \leq k} \phi'_{2,j} \mathrel{\not=\!\mid} t_2$ because $\Delta_{1,a}(t') > \Delta_{2,a}(t')$ and $\Delta_{2,a}(t')$ is the maximum probabilistic lower bound for which t_2 satisfies a formula of that form. Notice that the resulting formula, which we denote by ϕ for short, satisfies:

$$\begin{aligned} \mathit{depth}(\phi) &= 1 + \max\nolimits_{1 \leq j \leq k} \mathit{depth}(\phi'_{2,j}) \\ &\leq 1 + \max\nolimits_{1 \leq j \leq k} \max(\mathit{height}(t'), \mathit{height}(t'_{2,j})) \\ &= 1 + \max(\mathit{height}(t'), \max\nolimits_{1 \leq j \leq k} \mathit{height}(t'_{2,j})) \\ &= \max(1 + \mathit{height}(t'), 1 + \max\nolimits_{1 \leq j \leq k} \mathit{height}(t'_{2,j})) \\ &\leq \max(\mathit{height}(t_1), \mathit{height}(t_2)) \end{aligned} \qquad \blacksquare$$

From the proof of Thm. 4.3, we directly derive a variant of Cleaveland algorithm [6] that computes a distinguishing $\mathrm{PML}_{\neg\wedge}$ formula for two $\sim_{\mathrm{PB}}$-inequivalent states s_1 and s_2 of a finite-state RPLTS $(S, A, \longrightarrow)$. The definition of *disting_formula*, which was left unspecified in Table 1, is now provided in Table 2; its correctness immediately stems from the proof of Thm. 4.3.

It starts with the identification in the block tree of the deepest block B to which s_1 and s_2 both belong. The result of *deepest_common_block* is the point at which s_1 and s_2 are separated, because they reach with different probabilities p_1 and p_2, respectively, a block B' after performing a, where a, B', p_1, and p_2 are retrieved through *branch_label* (the state acting as its third argument uniquely identifies a sub-block resulting from the splitting of B with respect to a and B'). We suppose that $p_1 > p_2$, otherwise the roles of s_1 and s_2 are exchanged through *swap* and the distinguishing formula is finally negated. If s_2 has no a-transition, then *disting_formula* returns $\langle a \rangle_1$ because $s_1 \models \langle a \rangle_1 \mathrel{\not=\!\mid} s_2$. Assume now that $s_1 \xrightarrow{a} \Delta_{1,a}$ and $s_2 \xrightarrow{a} \Delta_{2,a}$. Since $\Delta_{1,a}(B') = p_1 > p_2 = \Delta_{2,a}(B')$, block B' contains an equivalence class C' such that $\Delta_{1,a}(C') > \Delta_{2,a}(C')$. Notice that C', which is a leaf of the block tree and may coincide with the entire B', has the same role played by t' in the proof of Thm. 4.3. It holds that $\mathit{supp}(\Delta_{2,a}) \setminus C' = \{s'_{2,1}, s'_{2,2}, \ldots, s'_{2,k}\} \neq \emptyset$ because $\Delta_{2,a}(C') < \Delta_{1,a}(C') \leq 1$. In this case, fixing $s' \in C'$, *disting_formula* recursively computes $\phi'_{2,j}$ such that $s' \models \phi'_{2,j} \mathrel{\not=\!\mid} s'_{2,j}$ for each $j = 1, 2, \ldots, k$. Since several states $s'_{2,j}$ may belong to the same equivalence class, which is a leaf of the block tree retrievable through *leaf*, in the construction of ϕ we keep track of the classes that have already been considered. Formula $\langle a \rangle_{\Delta_{1,a}(C')} \bigwedge_j \phi'_{2,j}$ is then returned because every state in C' satisfies $\bigwedge_j \phi'_{2,j}$ and hence $s_1 \models \langle a \rangle_{\Delta_{1,a}(C')} \bigwedge_j \phi'_{2,j} \mathrel{\not=\!\mid} s_2$. This would not necessarily hold if we focussed on B' and took an arbitrary state in B', in the case that B' contained several equivalence classes.

```
$disting\_formula(s_1, s_2, \longrightarrow, block\_tree)$
    begin
        $B := deepest\_common\_block(block\_tree, s_1, s_2)$;
        $(a, B', p_1) := branch\_label(block\_tree, B, s_1)$;
        $(a, B', p_2) := branch\_label(block\_tree, B, s_2)$;
        if $(p_1 > p_2)$ then
            $negate\_formula := false$;
        else begin
            $negate\_formula := true$;
            $swap(s_1, s_2)$;
        end
        if $(s_2 \not\xrightarrow{a})$ then
            $\phi := \langle a \rangle_1$;
        else begin
            let $s_1 \xrightarrow{a} \Delta_{1,a}$;
            let $s_2 \xrightarrow{a} \Delta_{2,a}$;
            for each $C' \in leaves(block\_tree)$ do
                if $(C' \subseteq B' \wedge \Delta_{1,a}(C') > \Delta_{2,a}(C'))$ then
                    break;
            let $s' \in C'$;
            $\phi := true$;
            $considered\_class\_states := C'$;
            for each $s'_2 \in supp(\Delta_{2,a}) \setminus considered\_class\_states$ do begin
                $\phi := \phi \wedge disting\_formula(s', s'_2, \longrightarrow, block\_tree)$;
                $considered\_class\_states := considered\_class\_states \cup leaf(block\_tree, s'_2)$;
            end
            $\phi := \langle a \rangle_{\Delta_{1,a}(C')} \phi$;
        end
        if $(negate\_formula = true)$ then
            $\phi := \neg \phi$;
        return $\phi$;
    end
```

Table 2: Variant of Cleaveland algorithm for $\sim_{\mathrm{PB}}$ and $\mathrm{PML}_{\neg\wedge}$

4.4. $\mathrm{PML}_{\neg\vee}$ Characterizes $\sim_{\mathrm{PB}}$: Adapting the Proof

Since $\phi_1 \wedge \phi_2$ is logically equivalent to $\neg(\neg\phi_1 \vee \neg\phi_2)$, it is not surprising that $\mathrm{PML}_{\neg\vee}$ characterizes $\sim_{\mathrm{PB}}$ too. However, the proof of this result will be useful to develop the proof of the fact that $\mathrm{PML}_{\vee}$ characterizes $\sim_{\mathrm{PB}}$ as well. Similar to Thm. 4.3, also in the interesting part of the proof for $\mathrm{PML}_{\neg\vee}$ we reason on the contrapositive and proceed by induction. Given t_1 and t_2 such that $t_1 \neq t_2$, we again exploit negation so to ensure that certain distinguishing formulas are *not satisfied* by a specific derivative t' of t_1 rather than the derivatives of t_2 different from t', then take the *disjunction* of those formulas preceded by a diamond decorated with the probability for t_2 of *not reaching* t'.

In the only non-trivial case, $t' \in supp(\Delta_{1,a})$ is such that $\Delta_{1,a}(t') > \Delta_{2,a}(t')$. By the induction hypothesis, there exists a formula $\phi'_{2,j}$ distinguishing t' from a specific $t'_{2,j} \in supp(\Delta_{2,a}) \setminus \{t'\}$. We can assume that $t' \not\models \phi'_{2,j} \dashv t'_{2,j}$ (otherwise we consider $\neg\phi'_{2,j}$ as negation is in $\mathrm{PML}_{\neg\vee}$). Thus, $t_1 \not\models \langle a \rangle_{1-\Delta_{2,a}(t')} \bigvee_j \phi'_{2,j} \dashv t_2$ because $1-\Delta_{2,a}(t') > 1-\Delta_{1,a}(t')$ and the maximum probabilistic lower bound for which t_1 satisfies a formula of that form cannot exceed $1 - \Delta_{1,a}(t')$. Note that $1 - \Delta_{2,a}(t')$ is the *maximum* lower bound for which t_2 satisfies such a formula, because it is the probability with which t_2 does not reach t' after performing a.

Theorem 4.4. Let $(T, A, \longrightarrow)$ be in RPT_f and $t_1, t_2 \in T$. Then $t_1 = t_2$ iff $t_1 \models \phi \iff t_2 \models \phi$ for all $\phi \in \mathrm{PML}_{\neg\vee}$. Moreover, if $t_1 \neq t_2$, then there exists $\phi \in \mathrm{PML}_{\neg\vee}$ distinguishing t_1 from t_2 such that $\mathit{depth}(\phi) \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$.

Proof The proof is similar to the one of Thm. 4.3, apart from the final part of the last subcase, which changes as follows.
By the induction hypothesis, for each $j = 1, 2, \dots, k$ there exists $\phi'_{2,j} \in \mathrm{PML}_{\neg\vee}$, with $\mathit{depth}(\phi'_{2,j}) \leq \max(\mathit{height}(t'), \mathit{height}(t'_{2,j}))$, such that it is not the case that $t' \models \phi'_{2,j} \iff t'_{2,j} \models \phi'_{2,j}$. Since $\mathrm{PML}_{\neg\vee}$ includes negation, without loss of generality we can assume that $t' \not\models \phi'_{2,j} \Dashv t'_{2,j}$. Therefore, it holds that $t_1 \not\models \langle a \rangle_{1-\Delta_{2,a}(t')} \bigvee_{1 \leq j \leq k} \phi'_{2,j} \Dashv t_2$ because $1 - \Delta_{2,a}(t') > 1 - \Delta_{1,a}(t')$ and the maximum probabilistic lower bound for which t_1 satisfies a formula of that form cannot exceed $1 - \Delta_{1,a}(t')$. ■

The proof of Thm. 4.4 directly leads to a variant of Cleaveland algorithm [6] that computes a distinguishing $\mathrm{PML}_{\neg\vee}$ formula for two $\sim_{\mathrm{PB}}$-inequivalent states s_1 and s_2 of a finite-state RPLTS $(S, A, \longrightarrow)$. The definition of *disting_formula* is similar to the one provided in Table 2. The only differences are that ϕ is incrementally built as $\phi := \phi \vee \mathit{disting\_formula}(s', s'_2, \longrightarrow, \mathit{block\_tree})$ and at the end of the last for loop $\phi := \langle a \rangle_{1-\Delta_{2,a}(C')}\phi$.

4.5. Also PML$_\vee$ Characterizes $\sim_{\mathrm{PB}}$

The proof that $\mathrm{PML}_\vee$ characterizes $\sim_{\mathrm{PB}}$ is inspired by the one for $\mathrm{PML}_{\neg\vee}$, thus in its interesting part it considers the contrapositive and proceeds by induction. In the only non-trivial case, we will arrive at a situation in which $t_1 \not\models \langle a \rangle_{1-(\Delta_{2,a}(t')+p)} \bigvee_{j \in J} \phi'_{2,j} \Dashv t_2$, where:

- t' is a derivative of t_1 such that $\Delta_{1,a}(t') > \Delta_{2,a}(t')$ and $t' \not\models \bigvee_{j \in J} \phi'_{2,j}$;
- J is an index set identifying the derivatives t'' of t_2 other than t' such that $\Delta_{1,a}(t'') \neq \Delta_{2,a}(t'')$;
- p is the probability that t_2 reaches nodes $t''' \notin J$ such that $t''' \not\models \bigvee_{j \in J} \phi'_{2,j}$.

The choice of t' is crucial, because negation is no longer available in $\mathrm{PML}_\vee$. Unlike the case of $\mathrm{PML}_{\neg\vee}$, this induces the limitation to J and the introduction of p in the format of the distinguishing $\mathrm{PML}_\vee$ formula. An important observation for properly setting up all these parameters is that, in many cases, a disjunctive distinguishing formula can be obtained from a conjunctive one by suitably *increasing* some probabilistic lower bounds.

Example 4.5. The nodes t_1 and t_2 in Fig. 1(a) cannot be distinguished by any formula in which neither conjunction nor disjunction occurs. It holds that:

$$\begin{array}{lcl} t_1 & \models & \langle a \rangle_{0.5}\, (\langle b \rangle_1 \wedge \langle c \rangle_1) \;\not\Dashv\; t_2 \\ t_1 & \not\models & \langle a \rangle_{1.0}\, (\langle b \rangle_1 \vee \langle c \rangle_1) \;\Dashv\; t_2 \end{array}$$

Notice that, when moving from the conjunctive formula to the disjunctive one, the probabilistic lower bound decorating the a-diamond increases from 0.5 to 1 and the roles of t_1 and t_2 with respect to $\models$ are inverted.

The situation is similar for the nodes t_3 and t_4 in Fig. 1(b). Two occurrences of conjunction/disjunction are necessary:

$$\begin{array}{lcl} t_3 & \models & \langle a \rangle_{0.2}\, (\langle b \rangle_1 \wedge \langle c \rangle_1 \wedge \langle d \rangle_1) \;\not\Dashv\; t_4 \\ t_3 & \models & \langle a \rangle_{0.9}\, (\langle b \rangle_1 \vee \langle c \rangle_1 \vee \langle d \rangle_1) \;\not\Dashv\; t_4 \end{array}$$

but the roles of t_3 and t_4 with respect to $\models$ cannot be inverted. ■

However, increasing some of the probabilistic lower bounds in a conjunctive distinguishing formula does not always yield a disjunctive one. This is the case when the use of conjunction/disjunction is not necessary for telling two different nodes apart.

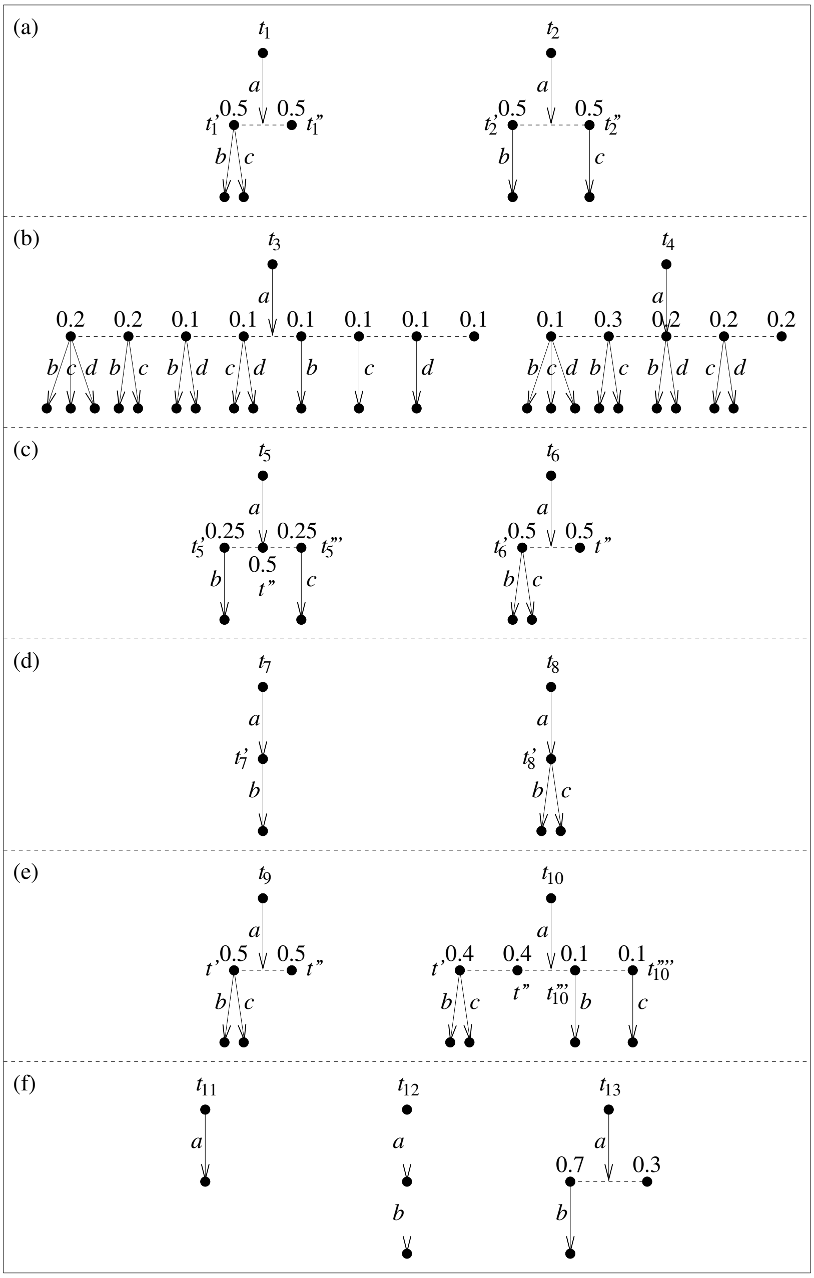

Figure 1: RPT_f models used in the examples of Sects. 4.5 and 4.6

Example 4.6. For the nodes t_5 and t_6 in Fig. 1(c), it holds that:

$$t_5 \not\models \langle a \rangle_{0.5} \, (\langle b \rangle_1 \wedge \langle c \rangle_1) \;=\!\!\mid\; t_6$$

If we replace conjunction with disjunction and we vary the probabilistic lower bound between 0.5 and 1, we produce no disjunctive formula capable of discriminating between t_5 and t_6. Nevertheless, a distinguishing formula belonging to $\mathrm{PML}_\vee$ exists because:

$$t_5 \not\models \langle a \rangle_{0.5} \, \langle b \rangle_1 \;=\!\!\mid\; t_6$$

where disjunction does not occur at all. ■

The examples above show that the increase of some probabilistic lower bounds when moving from conjunctive distinguishing formulas to disjunctive ones takes place only in the case that the probabilities of reaching certain nodes have to be *summed up*. Additionally, we recall that, in order for two nodes to be related by $\sim_{\mathrm{PB}}$, they must enable the same actions, so focussing on a *single* action is enough for discriminating when only disjunction is available.

Bearing this in mind, given a reactive probabilistic tree of finite height, for any node t we define the set $\Phi_\vee(t)$ of $\mathrm{PML}_\vee$ formulas satisfied by t each of which features:

- probabilistic lower bounds of diamonds that are *maximal* with respect to the satisfiability of a formula of that format by t (this is consistent with the observation in the last sentence before Thm. 4.4, and keeps the set $\Phi_\vee(t)$ finite);
- diamonds that arise only from *existing* transitions that depart from t (so to avoid useless diamonds in disjunctions and hence keep the set $\Phi_\vee(t)$ finite);
- disjunctions that stem only from *single* transitions of *different* nodes in the support of a distribution reached by t (transitions departing from the same node would result in formulas like $\bigvee_{h \in H} \langle a_h \rangle_{p_h} \phi_h$, with $a_{h_1} \neq a_{h_2}$ for $h_1 \neq h_2$, which are useless for discriminating with respect to $\sim_{\mathrm{PB}}$) and are preceded by a diamond decorated with the *sum* of the probabilities assigned to those nodes by the distribution reached by t.

Definition 4.7. The set $\Phi_\vee(t)$ for a node t of finite height is defined by induction on *height*(t) as follows:

- If *height*$(t) = 0$, then $\Phi_\vee(t) = \emptyset$.
- If *height*$(t) \geq 1$ for t having transitions of the form $t \xrightarrow{a_i} \Delta_i$ with $\mathit{supp}(\Delta_i) = \{t'_{i,j} \mid j \in J_i\}$ and $i \in I \neq \emptyset$, then:

$$\begin{aligned} \Phi_\vee(t) \;=\; & \{\langle a_i \rangle_1 \mid i \in I\} \\ & \cup \bigcup_{i \in I} \mathit{hplb}\Big(\bigcup_{\emptyset \neq J' \subseteq J_i} \{\langle a_i \rangle_{\sum\limits_{j \in J'} \Delta_i(t'_{i,j})} \dot{\bigvee_{j \in J'}} \phi'_{i,j,k} \mid t'_{i,j} \in \mathit{supp}(\Delta_i), \phi'_{i,j,k} \in \Phi_\vee(t'_{i,j})\}\Big) \end{aligned}$$

where operator $\dot{\vee}$ is a variant of $\vee$ in which identical operands are not admitted (i.e., idempotence is forced) and function *hplb* keeps only the formula with the highest probabilistic lower bound decorating the initial a_i-diamond among the formulas differring only for that bound. ■

To illustrate the definition given above, we exhibit some examples showing the usefulness of $\Phi_\vee$-sets for discrimination purposes. In particular, let us reconsider the non-trivial case mentioned at the beginning of this subsection. Given two different nodes that with the same action reach two different distributions, a good criterion for choosing t' (a derivative of the first node not satisfying certain formulas, to which the first distribution assigns a probability greater than the second one) seems to be the *minimality* of the $\Phi_\vee$-set.

Example 4.8. For the nodes t_7 and t_8 in Fig. 1(d), we have:

$$\begin{array}{rcl} \Phi_{\vee}(t_7) & = & \{\langle a\rangle_1,\ \langle a\rangle_1\langle b\rangle_1\} \\ \Phi_{\vee}(t_8) & = & \{\langle a\rangle_1,\ \langle a\rangle_1\langle b\rangle_1,\ \langle a\rangle_1\langle c\rangle_1\} \end{array}$$

A formula like $\langle a\rangle_1\,(\langle b\rangle_1 \vee \langle c\rangle_1)$ is useless for discriminating between t_7 and t_8, because disjunction is between two actions enabled by the same node and hence constituting a nondeterministic choice. Indeed, such a formula is not part of $\Phi_{\vee}(t_8)$. While in the case of conjunction it is often necessary to concentrate on several alternative actions, in the case of disjunction it is convenient to focus on a single action per node when aiming at producing a distinguishing formula.

The fact that $\langle a\rangle_1\langle c\rangle_1 \in \Phi_{\vee}(t_8)$ is a distinguishing formula can be retrieved as follows. Starting from the two identically labeled transitions $t_7 \xrightarrow{a} \Delta_{7,a}$ and $t_8 \xrightarrow{a} \Delta_{8,a}$ where $\Delta_{7,a}(t'_7) = 1 = \Delta_{8,a}(t'_8)$ and $\Delta_{7,a}(t'_8) = 0 = \Delta_{8,a}(t'_7)$, we have:

$$\begin{array}{rcl} \Phi_{\vee}(t'_7) & = & \{\langle b\rangle_1\} \\ \Phi_{\vee}(t'_8) & = & \{\langle b\rangle_1,\ \langle c\rangle_1\} \end{array}$$

If we focus on t'_7 because $\Delta_{7,a}(t'_7) > \Delta_{8,a}(t'_7)$ and its $\Phi_{\vee}$-set is minimal, then $t'_7 \not\models \langle c\rangle_1 \Dashv t'_8$ with $\langle c\rangle_1 \in \Phi_{\vee}(t'_8) \setminus \Phi_{\vee}(t'_7)$. As a consequence, $t_7 \not\models \langle a\rangle_1\langle c\rangle_1 \Dashv t_8$ where the value 1 decorating the a-diamond stems from $1 - \Delta_{8,a}(t'_7)$. ■

Example 4.9. For the nodes t_1 and t_2 in Fig. 1(a), we have:

$$\begin{array}{rcl} \Phi_{\vee}(t_1) & = & \{\langle a\rangle_1,\ \langle a\rangle_{0.5}\langle b\rangle_1,\ \langle a\rangle_{0.5}\langle c\rangle_1\} \\ \Phi_{\vee}(t_2) & = & \{\langle a\rangle_1,\ \langle a\rangle_{0.5}\langle b\rangle_1,\ \langle a\rangle_{0.5}\langle c\rangle_1,\ \langle a\rangle_1\,(\langle b\rangle_1 \vee \langle c\rangle_1)\} \end{array}$$

The formulas with two diamonds and no disjunction are identical in the two sets, so their disjunction $\langle a\rangle_{0.5}\langle b\rangle_1 \vee \langle a\rangle_{0.5}\langle c\rangle_1$ is useless for discriminating between t_1 and t_2. Indeed, such a formula is part of neither $\Phi_{\vee}(t_1)$ nor $\Phi_{\vee}(t_2)$. In contrast, their disjunction in which decorations of identical diamonds are summed up, i.e., $\langle a\rangle_1\,(\langle b\rangle_1 \vee \langle c\rangle_1)$, is fundamental. It belongs only to $\Phi_{\vee}(t_2)$ because in the case of t_1 the b-transition and the c-transition depart from the same node, hence no probabilities can be added.

The fact that $\langle a\rangle_1\,(\langle b\rangle_1 \vee \langle c\rangle_1) \in \Phi_{\vee}(t_2)$ is a distinguishing formula can be retrieved as follows. Starting from the two identically labeled transitions $t_1 \xrightarrow{a} \Delta_{1,a}$ and $t_2 \xrightarrow{a} \Delta_{2,a}$ where $\Delta_{1,a}(t'_1) = \Delta_{1,a}(t''_1) = 0.5 = \Delta_{2,a}(t'_2) = \Delta_{2,a}(t''_2)$ and $\Delta_{1,a}(t'_2) = \Delta_{1,a}(t''_2) = 0 = \Delta_{2,a}(t'_1) = \Delta_{2,a}(t''_1)$, we have:

$$\begin{array}{rclcrcl} \Phi_{\vee}(t'_1) & = & \{\langle b\rangle_1,\ \langle c\rangle_1\} & \qquad & \Phi_{\vee}(t''_1) & = & \emptyset \\ \Phi_{\vee}(t'_2) & = & \{\langle b\rangle_1\} & & \Phi_{\vee}(t''_2) & = & \{\langle c\rangle_1\} \end{array}$$

If we focus on t''_1 because $\Delta_{1,a}(t''_1) > \Delta_{2,a}(t''_1)$ and its $\Phi_{\vee}$-set is minimal, then $t''_1 \not\models \langle b\rangle_1 \Dashv t'_2$ with $\langle b\rangle_1 \in \Phi_{\vee}(t'_2) \setminus \Phi_{\vee}(t''_1)$ as well as $t''_1 \not\models \langle c\rangle_1 \Dashv t''_2$ with $\langle c\rangle_1 \in \Phi_{\vee}(t''_2) \setminus \Phi_{\vee}(t''_1)$. As a consequence, $t_1 \not\models \langle a\rangle_1\,(\langle b\rangle_1 \vee \langle c\rangle_1) \Dashv t_2$ where the value 1 decorating the a-diamond stems from $1 - \Delta_{2,a}(t''_1)$. ■

Example 4.10. For the nodes t_5 and t_6 in Fig. 1(c), we have:

$$\begin{array}{l} \Phi_{\vee}(t_5) = \{\langle a\rangle_1,\ \langle a\rangle_{0.25}\langle b\rangle_1,\ \langle a\rangle_{0.25}\langle c\rangle_1,\ \langle a\rangle_{0.5}\,(\langle b\rangle_1 \vee \langle c\rangle_1)\} \\ \Phi_{\vee}(t_6) = \{\langle a\rangle_1,\ \langle a\rangle_{0.5}\langle b\rangle_1,\ \langle a\rangle_{0.5}\langle c\rangle_1\} \end{array}$$

The formulas with two diamonds and no disjunction are different in the two sets, so they are enough for discriminating between t_5 and t_6. In contrast, the only formula with disjunction, which belongs to $\Phi_{\vee}(t_5)$, is useless because the probability decorating its a-diamond is equal to the probability decorating the a-diamond of each of the two formulas with two diamonds in $\Phi_{\vee}(t_6)$.

The fact that $\langle a\rangle_{0.5}\langle b\rangle_1 \in \Phi_{\vee}(t_6)$ is a distinguishing formula can be retrieved as follows. Starting from the two identically labeled transitions $t_5 \xrightarrow{a} \Delta_{5,a}$ and $t_6 \xrightarrow{a} \Delta_{6,a}$ where $\Delta_{5,a}(t'_5) = \Delta_{5,a}(t'''_5) = 0.25$, $\Delta_{5,a}(t'') = 0.5 = \Delta_{6,a}(t'_6) = \Delta_{6,a}(t'')$, and $\Delta_{5,a}(t'_6) = 0 = \Delta_{6,a}(t'_5) = \Delta_{6,a}(t'''_5)$, we have:

$$\begin{array}{rclcrcl} \Phi_{\vee}(t'_5) & = & \{\langle b\rangle_1\} & \qquad & \Phi_{\vee}(t'''_5) & = & \{\langle c\rangle_1\} \\ \Phi_{\vee}(t'_6) & = & \{\langle b\rangle_1,\ \langle c\rangle_1\} & & \Phi_{\vee}(t'') & = & \emptyset \end{array}$$

Notice that t'' might be useless for discriminating purposes because it has the same probability in both distributions, so we exclude it. If we focus on t'''_5 because $\Delta_{5,a}(t'''_5) > \Delta_{6,a}(t'''_5)$ and its $\Phi_\vee$-set is minimal after the exclusion of t'', then $t'''_5 \not\models \langle b\rangle_1 \Dashv t'_6$ with $\langle b\rangle_1 \in \Phi_\vee(t'_6) \setminus \Phi_\vee(t'''_5)$, while no distinguishing formula is considered with respect to t'' as element of $supp(\Delta_{6,a})$ due to the exclusion of t'' itself. As a consequence, $t_5 \not\models \langle a\rangle_{0.5}\langle b\rangle_1 \Dashv t_6$ where the value 0.5 decorating the a-diamond stems from $1 - (\Delta_{6,a}(t'''_5) + p)$ with $p = \Delta_{6,a}(t'')$. The reason for subtracting the probability that t_6 reaches t'' after performing a is that $t'' \not\models \langle b\rangle_1$.

We conclude by observing that focussing on t'' as derivative with the minimum $\Phi_\vee$-set is indeed problematic, because it would result in $\langle a\rangle_{0.5}\langle b\rangle_1$ when considering t'' as derivative of t_5, but it would result in $\langle a\rangle_{0.5}\,(\langle b\rangle_1 \vee \langle c\rangle_1)$ when considering t'' as derivative of t_6, with the latter formula not distinguishing between t_5 and t_6. Moreover, when focussing on t'''_5, no formula ϕ' could have been found such that $t'''_5 \not\models \phi' \Dashv t''$ as $\Phi_\vee(t'') \subsetneq \Phi_\vee(t'''_5)$. ■

The last example shows that, in the general format for the distinguishing PML$_\vee$ formula mentioned at the beginning of this subsection, i.e., $\langle a\rangle_{1-(\Delta_{2,a}(t')+p)} \bigvee_{j\in J} \phi'_{2,j}$, the set J only contains any derivative of the second node different from t' to which the two distributions assign two *different* probabilities. No derivative of the two original nodes having the same probability in both distributions is taken into account even if its $\Phi_\vee$-set is minimal – because it might be useless for discriminating purposes – nor is it included in J – because there might be no formula satisfied by this node when viewed as a derivative of the second node, which is not satisfied by t'. Furthermore, the value p is the probability that the second node reaches the excluded derivatives that do *not* satisfy $\bigvee_{j\in J} \phi'_{2,j}$; note that the first node reaches those derivatives with the same probability p.

We present two additional examples illustrating some technicalities of Def. 4.7. The former example shows the usefulness of the operator $\dot\vee$ and of the function *hplb* for selecting the right t' on the basis of the minimality of its $\Phi_\vee$-set among the derivatives of the first node to which the first distribution assigns a probability greater than the second one. The latter example emphasizes the role played, for the same purpose as before, by formulas occurring in a $\Phi_\vee$-set whose number of nested diamonds is not maximal.

Example 4.11. For the nodes t_9 and t_{10} in Fig. 1(e), we have:

$$\begin{aligned}
\Phi_\vee(t_9) &= \{\langle a\rangle_1,\ \langle a\rangle_{0.5}\langle b\rangle_1,\ \langle a\rangle_{0.5}\langle c\rangle_1\} \\
\Phi_\vee(t_{10}) &= \{\langle a\rangle_1,\ \langle a\rangle_{0.5}\langle b\rangle_1,\ \langle a\rangle_{0.5}\langle c\rangle_1,\ \langle a\rangle_{0.6}\,(\langle b\rangle_1 \vee \langle c\rangle_1)\}
\end{aligned}$$

Starting from the two identically labeled transitions $t_9 \xrightarrow{a} \Delta_{9,a}$ and $t_{10} \xrightarrow{a} \Delta_{10,a}$ where $\Delta_{9,a}(t') = \Delta_{9,a}(t'')$ $= 0.5$, $\Delta_{10,a}(t') = \Delta_{10,a}(t'') = 0.4$, $\Delta_{10,a}(t'''_{10}) = \Delta_{10,a}(t''''_{10}) = 0.1$, and $\Delta_{9,a}(t'''_{10}) = \Delta_{9,a}(t''''_{10}) = 0$, we have:

$$\begin{aligned}
\Phi_\vee(t') &= \{\langle b\rangle_1,\ \langle c\rangle_1\} & \Phi_\vee(t'') &= \emptyset \\
\Phi_\vee(t'''_{10}) &= \{\langle b\rangle_1\} & \Phi_\vee(t''''_{10}) &= \{\langle c\rangle_1\}
\end{aligned}$$

If we focus on t'' because $\Delta_{9,a}(t'') > \Delta_{10,a}(t'')$ and its $\Phi_\vee$-set is minimal, then $t'' \not\models \langle b\rangle_1 \Dashv t'$ with $\langle b\rangle_1 \in \Phi_\vee(t') \setminus \Phi_\vee(t'')$, $t'' \not\models \langle b\rangle_1 \Dashv t'''_{10}$ with $\langle b\rangle_1 \in \Phi_\vee(t'''_{10}) \setminus \Phi_\vee(t'')$, and $t'' \not\models \langle c\rangle_1 \Dashv t''''_{10}$ with $\langle c\rangle_1 \in \Phi_\vee(t''''_{10}) \setminus \Phi_\vee(t'')$. As a consequence, $t_9 \not\models \langle a\rangle_{0.6}\,(\langle b\rangle_1 \vee \langle c\rangle_1) \Dashv t_{10}$ where the formula belongs to $\Phi_\vee(t_{10})$ and the value 0.6 decorating the a-diamond stems from $1 - \Delta_{10,a}(t'')$.

If $\vee$ were used in place of $\dot\vee$, then in $\Phi_\vee(t_{10})$ we would also have formulas like $\langle a\rangle_{0.5}\,(\langle b\rangle_1 \vee \langle b\rangle_1)$ and $\langle a\rangle_{0.5}\,(\langle c\rangle_1 \vee \langle c\rangle_1)$. These are useless in that logically equivalent to other formulas already in $\Phi_\vee(t_{10})$ in which disjunction does not occur. Most importantly, they would apparently augment the size of $\Phi_\vee(t_{10})$, an inappropriate fact if t_{10} were a derivative of some other node instead of being the root of a tree.

If *hplb* were not used, then in $\Phi_\vee(t_{10})$ we would also have formulas like $\langle a\rangle_{0.1}\langle b\rangle_1$, $\langle a\rangle_{0.4}\langle b\rangle_1$, $\langle a\rangle_{0.1}\langle c\rangle_1$, and $\langle a\rangle_{0.4}\langle c\rangle_1$, in which the probabilistic lower bounds of the a-diamonds are not maximal with respect to the satisfiability of formulas of that form by t_{10}; those with maximal probabilistic lower bounds associated with a-diamonds are $\langle a\rangle_{0.5}\langle b\rangle_1$ and $\langle a\rangle_{0.5}\langle c\rangle_1$, which already belong to $\Phi_\vee(t_{10})$. In the case that t_9 and t_{10} were derivatives of two nodes under comparison instead of being the roots of two trees, the presence of those additional formulas in $\Phi_\vee(t_{10})$ may lead to focus on t_{10} instead of t_9 – for reasons that will be clear in Ex. 4.13 – thereby producing no distinguishing formula. ■

Example 4.12. For the nodes t_{11}, t_{12}, t_{13} in Fig. 1(f), we have:

$$\begin{array}{rcl}\Phi_{\vee}(t_{11}) & = & \{\langle a\rangle_1\} \\ \Phi_{\vee}(t_{12}) & = & \{\langle a\rangle_1,\ \langle a\rangle_1\langle b\rangle_1\} \\ \Phi_{\vee}(t_{13}) & = & \{\langle a\rangle_1,\ \langle a\rangle_{0.7}\langle b\rangle_1\}\end{array}$$

Let us view them as derivatives of other nodes, rather than roots of trees. The presence of formula $\langle a\rangle_1$ in $\Phi_{\vee}(t_{12})$ and $\Phi_{\vee}(t_{13})$ – although it has not the maximum number of nested diamonds in those two sets – ensures the minimality of $\Phi_{\vee}(t_{11})$ and hence that t_{11} is selected for building a distinguishing formula. If $\langle a\rangle_1$ were not in $\Phi_{\vee}(t_{12})$ and $\Phi_{\vee}(t_{13})$, then t_{12} and t_{13} could be selected, but no distinguishing formula satisfied by t_{11} could be obtained. ■

The criterion for selecting the right t' based on the minimality of its $\Phi_{\vee}$-set has to take into account a further aspect related to *formulas without disjunctions*. If two derivatives – with different probabilities in the two distributions – have the same formulas without disjunctions in their $\Phi_{\vee}$-sets, then a distinguishing formula for the two nodes will have disjunctions in it (see Exs. 4.9 and 4.11). In contrast, if the formulas without disjunctions are different between the two $\Phi_{\vee}$-sets, then one of those formulas will tell the two derivatives apart (see Ex. 4.8).

A particular instance of the second case is the one in which for each formula without disjunctions in one of the two $\Phi_{\vee}$-sets there is a variant in the other $\Phi_{\vee}$-set – i.e., a formula without disjunctions that has the same format but may differ for the values of some probabilistic lower bounds – and vice versa. In this event, *regardless of the minimality* of the $\Phi_{\vee}$-sets, it has to be selected the derivative such that (i) for each formula without disjunctions in its $\Phi_{\vee}$-set there exists a variant in the $\Phi_{\vee}$-set of the other derivative such that the probabilistic lower bounds in the former formula are $\leq$ than the corresponding bounds in the latter formula and (ii) at least one probabilistic lower bound in a formula without disjunctions in the $\Phi_{\vee}$-set of the selected derivative is $<$ than the corresponding bound in the corresponding variant in the $\Phi_{\vee}$-set of the other derivative. We say that the $\Phi_{\vee}$-set of the selected derivative is a $(\leq, <)$*-variant* of the $\Phi_{\vee}$-set of the other derivative.

Example 4.13. Let us view the nodes t_5 and t_6 in Fig. 1(c) as derivatives of other nodes, rather than roots of trees. Based on their $\Phi_{\vee}$-sets shown in Ex. 4.10, we should focus on t_6 because $\Phi_{\vee}(t_6)$ contains fewer formulas. However, by so doing, we would be unable to find a distinguishing formula in $\Phi_{\vee}(t_5)$ that is not satisfied by t_6. Indeed, if we look carefully at the formulas without disjunctions in $\Phi_{\vee}(t_5)$ and $\Phi_{\vee}(t_6)$, we note that they differ only for their probabilistic lower bounds: $\langle a\rangle_1 \in \Phi_{\vee}(t_6)$ is a variant of $\langle a\rangle_1 \in \Phi_{\vee}(t_5)$, $\langle a\rangle_{0.5}\langle b\rangle_1 \in \Phi_{\vee}(t_6)$ is a variant of $\langle a\rangle_{0.25}\langle b\rangle_1 \in \Phi_{\vee}(t_5)$, and $\langle a\rangle_{0.5}\langle c\rangle_1 \in \Phi_{\vee}(t_6)$ is a variant of $\langle a\rangle_{0.25}\langle c\rangle_1 \in \Phi_{\vee}(t_5)$. Therefore, we must focus on t_5 because $\Phi_{\vee}(t_5)$ contains formulas without disjunctions such as $\langle a\rangle_{0.25}\langle b\rangle_1$ and $\langle a\rangle_{0.25}\langle c\rangle_1$ having smaller bounds: $\Phi_{\vee}(t_5)$ is a $(\leq, <)$-variant of $\Phi_{\vee}(t_6)$.

Consider now the nodes t_9 and t_{10} in Fig. 1(e), whose $\Phi_{\vee}$-sets are shown in Ex. 4.11. If function *hplb* were not used and hence $\Phi_{\vee}(t_{10})$ also contained $\langle a\rangle_{0.1}\langle b\rangle_1$, $\langle a\rangle_{0.4}\langle b\rangle_1$, $\langle a\rangle_{0.1}\langle c\rangle_1$, and $\langle a\rangle_{0.4}\langle c\rangle_1$, then the formulas without disjunctions in $\Phi_{\vee}(t_9)$ would no longer be equal to those in $\Phi_{\vee}(t_{10})$. More precisely, the formulas without disjunctions would be similar between the two sets, with those in $\Phi_{\vee}(t_{10})$ having smaller probabilistic lower bounds, so that we would erroneously focus on t_{10}. ■

Summing up, in the construction of the distinguishing $\mathrm{PML}_{\vee}$ formula mentioned at the beginning of this subsection, i.e., $\langle a\rangle_{1-(\Delta_{2,a}(t')+p)} \bigvee_{j\in J} \phi'_{2,j}$, the steps for choosing the derivative t', on the basis of which each subformula $\phi'_{2,j}$ is then generated so that it is not satisfied by t', are the following:

1. Consider only derivatives to which $\Delta_{1,a}$ assigns a probability greater than the one assigned by $\Delta_{2,a}$.

2. Within the previous set, eliminate all the derivatives whose $\Phi_{\vee}$-sets have $(\leq, <)$-variants.

3. Among the remaining derivatives, focus on one of those having a minimal $\Phi_{\vee}$-set.

Theorem 4.14. Let $(T, A, \longrightarrow)$ be in RPT_f and $t_1, t_2 \in T$. Then $t_1 = t_2$ iff $t_1 \models \phi \iff t_2 \models \phi$ for all $\phi \in \mathrm{PML}_\vee$. Moreover, if $t_1 \neq t_2$, then there exists $\phi \in \mathrm{PML}_\vee$ distinguishing t_1 from t_2 such that $\mathit{depth}(\phi) \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$.

Proof Given $t_1, t_2 \in T$, we proceed as follows:

- If $t_1 = t_2$, then obviously $t_1 \models \phi \iff t_2 \models \phi$ for all $\phi \in \mathrm{PML}_\vee$.
- Assuming that $t_1 \neq t_2$, we show that there exists $\phi \in \Phi_\vee(t_1) \cup \Phi_\vee(t_2)$, which ensures that $\mathit{depth}(\phi) \leq \max(\mathit{height}(t_1), \mathit{height}(t_2))$, such that it is not the case that $t_1 \models \phi \iff t_2 \models \phi$ by proceeding by induction on $\mathit{height}(t_1) \in \mathbb{N}$. The proof is similar to the one of Thm. 4.4, in particular in the cases $\mathit{height}(t_1) = 0$ and $\mathit{height}(t_1) = n + 1$ with $\mathit{init}(t_1) \neq \mathit{init}(t_2)$ it benefits from the presence of $\{\langle a_i \rangle 1 \mid i \in I\}$ in $\Phi_\vee(t)$ as of Def. 4.7. However, it changes as follows before the application of the induction hypothesis in the case $\mathit{height}(t_1) = n + 1$ with $\mathit{init}(t_1) = \mathit{init}(t_2) \neq \emptyset$ and $t_1 \xrightarrow{a} \Delta_{1,a}$, $t_2 \xrightarrow{a} \Delta_{2,a}$, and $\Delta_{1,a} \neq \Delta_{2,a}$ for some $a \in \mathit{init}(t_1)$.
  Let $\mathit{supp}_a = \mathit{supp}(\Delta_{1,a}) \cup \mathit{supp}(\Delta_{2,a})$, which can be partitioned into $\mathit{supp}_{a,\neq} = \{t' \in \mathit{supp}_a \mid \Delta_{1,a}(t') \neq \Delta_{2,a}(t')\}$ and $\mathit{supp}_{a,=} = \{t' \in \mathit{supp}_a \mid \Delta_{1,a}(t') = \Delta_{2,a}(t')\}$ with $|\,\mathit{supp}_{a,\neq}\,| \geq 2$ because $\Delta_{1,a} \neq \Delta_{2,a}$ and $|\,\mathit{supp}_{a,=}\,| \geq 0$. We recall that $\Phi_\vee(t'')$ is a $(\leq, <)$-variant of $\Phi_\vee(t')$ iff:

  – For each formula without disjunctions in one of the two $\Phi_\vee$-sets, there exists a variant in the other $\Phi_\vee$-set – i.e., a formula without disjunctions that has the same format but may differ for the values of some probabilistic lower bounds – and vice versa (this means that there exists a bijection between the formulas without disjunctions in the two $\Phi_\vee$-sets, because the maximality of the probabilistic lower bounds in a $\Phi_\vee$-set implies the existence of at most one formula with a given format in the $\Phi_\vee$-set).

  – For each formula without disjunctions in $\Phi_\vee(t'')$, there exists a variant in $\Phi_\vee(t')$ such that the probabilistic lower bounds in the former formula are $\leq$ than the corresponding bounds in the latter formula.

  – At least one probabilistic lower bound in a formula without disjunctions in $\Phi_\vee(t'')$ is $<$ than the corresponding bound in the corresponding variant in $\Phi_\vee(t')$.

  Among all the nodes in $\mathit{supp}_{a,\neq}$, there is one denoted by t' such that, for all $t'' \in \mathit{supp}_{a,\neq} \setminus \{t'\}$, $\Phi_\vee(t'')$ is not a $(\leq, <)$-variant of $\Phi_\vee(t')$ as we now prove by proceeding by induction on $|\,\mathit{supp}_{a,\neq}\,| \in \mathbb{N}_{\geq 2}$:

  – If $|\,\mathit{supp}_{a,\neq}\,| = 2$ – hence $\mathit{supp}_{a,\neq} = \{t', t''\}$ – then trivially at least one of $\Phi_\vee(t')$ and $\Phi_\vee(t'')$ is not a $(\leq, <)$-variant of the other. Indeed, if this were not the case, then at least one probabilistic lower bound in a formula without disjunctions in $\Phi_\vee(t'')$ would be $<$ than the corresponding bound in the corresponding variant in $\Phi_\vee(t')$ *and* at least one probabilistic lower bound in a formula without disjunctions in $\Phi_\vee(t')$ would be $<$ than the corresponding bound in the corresponding variant in $\Phi_\vee(t'')$, but then it would *not* hold that for each formula without disjunctions in $\Phi_\vee(t')$ there exists a variant in $\Phi_\vee(t'')$ such that the probabilistic lower bounds in the former formula are $\leq$ than the corresponding bounds in the latter formula *and* for each formula without disjunctions in $\Phi_\vee(t'')$ there exists a variant in $\Phi_\vee(t')$ such that the probabilistic lower bounds in the former formula are $\leq$ than the corresponding bounds in the latter formula.

  – Let $|\,\mathit{supp}_{a,\neq}\,| = n + 1$ for some $n \in \mathbb{N}_{\geq 2}$ and suppose that the result holds for each subset of $\mathit{supp}_{a,\neq}$ of cardinality between 2 and n. Assuming that $\mathit{supp}_{a,\neq} = \{t'_1, t'_2, \dots, t'_{n+1}\}$, we denote by t' the node in $\mathit{supp}_{a,\neq} \setminus \{t'_{n+1}\}$ that, by the induction hypothesis, enjoys the property over that subset. There are two cases:

    ∗ If $\Phi_\vee(t'_{n+1})$ is not a $(\leq, <)$-variant of $\Phi_\vee(t')$ either, then t' enjoys the property over the entire set $\mathit{supp}_{a,\neq}$.

  - * Suppose that $\Phi_{\vee}(t'_{n+1})$ is a $(\leq,<)$-variant of $\Phi_{\vee}(t')$, which implies that $\Phi_{\vee}(t')$ cannot be a $(\leq,<)$-variant of $\Phi_{\vee}(t'_{n+1})$. From the fact that, for all $t'' \in \mathit{supp}_{a,\neq} \setminus \{t', t'_{n+1}\}$, $\Phi_{\vee}(t'')$ is not a $(\leq,<)$-variant of $\Phi_{\vee}(t')$ by the induction hypothesis, it follows that $\Phi_{\vee}(t'')$ is not a $(\leq,<)$-variant of $\Phi_{\vee}(t'_{n+1})$. Indeed, for each such t'' the set $\Phi_{\vee}(t'')$ contains at least a formula without disjunctions that is not a variant of any formula without disjunctions in $\Phi_{\vee}(t')$, or all formulas without disjunctions in $\Phi_{\vee}(t'')$ are identical to formulas without disjunctions in $\Phi_{\vee}(t')$, hence this holds true with respect to $\Phi_{\vee}(t'_{n+1})$ too, given that $\Phi_{\vee}(t'_{n+1})$ is a $(\leq,<)$-variant of $\Phi_{\vee}(t')$. As a consequence, t'_{n+1} enjoys the property over the entire set $\mathit{supp}_{a,\neq}$.

Within the set of all the nodes $t' \in \mathit{supp}_{a,\neq}$ enjoying the property above, we select one with a minimal $\Phi_{\vee}$-set, which we denote by $t'_{\min}$. Suppose that $\Delta_{1,a}(t'_{\min}) > \Delta_{2,a}(t'_{\min})$ and let $t'_{2,j}$ be an arbitrary node belonging to $\mathit{supp}_{a,\neq,2} = (\mathit{supp}_{a,\neq} \setminus \{t'_{\min}\}) \cap \mathit{supp}(\Delta_{2,a})$; if $\Delta_{1,a}(t'_{\min}) < \Delta_{2,a}(t'_{\min})$, we would focus on an arbitrary node $t'_{1,j} \in \mathit{supp}_{a,\neq,1}$. By the induction hypothesis, from $t'_{2,j} \neq t'$ it follows that there exists $\phi'_{2,j} \in \Phi_{\vee}(t'_{\min}) \cup \Phi_{\vee}(t'_{2,j})$ such that it is not the case that $t'_{\min} \models \phi'_{2,j} \iff t'_{2,j} \models \phi'_{2,j}$. In particular, it holds that $t'_{\min} \not\models \phi'_{2,j} \mathrel{=\!\!|} t'_{2,j}$ because $\phi'_{2,j} \in \Phi_{\vee}(t'_{2,j})$, as can be seen by considering the following two cases based on the fact that $\Phi_{\vee}(t'_{2,j})$ is not a $(\leq,<)$-variant of $\Phi_{\vee}(t'_{\min})$:

- If at least one formula without disjunctions in $\Phi_{\vee}(t'_{2,j})$ is not a variant of any formula without disjunctions in $\Phi_{\vee}(t'_{\min})$, then such a formula can be taken as $\phi'_{2,j}$ given the maximality of the probabilistic lower bounds of any basic formula in $\Phi_{\vee}(t'_{\min})$.
- If all basic formulas in $\Phi_{\vee}(t'_{2,j})$ are identical to basic formulas in $\Phi_{\vee}(t'_{\min})$, then $\Phi_{\vee}(t'_{2,j})$ must contain some more formulas (with disjunctions) not in $\Phi_{\vee}(t'_{\min})$ given the minimality of the latter set, otherwise we would have selected $t'_{2,j}$ in place of $t'_{\min}$. One of the additional formulas (with disjunctions) in $\Phi_{\vee}(t'_{2,j})$ can be taken as $\phi'_{2,j}$.

Letting $\mathit{supp}_{a,=,\not\models} = \{t' \in \mathit{supp}_{a,=} \mid t' \not\models \bigvee_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \phi'_{2,j}\}$ as well as $p_{\not\models} = \Delta_{2,a}(\mathit{supp}_{a,=,\not\models}) = \Delta_{1,a}(\mathit{supp}_{a,=,\not\models})$, we have that $t_1 \not\models \langle a \rangle_{1-(\Delta_{2,a}(t'_{\min})+p_{\not\models})} \bigvee_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \phi'_{2,j} \mathrel{=\!\!|} t_2$ as $1-(\Delta_{2,a}(t'_{\min}) + p_{\not\models}) > 1-(\Delta_{1,a}(t'_{\min})+p_{\not\models})$ and the maximum probabilistic lower bound for which t_1 satisfies a formula of that form cannot exceed $1-(\Delta_{1,a}(t'_{\min}) + p_{\not\models})$. The distinguishing $\mathrm{PML}_{\vee}$ formula above may not be in $\Phi_{\vee}(t_2)$, but it is logically implied by, or equivalent to, a distinguishing formula in $\Phi_{\vee}(t_2)$ for the following reasons:

- Each $t'_{2,j}$ belongs to $\mathit{supp}(\Delta_{2,a})$.
- Each $\phi'_{2,j}$ belongs to $\Phi_{\vee}(t'_{2,j})$.
- The probabilistic lower bound $1-(\Delta_{2,a}(t'_{\min}) + p_{\not\models})$ is equal to $\sum_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \Delta_{2,a}(t'_{2,j}) + \Delta_{2,a}(\mathit{supp}_{a,=,\models})$, so in the distinguishing $\mathrm{PML}_{\vee}$ formula it is enough to replace $\bigvee_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \phi'_{2,j}$ with $\dot{\bigvee}_{t' \in \mathit{supp}_{a,\neq,2} \cup \mathit{supp}_{a,=,\models}} \phi_{t'}$ where:
  - * $\phi_{t'} = \phi'_{2,j}$ if $t' = t'_{2,j}$ for some j;
  - * $\phi_{t'} = \phi' \in \Phi_{\vee}(t')$ if $t' \neq t'_{2,j}$ for all j, where $\phi' \implies \phi'_{2,j}$ for some j and the existence of such a ϕ' in $\Phi_{\vee}(t')$ stems from $t' \in \mathit{supp}_{a,=,\models}$, i.e., $t' \models \bigvee_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \phi'_{2,j}$. ■

From the proof of Thm. 4.14, we directly derive a variant of Cleaveland algorithm [6] that computes a distinguishing $\mathrm{PML}_{\vee}$ formula for two $\sim_{\mathrm{PB}}$-inequivalent states s_1 and s_2 of a finite-state RPLTS $(S, A, \longrightarrow)$. The definition of *disting_formula* is provided in Table 3, with $C'_{\min}$ playing the same role as $t'_{\min}$, and its correctness directly stems from the proof of Thm. 4.14. A major difference with respect to the definition in Table 2 is that, in the case in which both s_1 and s_2 have an outgoing a-transition, the calculation of the distinguishing formula is no longer recursive as it relies on *discriminate_$\phi_{\vee}$_sets*. This discriminates two states based on the formulas in the $\Phi_{\vee}$-sets of their finite reactive probabilistic tree approximations as established towards the end of the proof of Thm. 4.14.

```
$disting\_formula(s_1, s_2, \longrightarrow, block\_tree)$
    begin
        $B := deepest\_common\_block(block\_tree, s_1, s_2)$;
        $(a, B', p_1) := branch\_label(block\_tree, B, s_1)$;
        $(a, B', p_2) := branch\_label(block\_tree, B, s_2)$;
        if $(s_1 \overset{a}{\not\longrightarrow} \vee s_2 \overset{a}{\not\longrightarrow})$ then
            $\phi := \langle a \rangle_1$;
        else begin
            let $s_1 \xrightarrow{a} \Delta_{1,a}$;
            let $s_2 \xrightarrow{a} \Delta_{2,a}$;
            $supp\_classes_{a,\neq} := \emptyset$;
            $supp\_classes_{a,=} := \emptyset$;
            for each $C \in leaves(block\_tree)$ do
                if $(\Delta_{1,a}(C) \neq \Delta_{2,a}(C))$ then
                    $supp\_classes_{a,\neq} := supp\_classes_{a,\neq} \cup \{C\}$;
                else
                    $supp\_classes_{a,=} := supp\_classes_{a,=} \cup \{C\}$;
            $supp\_classes_{a,\neq,nv} := find\_classes\_with\_no\_variants(supp\_classes_{a,\neq})$;
            $C'_{\min} := find\_class\_min\_\Phi_\vee\_set(supp\_classes_{a,\neq,nv})$;
            if $(\Delta_{1,a}(C'_{\min}) < \Delta_{2,a}(C'_{\min}))$ then
                $swap(s_1, s_2)$;
            let $s' \in C'_{\min}$;
            $\phi := true$;
            $considered\_class\_states := C'_{\min}$;
            for each $s'_2 \in supp(\Delta_{2,a}) \setminus considered\_class\_states$ do begin
                $\phi := \phi \vee discriminate\_\Phi_\vee\_sets(s', s'_2, \longrightarrow)$;
                $considered\_class\_states := considered\_class\_states \cup leaf(block\_tree, s'_2)$;
            end
            $supp\_classes_{a,=,\not\models} := check(supp\_classes_{a,=}, \phi)$;
            $p_{\not\models} := \Delta_{2,a}(\bigcup supp\_classes_{a,=,\not\models})$;
            $\phi := \langle a \rangle_{1-(\Delta_{2,a}(C'_{\min})+p_{\not\models})}\phi$;
        end
        return $\phi$;
    end
```

Table 3: Variant of Cleaveland algorithm for $\sim_{\mathrm{PB}}$ and $\mathrm{PML}_\vee$

4.6. PML$_\wedge$ Characterizes $\sim_{\rm PB}$: A Direct Proof for Discrete Systems without Measure Theory

In our setting, we can also demonstrate that PML$_\wedge$ characterizes $\sim_{\rm PB}$ by working directly on *discrete* state spaces in a way that, unlike [11, 9], does not make any use of measure theory. The idea is to obtain $t_1 \models \langle a \rangle_{\Delta_{1,a}(t')+p} \bigwedge_{j \in J} \phi'_{2,j} \not\models t_2$ by adapting the proof of Thm. 4.14 consistently with the proof of Thm. 4.3.

To this purpose, given a reactive probabilistic tree of finite height, for any node t we define the set $\Phi_\wedge(t)$ of PML$_\wedge$ formulas satisfied by t featuring – in addition to maximal probabilistic lower bounds and diamonds arising only from transitions of t as for $\Phi_\vee(t)$ – conjunctions that (i) stem only from transitions departing from the *same node* in the support of a distribution reached by t and (ii) are preceded by a diamond decorated with the *sum* of the probabilities assigned by that distribution to that node and other nodes with the *same transitions* considered for that node. Formally, given t having transitions of the form $t \xrightarrow{a_i} \Delta_i$ with $supp(\Delta_i) = \{t'_{i,j} \mid j \in J_i\}$ and $i \in I \neq \emptyset$, we let:

$$\begin{array}{rcl} \Phi_\wedge(t) & = & \{\langle a_i \rangle_1 \mid i \in I\} \\ & \cup & \bigcup\limits_{i \in I} splb(\{\!| \, \langle a_i \rangle_{\Delta_i(t'_{i,j})} \bigwedge\limits_{k \in K'} \phi'_{i,j,k} \mid \emptyset \neq K' \subseteq K_{i,j}, t'_{i,j} \in supp(\Delta_i), \phi'_{i,j,k} \in \Phi_\wedge(t'_{i,j}) \, |\!\}) \end{array}$$

where $\{\!|$ and $|\!\}$ are multiset parentheses, $K_{i,j}$ is the index set for $\Phi_\wedge(t'_{i,j})$, and function *splb* merges all formulas possibly differring only for the probabilistic lower bound decorating their initial a_i-diamond by summing up those bounds (notice that such formulas stem from different nodes in $supp(\Delta_i)$).

We now provide some examples illustrating the technicalities of the definition above, as well as the fact that a good criterion for choosing t' occurring in the distinguishing PML$_\wedge$ formula at the beginning of this subsection is the *maximality* of the $\Phi_\wedge$-set.

Example 4.15. In Fig. 1(b), the multiset giving rise to $\Phi_\wedge(t_3)$ contains two occurrences of $\langle a \rangle_{0.2} \langle b \rangle_1$ and two occurrences of $\langle a \rangle_{0.1} \langle b \rangle_1$, which are merged into $\langle a \rangle_{0.6} \langle b \rangle_1$ by function *splb*. Likewise, the multiset behind $\Phi_\wedge(t_4)$ contains formulas $\langle a \rangle_{0.1} \langle b \rangle_1$, $\langle a \rangle_{0.3} \langle b \rangle_1$, and $\langle a \rangle_{0.2} \langle b \rangle_1$, which are merged into $\langle a \rangle_{0.6} \langle b \rangle_1$. ■

Example 4.16. For the nodes t_1 and t_2 in Fig. 1(a), we have:

$$\begin{array}{rcl} \Phi_\wedge(t_1) & = & \{\langle a \rangle_1, \, \langle a \rangle_{0.5} \langle b \rangle_1, \, \langle a \rangle_{0.5} \langle c \rangle_1, \, \langle a \rangle_{0.5} \, (\langle b \rangle_1 \wedge \langle c \rangle_1)\} \\ \Phi_\wedge(t_2) & = & \{\langle a \rangle_1, \, \langle a \rangle_{0.5} \langle b \rangle_1, \, \langle a \rangle_{0.5} \langle c \rangle_1\} \end{array}$$

The conjunction $\langle a \rangle_{0.5} \langle b \rangle_1 \wedge \langle a \rangle_{0.5} \langle c \rangle_1$ is useless for discriminating between t_1 and t_2 – it is part of neither $\Phi_\wedge(t_1)$ nor $\Phi_\wedge(t_2)$ – while $\langle a \rangle_{0.5} \, (\langle b \rangle_1 \wedge \langle c \rangle_1)$ is the only distinguishing formula and belongs only to $\Phi_\wedge(t_1)$, because in the case of t_2 the b-transition and the c-transition depart from two different nodes. Starting from the two identically labeled transitions $t_1 \xrightarrow{a} \Delta_{1,a}$ and $t_2 \xrightarrow{a} \Delta_{2,a}$ where $\Delta_{1,a}(t'_1) = \Delta_{1,a}(t''_1) = 0.5 = \Delta_{2,a}(t'_2) = \Delta_{2,a}(t''_2)$ and $\Delta_{1,a}(t'_2) = \Delta_{1,a}(t''_2) = 0 = \Delta_{2,a}(t'_1) = \Delta_{2,a}(t''_1)$, we have:

$$\begin{array}{rclcrcl} \Phi_\wedge(t'_1) & = & \{\langle b \rangle_1, \, \langle c \rangle_1\} & \qquad & \Phi_\wedge(t''_1) & = & \emptyset \\ \Phi_\wedge(t'_2) & = & \{\langle b \rangle_1\} & & \Phi_\wedge(t''_2) & = & \{\langle c \rangle_1\} \end{array}$$

If we focus on t'_1 because $\Delta_{1,a}(t'_1) > \Delta_{2,a}(t'_1)$ and its $\Phi_\wedge$-set is maximal, then $t'_1 \models \langle c \rangle_1 \not\models t'_2$ with $\langle c \rangle_1 \in \Phi_\wedge(t'_1) \setminus \Phi_\wedge(t'_2)$ as well as $t'_1 \models \langle b \rangle_1 \not\models t''_2$ with $\langle b \rangle_1 \in \Phi_\wedge(t'_1) \setminus \Phi_\wedge(t''_2)$. As a consequence, $t_1 \models \langle a \rangle_{0.5} \, (\langle b \rangle_1 \wedge \langle c \rangle_1) \not\models t_2$ where the value 0.5 decorating the a-diamond stems from $\Delta_{1,a}(t'_1)$. ■

As far as the other two variables occurring in the distinguishing PML$_\wedge$ formula at the beginning of this subsection are concerned, J only contains any derivative of the second node different from t' to which the two distributions assign two *different* probabilities, while p is the probability of reaching derivatives having the *same* probability in both distributions that *satisfy* $\bigwedge_{j \in J} \phi'_{2,j}$. Moreover, when selecting t', we have to leave out all the derivatives whose $\Phi_\wedge$-sets have $(\leq, <)$-variants.

Theorem 4.17. Let $(T, A, \longrightarrow)$ be in RPT_f and $t_1, t_2 \in T$. Then $t_1 = t_2$ iff $t_1 \models \phi \iff t_2 \models \phi$ for all $\phi \in$ PML$_\wedge$. Moreover, if $t_1 \neq t_2$, then there exists $\phi \in$ PML$_\wedge$ distinguishing t_1 from t_2 such that $depth(\phi) \leq \max(height(t_1), height(t_2))$.

Proof Similar to that of Thm. 4.14, with these differences:

- We select $t'_{\max}$ as one of the nodes with maximal $\Phi_{\wedge}$-set in $\mathit{supp}_{a,\neq}$ having no $(\leq, <)$-variants.
- It holds that $t'_{\max} \models \phi'_{2,j} \not\dashv t'_{2,j}$ for all $t'_{2,j} \in \mathit{supp}_{a,\neq,2}$ because $\phi'_{2,j} \in \Phi_{\wedge}(t'_{\max})$ thanks to the maximality of $\Phi_{\wedge}(t'_{\max})$.
- Letting $\mathit{supp}_{a,=,\models} = \{t' \in \mathit{supp}_{a,=} \mid t' \models \bigwedge_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \phi'_{2,j}\}$ as well as $p_{\models} = \Delta_{1,a}(\mathit{supp}_{a,=,\models}) = \Delta_{2,a}(\mathit{supp}_{a,=,\models})$, we have that $t_1 \models \langle a \rangle_{\Delta_{1,a}(t'_{\max})+p_{\models}} \bigwedge_{t'_{2,j} \in \mathit{supp}_{a,\neq,2}} \phi'_{2,j} \not\dashv t_2$ because $\Delta_{1,a}(t'_{\max})+p_{\models} > \Delta_{2,a}(t'_{\max}) + p_{\models}$ and the maximum probabilistic lower bound for which t_2 satisfies a formula of that form cannot exceed $\Delta_{2,a}(t'_{\max}) + p_{\models}$.
- The distinguishing $\mathrm{PML}_{\wedge}$ formula is in $\Phi_{\wedge}(t_1)$ due to *splb*. ■

The proof of Thm. 4.17 directly leads to a variant of Cleaveland algorithm [6] that computes a distinguishing $\mathrm{PML}_{\wedge}$ formula for two $\sim_{\mathrm{PB}}$-inequivalent states s_1 and s_2 of a finite-state RPLTS $(S, A, \longrightarrow)$. The definition of *disting_formula* is similar to the one provided in Table 3. The only differences are that $\Phi_{\wedge}$-sets are considered in place of $\Phi_{\vee}$-sets, $C'_{\max}$ is used instead of $C'_{\min}$, ϕ is incrementally built as $\phi := \phi \wedge \mathit{discriminate\_\Phi_{\wedge}\_sets}(s', s'_2, \longrightarrow)$, $\mathit{supp\_classes}_{a,=,\models}$ is used instead of $\mathit{supp\_classes}_{a,=,\not\models}$, $p_{\models}$ is used instead of $p_{\not\models}$, and at the end $\phi := \langle a \rangle_{\Delta_{1,a}(C'_{\max})+p_{\models}} \phi$.

For the sake of completeness, we mention that a different Cleaveland-inspired algorithm for computing a distinguishing $\mathrm{PML}_{\wedge}$ formula in the finite-state case can be found in [11]. Different from ours, this latter algorithm does *not* directly arise from the proof – based on measure-theoretic arguments – of the logical characterization result in [11].

5. Conclusions

In this paper, we have studied modal logic characterizations of bisimilarity over reactive probabilistic systems by introducing of the fully abstract and compact model of *reactive probabilistic trees*. This has allowed us to develop uniform and constructive alternative proofs of the logical characterization results of Larsen and Skou [22] and Desharnais, Edalat, and Panangaden [11] without making use of measure-theoretic arguments, and also to show that conjunction can be replaced by disjunction without having to reintroduce negation.

The intuition behind our result for $\mathrm{PML}_{\vee}$ is that from a conjunctive distinguishing formula it is often possible to derive a disjunctive one by suitably increasing some probabilistic lower bounds. On the model side, this corresponds to summing up the probabilities of reaching certain states that are in the support of a target distribution. In fact, a state of an RPLTS can be given a semantics as a reactive probabilistic tree, and hence it is characterized by the countable set of formulas (approximated by the $\Phi_{\vee}$-set) obtained by doing finite visits of the tree.

In summary, our new result for $\mathrm{PML}_{\vee}$ completes the picture for reactive probabilistic systems by emphasizing that conjunction and disjunction are *interchangeable* to characterize (bi)simulation equivalence. In contrast, the two logical connectives are *both necessary* for characterizing simulation preorder, as established in [12].

As an application of the results in this paper, the $\mathrm{PML}_{\vee}$-based characterization of bisimilarity may help to prove a conjecture in [4]. That work studies the discriminating power of three different testing equivalences over reactive probabilistic systems, respectively using reactive probabilistic tests, fully nondeterministic tests, and nondeterministic and probabilistic tests. Numerous examples lead to conjecture that testing equivalence based on nondeterministic and probabilistic tests may have the same discriminating power as bisimilarity. Given two $\sim_{\mathrm{PB}}$-inequivalent reactive probabilistic systems, the idea of the tentative proof is to build a distinguishing nondeterministic and probabilistic test from a distinguishing $\mathrm{PML}_{\wedge}$ formula. One of the main difficulties with carrying out such a proof is the fact that choices within tests fit well together with disjunction (available in $\mathrm{PML}_{\vee}$) rather than conjunction (available in $\mathrm{PML}_{\wedge}$). This may now be overcome by starting the tentative proof from a distinguishing $\mathrm{PML}_{\vee}$ formula.

Acknowledgment: We are grateful to the anonymous reviewers for their comments and suggestions.

References

[1] P. Aczel and N. Mendler. A final coalgebra theorem. In *Proc. of the 3rd Conf. on Category Theory and Computer Science (CTCS 1989)*, volume 389 of *LNCS*, pages 357–365. Springer, 1989.

[2] C. Baier and M. Kwiatkowska. Domain equations for probabilistic processes. *Mathematical Structures in Computer Science*, 10:665–717, 2000.

[3] M. Bernardo and M. Miculan. Disjunctive probabilistic modal logic is enough for bisimilarity on reactive probabilistic systems. In *Proc. of the 17th Italian Conf. on Theoretical Computer Science (ICTCS 2016)*, volume 1720, pages 203–220. CEUR-WS, 2016.

[4] M. Bernardo, D. Sangiorgi, and V. Vignudelli. 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)*, volume 8657 of *LNCS*, pages 281–296. Springer, 2014.

[5] P. Billingsley. *Probability and Measure.* Wiley-Interscience, 1995.

[6] R. Cleaveland. On automatically explaining bisimulation inequivalence. In *Proc. of the 2nd Int. Workshop on Computer Aided Verification (CAV 1990)*, volume 531 of *LNCS*, pages 364–372. Springer, 1990.

[7] R. Cleaveland and O. Sokolsky. Equivalence and preorder checking for finite-state systems. In *Handbook of Process Algebra*, pages 391–424. Elsevier, 2001.

[8] V. Danos, J. Desharnais, F. Laviolette, and P. Panangaden. Bisimulation and cocongruence for probabilistic systems. *Information and Computation*, 204:503–523, 2006.

[9] Y. Deng and H. Wu. Modal characterisations of probabilistic and fuzzy bisimulations. In *Proc. of the 16th Int. Conf. on Formal Engineering Methods (ICFEM 2014)*, volume 8829 of *LNCS*, pages 123–138. Springer, 2014.

[10] C. Derman. *Finite State Markovian Decision Processes.* Academic Press, 1970.

[11] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 179:163–193, 2002.

[12] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labelled Markov processes. *Information and Computation*, 184:160–200, 2003.

[13] N. Fijalkow, B. Klin, and P. Panangaden. Expressiveness of probabilistic modal logics, revisited. In *Proc. of the 44th Int. Coll. on Automata, Languages and Programming (ICALP 2017)*, volume 80 of *LIPIcs*, pages 105:1–105:12, 2017.

[14] R.J. van Glabbeek. The linear time – branching time spectrum I. In *Handbook of Process Algebra*, pages 3–99. Elsevier, 2001.

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

[16] H. Hansson. *Time and Probability in Formal Design of Distributed Systems.* PhD Thesis, 1992.

[17] M. Hennessy and R. Milner. Algebraic laws for nondeterminism and concurrency. *Journal of the ACM*, 32:137–162, 1985.

[18] B. Jacobs and A. Sokolova. Exemplaric expressivity of modal logics. *Journal of Logic and Computation*, 20:1041–1068, 2010.

[19] P.C. Kanellakis and S.A. Smolka. CCS expressions, finite state processes, and three problems of equivalence. *Information and Computation*, 86:43–68, 1990.

[20] R.M. Keller. Formal verification of parallel programs. *Communications of the ACM*, 19:371–384, 1976.

[21] J.G. Kemeny and J.L. Snell. *Finite Markov Chains.* Van Nostrand, 1960.

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

[23] K.G. Larsen and A. Skou. Compositional verification of probabilistic processes. In *Proc. of the 3rd Int. Conf. on Concurrency Theory (CONCUR 1992)*, volume 630 of *LNCS*, pages 456–471. Springer, 1992.

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

[25] P. Panangaden. Probabilistic bisimulation. In *Advanced Topics in Bisimulation and Coinduction*, pages 300–334. Cambridge University Press, 2011.

[26] M.O. Rabin. Probabilistic automata. *Information and Control*, 6:230–245, 1963.

[27] J.J.M.M. Rutten. Universal coalgebra: A theory of systems. *Theoretical Computer Science*, 249:3–80, 2000.

[28] R. Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems.* PhD Thesis, 1995.

[29] M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. of the 26th IEEE Symp. on Foundations of Computer Science (FOCS 1985)*, pages 327–338. IEEE-CS Press, 1985.

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

[31] J. Worrell. On the final sequence of a finitary set functor. *Theoretical Computer Science*, 338:184–199, 2005.