Disjunctive Probabilistic Modal Logic is Enough for Bisimilarity on Reactive Probabilistic Systems

Marco Bernardo
Dip. di Scienze Pure e Applicate, Università di Urbino, Italy

Marino Miculan
Dip. di Scienze Matematiche, Informatiche e Fisiche, Università di Udine, Italy

Abstract
Larsen and Skou characterized probabilistic bisimilarity over reactive probabilistic systems with a logic including true, negation, conjunction, and a diamond modality decorated with a probabilistic lower bound. Later on, Desharnais, Edalat, and Panangaden showed that negation is not necessary to characterize the same equivalence. In this paper, we prove that the logical characterization holds also when conjunction is replaced by disjunction, with negation still being not necessary. To this end, we introduce reactive probabilistic trees, a fully abstract model for reactive probabilistic systems that allows us to demonstrate expressivity of the disjunctive probabilistic modal logic, as well as of the previously mentioned logics, by means of a compactness argument.

1. Introduction
Since its introduction by Larsen and Skou [9], probabilistic bisimilarity has been used to compare probabilistic systems, like an implementation against its specification. It corresponds to Milner’s strong bisimilarity for nondeterministic systems, and coincides with lumpability for Markov chains. Remarkably, bisimilarity can be given a logical characterization: two processes are bisimilar if and only if they satisfy the same set of formulas of a suitable logic. This is important both from a fundamental and an implementation point of view: it allows to understand the logical complexity of the equivalence under scrutiny, and it can be used to define an algorithm for deciding bisimilarity of finite-state systems by constructing a formula that witnesses the failure of bisimilarity. Hence, the simpler the logic, the simpler the algorithm.

Larsen and Skou [9] proved that, in the case of reactive probabilistic systems, probabilistic bisimilarity can be characterized by a propositional modal logic similar to Hennessy-Milner logic [6]: it features the usual propositional constructs $\top$, $\land$, and $\neg$, plus a diamond modality $\langle a \rangle_p$ decorated with a probabilistic lower bound. Intuitively, a state satisfies $\langle a \rangle_p \varphi$ if, after having performed the action $a$, the probability of being in a state satisfying $\varphi$ is at least $p$. Later on, Desharnais, Edalat, and Panangaden [3] showed that negation is not necessary for discrimination purposes; the same result was subsequently reestablished by Jacobs and Sokolova [7] within the dual adjunction framework.

In this paper, we show that disjunction can be used in place of conjunction, without having to reintroduce negation. Thus, the constructs $\top$, $\lor$, and $\langle a \rangle_p$ suffice to characterize probabilistic bisimilarity on reactive probabilistic systems.

The idea of the proof is the following. First, 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 trees. These trees can be seen as the probabilistic counterpart of Winskel’s 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 a logic characterizes probabilistic bisimilarity, it suffices to prove that it allows to discriminate finite reactive probabilistic trees. This means that, given two different finite trees, we have to find a formula, which can be constructed by induction on the height of one of the trees, that tells them apart and has a depth not exceeding the height of the two trees.

Our technique is quite general and applies also to the probabilistic modal logics studied in the literature; in particular for the logics in [9] and [3] it allows us to provide simpler proofs of adequacy. More generally, this technique can be used in any computational model that has a compact fully abstract semantics.

Synopsis In Sect. 2, we recall the basic definitions about reactive probabilistic processes, bisimilarity, and logics. In Sect. 3, we characterize probabilistic bisimilarity in terms of finite reactive probabilistic trees. In Sect. 4, we prove that the various probabilistic modal logics considered in the paper can discriminate these finite trees, and hence characterize probabilistic bisimilarity. Conclusions and directions for future work are in Sect. 5.

2. Processes, Bisimilarity, and Logics
2.1 Reactive Probabilistic Processes and Bisimilarity
Probabilistic processes can be represented as labeled transitions systems [8] enriched with probabilistic information used to determine which action is executed or which state is reached. Following the terminology of [5], we focus on reactive probabilistic processes, where every state has at most one outgoing distribution for each action, and the choice among (differently labeled) distributions is nondeterministic. For a countable set $X$, the set of finitely supported (a.k.a. “simple”) probability distributions over $X$ is:

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

As usual, by “countable” we mean finite or countably infinite.
where the support is defined as \( \text{supp}(\Delta) \triangleq \{ x \in X \mid \Delta(x) > 0 \} \).

**Definition 2.1.** [RPLTS] A reactive probabilistic labeled transition system, RPLTS for short, is a triple \((S, A, \rightarrow)\) where:
- \(S\) is a countable set of states;
- \(A\) is a countable set of actions;
- \(\rightarrow \subseteq S \times A \times D(S)\) is a transition relation such that, whenever \((s, a, \Delta_1), (s, a, \Delta_2) \in \rightarrow\), then \(\Delta_1 = \Delta_2\).

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) and the numbers on them add up to 1. As usual, we denote \((s,a,\Delta) \in \rightarrow\) as \(s \xrightarrow{a} \Delta\), where the set of reachable states coincides with \(\text{supp}(\Delta)\). We also define cumulative reachability as \(\Delta(S') = \sum_{s' \in S'} \Delta(s')\) for all \(S' \subseteq S\). Probabilistic bisimilarity for the class of reactive probabilistic processes was introduced by Larsen and Skou [3].

**Definition 2.2.** [Probabilistic bisimilarity] Let \((S, A, \rightarrow)\) be an RPLTS. An equivalence relation \(\mathbb{R}\) over \(S\) is a probabilistic bisimulation iff, whenever \((s_1, s_2) \in \mathbb{R}\), then for all actions \(a \in A\):
- if \(s_1 \xrightarrow{a} \Delta_1\) then there exists \(s_2 \xrightarrow{a} \Delta_2\) such that \(\Delta_1(C) = \Delta_2(C)\) for all equivalence classes \(C \in S/B\);
- if \(s_2 \xrightarrow{a} \Delta_2\) then there exists \(s_1 \xrightarrow{a} \Delta_1\) such that \(\Delta_1(C) = \Delta_2(C)\) for all equivalence classes \(C \in S/B\).

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

### 2.2 Probabilistic Modal Logics

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

We are especially interested in probabilistic modal logics characterizing \(\sim_{PB}\). The logics we consider in this paper are similar to Hennessy-Milner logic [6], but the diamond modality is decorated with a probabilistic lower bound as follows:

\[
\begin{align*}
\text{PML}_{\sim} & : \phi \models T \iff \text{true} \\
\text{PML}_{\sim} & : \phi \models \neg \phi \iff \phi \not\models \phi \\
\text{PML}_{\sim} & : \phi \models \phi \land \phi \iff (\phi)_p \phi \\
\text{PML}_{\sim} & : \phi \models \phi \lor \phi \iff (\phi)_p \phi \\
\text{PML}_{\sim} & : (\phi)_p \phi \models T \iff \phi \models \phi \\
\text{PML}_{\sim} & : (\phi)_p \phi \models \neg \phi \iff \phi \not\models \phi \\
\text{PML}_{\sim} & : (\phi)_p \phi \models \phi \iff \phi \not\models \phi \\
\end{align*}
\]

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 as usual:

\[
\begin{align*}
s \models T & \iff \text{true} \\
ns & \models \neg \phi \\
s & \models \phi \land \phi \iff s = \phi_1 \text{ and } s = \phi_2 \\
s & \models \phi \lor \phi \iff s = \phi_1 \text{ or } s = \phi_2 \\
s & \models (\phi)_p \phi \iff \text{there exists } s \xrightarrow{a} \Delta \text{ such that } \Delta(s' \in S \mid s' \models \phi) \geq p \\
\end{align*}
\]

Larsen and Skou [9] proved that \(\text{PML}_{\sim}\) characterizes \(\sim_{PB}\). This holds true for \(\text{PML}_{\sim}\) as well, because \(\text{PML}_{\sim}\) is equivalent to \(\text{PML}_{\sim}\). Later on, Desharnais, Edalat, and Panangaden [3] proved that \(\text{PML}_{\sim}\) characterizes \(\sim_{PB}\) too, and hence negation is not necessary, a result subsequently reestablished by Jacobs and Sokolova [7] within the dual adjunction framework. The main aim of this paper is to show that \(\text{PML}_{\sim}\) suffices.

### 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.

#### 3.1 Coalgebras for Probabilistic Systems

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

\[
D(f) : D(X) \to D(Y) \quad D(f)(\Delta) = \lambda y. \Delta(f^{-1}(y))
\]

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

\[
B_{RP} : \text{Set} \to \text{Set} \quad B_{RP}(X) = (D(X) + 1)^A
\]

Indeed, given \(S = (S, A, \rightarrow)\), we define the corresponding coalgebra \((S, \sigma)\) as:

\[
\sigma : S \to B_{RP}(S) \quad \sigma(s) = \lambda a. \begin{cases} \Delta & \text{if } s \xrightarrow{a} \Delta \\ \text{otherwise} & \end{cases}
\]

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 \(\text{Coalg}(B_{RP})\) 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 \((S_1, \sigma_1)\) and \((S_2, \sigma_2)\) be \(B_{RP}\)-coalgebras. A relation \(\mathbb{R} \subseteq S_1 \times S_2\) is a \(B_{RP}\)-bisimulation iff there exists a coalgebra structure \(\rho : \mathbb{R} \to B_{RP}\mathbb{R}\) such that the projections \(\pi_1 : \mathbb{R} \to S_1\) and \(\pi_2 : \mathbb{R} \to S_2\) are homomorphisms (i.e., \(\sigma_1 \circ \pi_1 = B_{RP}(\pi_1 \circ \rho)\) for \(i = 1,2\)).

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

The following result shows that probabilistic bisimilarity corresponds to \(B_{RP}\)-bisimilarity.

**Proposition 3.2.** The probabilistic bisimilarity over an RPLTS \((S, A, \rightarrow)\) coincides with the \(B_{RP}\)-bisimilarity over the corresponding coalgebra \((S, \sigma)\).

**Proof.** An immediate consequence of [11] Lemma 4.4 and Thm. 4.5.

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 [11] 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; then the final coalgebra exists by the general result [10] Thm. 10.4).

Let \((Z, \zeta)\) be a 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.

**Proposition 3.4.** For all \(z_1, z_2 \in Z\): \(z_1 \sim z_2\) iff \(z_1 = z_2\).
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 $P_f$-coalgebra [12].

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

$$\frac{x \leq y}{x \leq z} \quad \text{in a partial order with a least element, called root, and for all } x \in X \text{ and } a \in A:$$

1. the set $\{y \in X \mid y \leq x\}$ is finite and well-ordered;
2. for all $(x_1, p_1), (x_2, p_2) \in \text{succ}(x,a)$, if $x_1 = x_2$ then $p_1 = p_2$;
3. for all $(x_1, p_1), (x_2, p_2) \in \text{succ}(x,a)$, if the subtrees rooted at $x_1$ and $x_2$ are isomorphic then $x_1 = x_2$;
4. if $\text{succ}(x,a) = 0$ then $\exists (\omega, p) \in \text{succ}(x,a) p = 1$.

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.

Reactive probabilistic trees are unordered trees where each node for each action has either no successors or a finite set of successors, which are labeled with positive real numbers that add up to 1; moreover, subtrees rooted at these successors are all different. See the forthcoming Fig. 1 for some examples. In particular, the trivial tree is $\text{nil} \triangleq (\{\cdot\}, x, a, \emptyset)$.

For $t = (X, \text{succ})$, we denote its root by $\perp$, its $a$-successors by $\beta(t) \triangleq \text{succ}^1(t,\perp)$, and the subtree rooted at $x \in X$ by $t[x] \triangleq (\{y \in X \mid x \leq y\}, \lambda y. \text{succ}(y, a))$; thus, $t[\perp] = \perp$.

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

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

where $\sup \emptyset = 0$; hence, $\text{height}(\text{nil}) = 0$. In particular, we denote by $RPT(t) \triangleq \{t \in RPT \mid \text{height}(t) < \omega\}$ 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. Since the resulting finite tree must be extensical, this operation has to collapse isomorphic subtrees resulting from the truncation. More formally:

**Definition 3.6.** For $n \in \mathbb{N}$, the pruning at $n$ is a function $\langle t \rangle_n : RPT \rightarrow RPT_f$ defined recursively on $n$ as follows, for $t = (X, \text{succ}) \in RPT$:

$$t|_0 \triangleq \text{nil} \quad t|_{n+1} \triangleq (\{\perp\} \cup Y, \text{succ'})$$

where:

$$Y = \bigcup Y' \mid ((X', \text{succ'}, p') \in q(t(a)), a \in A)$$

$$\text{succ'}(\perp, a) = (\{\perp, p'\} \mid (t', p') \in q(t(a)))$$

$$\text{succ'}(x,a) = \text{succ}(x,a) \text{ for } ((X',\text{succ'}), p') \in q(t(a)), x \in X'$$

and $q : \mathcal{P}_f(X \times \mathbb{R}_{\{0,1\}}) \rightarrow \mathcal{P}_f(RPT_f \times \mathbb{R}_{\{0,1\}})$ is a mutually defined auxiliary function collapsing isomorphic trees:

$$q(\emptyset) = \emptyset$$

$$q(S \uplus \{(x, p)\}) = \begin{cases} \{S' \uplus \{(t', p')\} \mid q(S) = S' \uplus \{(t', p')\} \text{ and } t' \cong t[x]_n \} & \text{if } q(S) \neq \emptyset \text{ and } q(S') \neq \emptyset \text{ and } q(S') \neq \emptyset \\
S' \uplus \{(t[x]_n, p)\} & \text{otherwise.} \end{cases}$$

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 reformatulate $B_{RP}$ in a slightly more “relational” format. We define a functor $D' : \text{Set} \rightarrow \text{Set}$ by letting for any set $X$:

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

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

**Proposition 3.7.** 1. $D' \cong D + 1$.
2. $D^A \cong B_{RP}$.
3. $\text{Coalg}(D^A) \cong \text{Coalg}(B_{RP})$.
4. The supports of the final $D^A$-coalgebra and of the final $B_{RP}$-coalgebra are isomorphic.

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

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

3: A $D^A$-coalgebra $(X, \sigma : X \rightarrow D'(X)^A)$ is mapped to $(X, \psi_X \circ \sigma : X \rightarrow B_{RP}(X))$; the vice versa is similar, using $\psi_X^{-1}$. 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 (up-to isomorphism). First, we observe that the set $RPT$ can be endowed with a $D^A$-coalgebra structure $\rho : RPT \rightarrow (D'(\mathcal{RP}))^A$ defined as follows, for $t = (X, \text{succ})$:

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

**Theorem 3.8.** (RP, $\rho$) is a final $B_{RP}$-coalgebra.

**Proof:** By Prop. 3.7 it suffices to prove that $(RPT, \rho)$ is the final $D^A$-coalgebra. To this end, we follow the construction given by Worrell in [12] Thm. 11. We define an ordinal-indexed final sequence of sets $(B_\alpha, \alpha)$ together with “projection functions” $f^\alpha : B_\beta \rightarrow B_\gamma$, $\gamma \leq \beta$:

$$B_0 = \{\text{nil}\} \cong 1 \quad f_0 = \emptyset$$

$$B_{\alpha+1} = D'(B_\alpha)^A, f_{\alpha+1} = D'(f_\alpha)^A$$

$$B_\lambda = \lim_{\alpha < \lambda} B_\alpha, f_\lambda = \pi_\alpha$$

for $\lambda$ a limit ordinal

the remaining $f_\alpha$ being given by suitable compositions. $D$ is $\omega$-accessible (because we restrict to finitely supported distributions), thus by [12] Thm. 13 and Prop. 3.7 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 $t = (t_0, t_1, \ldots)$ 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 $t$ the corresponding tree $u \in RPT$ is obtained by amalgamating $t$: 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 $t \in B_{\omega+\omega}$ 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 $\tau : B_{\omega+\omega} \rightarrow D'(B_{\omega+\omega})^A$ is given by $\tau(t)(a) = \{\{t\} \mid \forall i : t_i \in \text{succ}(t_i, a)\}$; therefore, $(B_{\omega+\omega}, \tau)$ and $(RPT, \rho)$ are isomorphic $D^A$-coalgebras, hence the thesis.
3.3 Full Abstraction and Compactness

By virtue of Thm. 3.8 given an RPLTS $S = (S, A, \rightarrow)$ there exists a unique coalgebra homomorphism $\phi : S \rightarrow RPT$, called the (final) semantics of $S$, which associates each state in $S$ with its behavior. This semantics is fully abstract.

Theorem 3.9. [Full abstraction] Let $(S, A, \rightarrow)$ be an RPLTS. For all $s_1, s_2 \in S$, $s_1 \sim_{PB} s_2$ iff $[s_1] = [s_2]$.

Proof It follows from Props. 3.2 and 3.4 and Thm. 3.8.

A key property of reactive probabilistic trees is that they are compact: two different trees can be distinguished by looking at their finite subtrees only. Let us formalize this principle:

Theorem 3.10. [Compactness] For all $t_1, t_2 \in RPT$, $t_1 = t_2$ iff for all $n \in \mathbb{N}$ : $t_1\mid_n = t_2\mid_n$.

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 $t_1\mid_n \neq t_2\mid_n$. Given a tree $t_n$, a finite path in $t_0$ is a sequence $(a_1, p_1, a_2, p_2, \ldots, a_n, p_n)$ such that for $i = 1, \ldots, n : (a_i, p_i) \in u_\rightarrow^{-1}(a_i)$ and $u_i = u[a_i]$. If $t_1 \neq t_2$, then there is a path of length $n$ in $t_1$, which cannot be rephrased into $t_2$ in $t_2$ we reach a tree $t_n$ such that for all $t$ : $(t, p_i) \notin t_n^{-1}(a_i)$. Therefore $t_1\mid_n \neq t_2\mid_n$.

Corollary 3.11. Let $(S, A, \rightarrow)$ be an RPLTS. For all $s_1, s_2 \in S$, $s_1 \sim_{PB} s_2$ iff for all $n \in \mathbb{N}$ : $[s_1]_n = [s_2]_n$.

4. Discriminating Power of PML$_\lambda$

By virtue of the construction in the previous section and Cor. 3.11 in order to prove that a modal logic characterizes $\sim_{PB}$ over reactive probabilistic processes, it is enough to show that it can discriminate all reactive probabilistic trees of finite height. However, a specific condition on the depth of distinguishing formulas has also to be satisfied, where $\text{depth}(\phi)$ is inductively defined as usual:

\[
\begin{align*}
\text{depth}(\top) &= 0 \\
\text{depth}(\lnot \phi) &= \text{depth}(\phi) \\
\text{depth}(\phi \land \phi_2) &= \max(\text{depth}(\phi), \text{depth}(\phi_2)) \\
\text{depth}(\phi \lor \phi_2) &= \max(\text{depth}(\phi), \text{depth}(\phi_2)) \\
\text{depth}(\Diamond \phi) &= 1 + \text{depth}(\phi)
\end{align*}
\]

Proposition 4.1. Let $L$ be one of the probabilistic modal logics in Sect. 2.2. If $L$ characterizes $\sim_{RPT}$ and for any two nodes $t_1$ and $t_2$ of an arbitrary RPT$_1$ model such that $t_1 \neq t_2$ there exists $\phi \in L$ distinguishing $t_1$ from $t_2$ such that:

\[
\text{depth}(\phi) \leq \max(\text{height}(t_1), \text{height}(t_2))
\]

then $L$ characterizes $\sim_{PB}$ over the set of RPLTS models.

Proof Given two states $s_1$ and $s_2$ of an RPLTS, if $s_1 \sim_{PB} s_2$ then for all $n \in \mathbb{N}$ it holds that $[s_1]_n = [s_2]_n$, thanks to Cor. 3.11 hence $s_1$ and $s_2$ satisfy the same formulas of $L$ because $L$ characterizes $\sim_{RPT}$. Suppose now that $s_1 \not\sim_{PB} s_2$ and then consider the minimal $n \in \mathbb{N}$ for which $[s_1]_n \neq [s_2]_n$. Then there exists $\phi \in L$ distinguishing $[s_1]_n$ from $[s_2]_n$ such that $\text{depth}(\phi) \leq \max(\text{height}(s_1)_n, \text{height}(s_2)_n) = n$, hence the same formula $\phi$ also distinguishes $s_1$ from $s_2$.

Notice that, in the proof above, if $\text{depth}(\phi)$ were greater than $n$ then, in general, $\phi$ may not distinguish higher prunings of $[s_1]_n$ and $[s_2]_n$, nor may any formula of depth at most $n$ and derivable from $\phi$ 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 $(a_1 \land (c_1) \land \lnot (d_1))$ 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 $(a_1 \lor (b_1)) \land \lnot (c_1)$ distinguishes $[s_1]_1$ from $[s_2]_1$, but this is no longer the case with the derived formula $(a_1 \lor (b_1))_1$ of depth 1 as both nodes satisfy it.

Based on the considerations above, in this section we show the main result of the paper: the logical equivalence induced by PML$_\lambda$ has the same discriminating power as $\sim_{PB}$.

This result is accomplished in three steps. Firstly, we demonstrate Larsen and Skou’s result for PML$_\lambda$ through $\sim_{RPT}$, which yields a proof that, with respect to the one in [36], 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. Secondly, we demonstrate that PML$_\lambda$ characterizes $\sim_{PB}$ by adapting the proof scheme to cope with the replacement of conjunction with disjunction. Thirdly, we demonstrate that PML$_\lambda$ characterizes $\sim_{PB}$ by further adapting the proof scheme to cope with the absence of negation.

Moreover, we demonstrate Desharnais, Edalat, and Panangaden’s result for PML$_\lambda$ through yet another adaptation of the proof scheme that, unlike the proof in [3], works directly with discrete space states.

4.1 PML$_\lambda$, Characterizes $\sim_{PB}$: A New Proof

To show that the logical equivalence induced by PML$_\lambda$ implies node equality $\simeq$, we reason on the contrapositive. 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$_\lambda$ 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 certain derivative $t'$ of $t_1$ (rather than the derivatives of $t_2$ different from $t'$), then take the conjunction of those formulas preceded by a diamond decorated with the probability for $t_1$ of reaching $t'$.

The 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 $\Delta_1(a) \neq \Delta_2(a)$ such that $\Delta_1(a) \neq \Delta_2(a)$, to obtain $\Delta_1(a)$ such that $\Delta_1(a)(t') > \Delta_2(a)(t')$, by the induction hypothesis there exists a PML$_\lambda$ formula $\phi_2$ that distinguishes $t'$ from a specific $\Delta_2(a)(t') \in \text{supp} \{ \Delta_2(a) \} \setminus \{ t' \}$. Thanks to the absence of negation in PML$_\lambda$, we can assume that $t' \not\vdash \phi_2 \land \phi_2 \not\vdash t_2$ because $\Delta_1(a)(t') > \Delta_2(a)(t')$ 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 in $\text{supp} \{ \Delta_1(a) \} \setminus \{ t' \}$ there might be other $\phi$-derivatives of $t_1$ that satisfy $\Delta_1(a)(t')$.

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

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 $\phi \in \text{PML}_\lambda$.
- Assuming that $t_1 \neq t_2$, we show that there exists $\phi \in \text{PML}_\lambda$, with $\text{depth}(\phi) \leq \max(\text{height}(t_1), \text{height}(t_2))$, such that it is not
Given main result of this paper, i.e., that $PML$ of this result will be useful to set up an outline of the proof of the certain derivative of $t_1$, we consider the case that after performing $\max(t'_1)$ with a consequence, $t_1$ and $t_2$ differ from each other, and $\phi_1$ becomes $\phi_2$ in $\phi_{12}$. Let $\phi_{12}$. If $\phi_{12} = \phi_1$ or $\phi_2$, then it holds that $t'_1 \not\equiv \phi_1 \iff t'_2 \not\equiv \phi_2$. Let $\phi_{12}$ be such that $\phi_{12} \not\equiv \phi_1$. Since $PML$ induces no loss of generality we can assume that $t'_1 \not\equiv \phi_1 \iff t'_2 \not\equiv \phi_2$. 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 $j = 1, 2, \ldots, k$ there exists $\phi_{2,j} \in PML_{\sim}$ with $\max(\phi_{2,1}) \leq \max(\height(t'_1), \height(t'_2))$, such that it is not the case that $t'_1 \not\equiv \phi_{2,j} \iff t'_2 \not\equiv \phi_{2,j}$.

4.3 PML$_{\sim}$ Characterizes $\sim_{PB}$

The proof that PML$_{\sim}$ characterizes $\sim_{PB}$ is inspired by the one for PML$_{\sim}$, thus considers the contrapositive and proceeds by induction. In the only non-trivial case, we will arrive at a situation in which $t_1 \not\equiv \langle a \rangle_{1-\Delta_2,(t'_1)+} \phi_{2,j} \equiv t_2$ for:

- a derivative $t'$ of $t_1$ such that $\Delta_1(a,t') > \Delta_2(a,t')$, which does not satisfy any subformula $\phi_{2,j}$;
- a suitable probabilistic value $p$ such that $\Delta_2(a,t') + p \leq 1$;
- an index set $J$ identifying certain derivatives of $t_2$ other than $t'$.

The choice of $t'$ is crucial, because negation is no longer available in PML$_{\sim}$, a fact that we are able to detect in the only non-trivial case for which $\sim_{PB}$ is available in PML$_{\sim}$. An important observation is that, in many cases, a disjunctive distinguishing formula can be obtained from a conjunctive one by suitably increasing some probabilistic lower bound of the latter.

**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:

\[
\begin{align*}
\phi_1 \land \phi_2 &\not\equiv \phi_3 \lor \phi_4, \\
\phi_1 \land \phi_2 &\not\equiv \phi_3 \lor \phi_4,
\end{align*}
\]

Notice that, when moving from the conjunctive formula to the disjunctive one, the probabilistic lower bound decorating the $\alpha$-diamond increases from 0.5 to 1 and the roles of $t_1$ and $t_2$ with respect to $\not\equiv$ are inverted.

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

\[
\begin{align*}
t_1 &\equiv \langle a \rangle_{0.5} \langle b \rangle_1 \land \langle c \rangle_1 \not\equiv t_2, \\
t_4 &\equiv \langle a \rangle_{0.9} \langle b \rangle_1 \lor \langle c \rangle_1 \not\equiv t_4
\end{align*}
\]

but the roles of $t_3$ and $t_4$ with respect to $\not\equiv$ 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.

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

\[
\begin{align*}
t_5 &\not\equiv \langle a \rangle_{0.5} \langle b \rangle_1 \land \langle c \rangle_3 \equiv t_6
\end{align*}
\]

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 PML$_{\sim}$ exists because:

\[
t_5 \not\equiv \langle a \rangle_{0.5} \langle b \rangle_1 \equiv t_6
\]

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 \( \sim_{PB} \), they must enable the same actions, so focusing on a single action is enough for discriminating when only disjunction is available.

Bearing this in mind, for any node \( t \) of finite height we define the set \( \Phi_{\vee}(t) \) of PML\( \vee \) formulas satisfied by \( t \) featuring:

- 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 \( \bigvee_{h \in H} \langle a_h \rangle_{\phi_h} \phi_h \), with \( a_{h_1} \neq a_{h_2} \) for \( h_1 \neq h_2 \), which are useless for discriminating with respect to \( \sim_{PB} \) 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 \( \text{height}(t) \) as follows:

- If \( \text{height}(t) = 0 \), then \( \Phi_{\vee}(t) = \emptyset \).
If $\text{height}(t) \geq 1$ for $t$ having transitions of the form $t \xrightarrow{a} \Delta$, with $\text{supp}(\Delta) = \{t'_j \mid j \in J\}$ and $i \in I \neq \emptyset$, then:

$$\Phi_\forall(t) = \{ (a_i)_i \mid i \in I \} \bigcup \left( \bigcup_{i \in I} \text{hpb}(\Phi_\forall, t_i) \bigcup \left\{ (a_i)_{i \in I} \bigcap \Delta_i(t'_j) \bigcap \bigvee_i \phi_{i,j,k} \mid t'_j \in \text{supp}(\Delta_i), \phi_{i,j,k} \in \Phi_\forall(t'_j) \right\} \right)$$

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

To illustrate the definition given above, we exhibit some examples showing the usefulness of $\Phi_\forall$-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 $\Phi_\forall$-set.

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

$$\Phi_\forall(t_7) = \{ (a_1), (a_1 \lor b_1) \}$$

$$\Phi_\forall(t_8) = \{ (a_1), (a_1 \lor b_1), (a_1 \lor c_1) \}$$

A formula like $(a_1) \lor (c_1)$, in which disjunction is between two actions enabled by the same node and hence constituting a nondeterministic choice, is useless for discriminating between $t_7$ and $t_8$. Indeed, such a formula is not part of $\Phi_\forall(t_8)$. 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 $(a_1 \lor c_1) \in \Phi_\forall(t_8)$ is a distinguishing formula can be retrieved as follows. Starting from the two identically labeled transitions $t_7 \xrightarrow{a} \Delta_{\Delta,a}$ and $t_8 \xrightarrow{a} \Delta_{\Delta,a}$ where $\Delta_{\Delta,a}(t'_5) = 1 = \Delta_{\Delta,a}(t_8')$, we have:

$$\Phi_\forall(t_7) = \{ (b_1) \}$$

$$\Phi_\forall(t_8) = \{ (b_1), (c_1) \}$$

If we focus on $t_7'$ because $\Delta_{a,a}(t'_7) > \Delta_{a,a}(t'_5)$ and its $\Phi_\forall$-set is minimal, then $t'_7 \not\equiv (c_1) \models t_7'$ with $(c_1) \in \Phi_\forall(t_7') \setminus \Phi_\forall(t_8)$. As a consequence, $t_7 \not\equiv (a_1 \lor c_1) \models t_7'$ where the value 1 decorating the $a$-diamond stems from $1 - \Delta_{a,a}(t'_7)$.

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

$$\Phi_\forall(t_1) = \{ (a_1), (a_0 \lor b_1), (a_0 \lor c_1) \}$$

$$\Phi_\forall(t_2) = \{ (a_1), (a_0 \lor b_1), (a_0 \lor c_1), (a_1 \lor b_1 \lor c_1) \}$$

The formulas with two diamonds and no disjunction are identical in the two sets, so their disjunction $\langle a_0, b_1 \lor a_0, c_1 \rangle$ is useless for discriminating between $t_1$ and $t_2$. Indeed, such a formula is part neither of $\Phi_\forall(t_1)$ nor $\Phi_\forall(t_2)$. In contrast, their disjunction in which decorations of identical diamonds are summed up, i.e., $(a_1 \lor b_1 \lor c_1)$, is fundamental. It belongs only to $\Phi_\forall(t_2)$ 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 $(a_1 \lor b_1 \lor c_1) \in \Phi_\forall(t_2)$ is a distinguishing formula can be retrieved as follows. Starting from the two identically labeled transitions $t_1 \xrightarrow{a} \Delta_{a,a}$ and $t_2 \xrightarrow{a} \Delta_{a,a}$ where $\Delta_{a,a}(t'_1) = 0.5 = \Delta_{a,a}(t'_2)$ and $\Delta_{a,a}(t_1') = 0.5 = \Delta_{a,a}(t_2')$, we have:

$$\Phi_\forall(t_1) = \{ (b_1), (c_1) \}$$

$$\Phi_\forall(t_2) = \{ (b_1) \}$$

$$\Phi_\forall(t_2) = \{ (c_1) \}$$

If we focus on $t'_1$ because $\Delta_{a,a}(t'_1) > \Delta_{a,a}(t'_2)$ and its $\Phi_\forall$-set is minimal, then $t'_1 \not\equiv (b_1) \models t_1'$ with $(b_1) \in \Phi_\forall(t'_1) \setminus \Phi_\forall(t'_2)$. As a consequence, $t_1 \not\equiv (a_1 \lor b_1 \lor c_1) \models t_1'$ where the value 1 decorating the $a$-diamond stems from $1 - \Delta_{a,a}(t'_1)$.

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

$$\Phi_\forall(t_5) = \{ (a_1), (a_0, b_1), (a_0, c_1), (a_0, 0, (b_1 \lor c_1)) \}$$

$$\Phi_\forall(t_6) = \{ (a_1), (a_0, b_1), (a_0, c_1) \}$$

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 $\Phi_\forall(t_6)$, 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 $\Phi_\forall(t_6)$.

The fact that $(a_0, b_1) \in \Phi_\forall(t_5)$ is a distinguishing formula can be retrieved as follows. Starting from the two identically labeled transitions $t_5 \xrightarrow{a} \Delta_{a,a}$ and $t_6 \xrightarrow{a} \Delta_{a,a}$ where $\Delta_{a,a}(t'_5) = 0.25 = \Delta_{a,a}(t'_6) = 0.25 = \Delta_{a,a}(t'_7)$, and $\Delta_{a,a}(t'_6) = 0 = \Delta_{a,a}(t'_8)$, we have:

$$\Phi_\forall(t_5) = \{ (b_1) \}$$

$$\Phi_\forall(t_6) = \{ (b_1), (c_1) \}$$

$$\Phi_\forall(t_7) = \emptyset$$

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 $\Delta_{a,a}(t''_5) > \Delta_{a,a}(t''_6)$ and its $\Phi_\forall$-set is minimal after the exclusion of $t''$, then $t''_5 \not\equiv (b_1) \models t''_5$ with $(b_1) \in \Phi_\forall(t''_5) \setminus \Phi_\forall(t''_6)$, while no distinguishing formula is considered with respect to $t''$ as element of $\text{supp}(\Delta_{a,a})$ due to the exclusion of $t''$ itself. As a consequence, $t_5 \not\equiv (a_0, b_1) \models t_5$ where the value 0.5 decorating the $a$-diamond stems from $1 - \Delta_{a,a}(t''_5) + p$ with $p = \Delta_{a,a}(t''_5)$. The reason for subtracting the probability that $t_5$ reaches $t''$ after performing $a$ is that $t'' \not\equiv (b_1)$.

We conclude by observing that focussing on $t''$ as derivative with the minimum $\Phi_\forall$-set is indeed problematic, because it would result in $(a_0, b_1)$ when considering $t''$ as derivative of $t_5$, but it would result in $(a_0, b_1 \lor c_1)$ when considering $t''$ as derivative of $t_6$, with the latter formula not distinguishing between $t_5$ and $t_6$. Moreover, when focussing on $t''$, no formula of was found such that $t'' \not\equiv \phi' \equiv t''$ as $\Phi_\forall(t'' \setminus \phi')$.

The last example shows that, in the general format for the PML $\forall$ distinguishing formula mentioned at the beginning of this subsection, i.e., $\langle a_1 \lor (\Delta_{a,a}(t'+p) \lor \bigvee_{j \in J} \phi_{k,j} \rangle$, 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 $\Phi_\forall$-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 $\bigvee_{j \in J} \phi_{k,j}$, note that the first node reaches those derivatives with the same probability $p$.

We present two additional examples illustrating some technicalities of Def. 4.2. The former example shows the usefulness of the operator $\lor$ and of the function $\text{hpb}$ for selecting the right $t'$ on the basis of the minimality of its $\Phi_\forall$-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 $\Phi_\forall$-set whose number of nested diamonds is not maximal.
EXAMPLE 4.11. For the nodes \( t_9 \) and \( t_{10} \) in Fig. 4(e), we have:

\[
\Phi_v(t_9) = \{ (a_1, (a_0,0.5,b_1), (a_0,0.5,c_1) ) \}
\]

\[
\Phi_v(t_{10}) = \{ (a_0,0.5,b_1), (a_0,0.5,c_1), (a_0,0.6,(b_1 \lor c_1)) \}
\]

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

\[
\Phi_v(t') = \{ (b_1), (c_1) \}
\]

\[
\Phi_v(t''') = \{ (b_1) \}
\]

\[
\Phi_v(t''') = \{ (c_1) \}
\]

If \( \omega \) is used in place of \( \Phi_v(t_{10}) \) and its \( \Phi_v \)-set is minimal, then \( t'' < t' \) with \( t'' \in \Phi_v(t') \setminus \Phi_v(t') \), \( t'' \not\subseteq t' \), \( t'' \not\subseteq \{ b_1 \} \) and \( (c_1) \in \Phi_v(t'') \setminus \Phi_v(t') \). As a consequence, \( t_9 \not\subseteq \{ a_0,0.6,(b_1 \lor c_1) \} \) and \( t_1 \) where the formula belongs to \( \Phi_v(t_{10}) \) and the value \( 0.6 \) decorating the a-diamond stems from \( 1 - \Delta_{10,a}(t') \).

EXAMPLE 4.12. For the nodes \( t_{11}, t_{12}, t_{13} \) in Fig. 4(f), we have:

\[
\Phi_v(t_{11}) = \{ (a_1) \}
\]

\[
\Phi_v(t_{12}) = \{ (a_1), (a_1, b_1) \}
\]

\[
\Phi_v(t_{13}) = \{ (a_1, a_0, b_1) \}
\]

Let us view them as derivatives of other nodes, rather than roots of trees. The presence of formula \( (a_1) \) in \( \Phi_v(t_{12}) \) and \( \Phi_v(t_{13}) \) – although it has not the maximum number of nested diamonds in those two sets – ensures the minimality of \( \Phi_v(t_{11}) \) and hence that \( t_{11} \) is selected for building a distinguishing formula. If \( (a_1) \) were not in \( \Phi_v(t_{12}) \) and \( \Phi_v(t_{13}) \), 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 \( \Phi_v \)-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 \( \Phi_v \)-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 \( \Phi_v \)-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 \( \Phi_v \)-sets there is a variant in the other \( \Phi_v \)-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 \( \Phi_v \)-sets, it has to be selected the derivative such that for each formula without disjunctions in its \( \Phi_v \)-set there exists a variant in the \( \Phi_v \)-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 \( \Phi_v \)-set of the selected derivative is \( < \) than the corresponding bound in the corresponding variant in the \( \Phi_v \)-set of the other derivative. We say that the \( \Phi_v \)-set of the selected derivative is a \((\leq, <)\)-variant of the \( \Phi_v \)-set of the other derivative.

EXAMPLE 4.13. Let us view the nodes \( t_9 \) and \( t_{10} \) in Fig. 4(e) as derivatives of other nodes, rather than roots of trees. Based on their \( \Phi_v \)-sets shown in Ex. 4.10, we need focus on \( t_9 \) because \( \Phi_v(t_{10}) \) contains fewer formulas. However, by so doing, we would be unable to find a distinguishing formula in \( \Phi_v(t_9) \) that is not satisfied by \( t_9 \). Indeed, if we look carefully at the formulas without disjunctions in \( \Phi_v(t_9) \) and \( \Phi_v(t_{10}) \), we note that they differ only for their probabilistic lower bounds: \( (a_1) \in \Phi_v(t_9) \) is a variant of \( (a_1) \in \Phi_v(t_{10}) \), \( (a_0,0.5,b_1) \in \Phi_v(t_9) \) is a variant of \( (a_0,0.5,b_1) \in \Phi_v(t_{10}) \), and \( (a_0,0.5,c_1) \in \Phi_v(t_9) \) is a variant of \( (a_0,0.5,c_1) \in \Phi_v(t_{10}) \). Therefore, we must focus on \( t_9 \) because \( \Phi_v(t_9) \) contains formulas without disjunctions such that \( \leq \) and \( < \) having smaller bounds: \( \Phi_v(t_9) \) is a \((\leq, <)\)-variant of \( \Phi_v(t_{10}) \).

Consider now the nodes \( t_9 \) and \( t_{10} \) in Fig. 4(e), whose \( \Phi_v \)-sets are shown in Ex. 4.11. If function \( hlp_b \) were not used and hence \( \Phi_v(t_{10}) \) also contained \( (a_1) \in (a_1) \in \Phi_v(t_9) \), \( (a_0,0.1,b_1) \in \Phi_v(t_9) \), and \( (a_0,0.1,c_1) \), then the formulas without disjunctions in \( \Phi_v(t_9) \) would no longer be equal to those in \( \Phi_v(t_{10}) \). More precisely, the formulas without disjunctions would be similar between the two sets, with those in \( \Phi_v(t_{10}) \) having smaller probabilistic lower bounds, so that we would erroneously focus on \( t_{10} \).

Summing up, in the construction of the PMLV distinguishing formula mentioned at the beginning of this subsection, i.e., \( \langle a_1 \rangle \cdot (\Delta_{a_1}(t_1) + \bigvee_{j \in J} \phi_2 \cdot j \rangle \) the steps for choosing the derivative \( t' \), on the basis of which each subformula \( \phi_2 \cdot j \) is determined so that it is not satisfied by \( t' \), are the following:

1. Consider only derivatives to which the first distribution assigns a probability greater than the second one.
2. Within the previous set, eliminate all the derivatives whose \( \Phi_v \)-sets have \((\leq, <)\)-variants.
3. Among the remaining derivatives, focus on one of those having a minimal \( \Phi_v \)-set.

THEOREM 4.14. Let \((T, A, \rightarrow)\) be in \( RPT \) and \( t_1, t_2 \in T \).

Then \( t_1 = t_2 \) iff \( t_1 \models \phi \iff t_2 \models \phi \) for all \( \phi \in PML\).

Moreover, if \( t_1 \neq t_2 \), there then exists \( \phi \in PMLv \) distinguishing \( t_1 \) from \( t_2 \) such that \( depth(\phi) \leq \max(\text{height}(t_1),\text{height}(t_2)) \).

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 \( \phi \in PMLv \).
- Assuming that \( t_1 \neq t_2 \), we show that there exists \( \phi \in \Phi_v(t_1) \cup \Phi_v(t_2) \), which ensures that \( depth(\phi) \leq \max(\text{height}(t_1),\text{height}(t_2)) \), such that it is not the case that \( t_1 \models \phi \iff t_2 \models \phi \) by proceeding by induction on \( \text{height}(t_1) \in N \).
- The proof is similar to the one of Thm. 4.3 in particular in the cases \( \text{height}(t_1) = 0 \) and \( \text{height}(t_1) = n + 1 \) with \( \text{init}(t_1) \neq \text{init}(t_2) \).

However, it changes as follows before the application of the induction hypothesis in the case \( \text{height}(t_1) = n + 1 \) with \( \text{init}(t_1) = \text{init}(t_2) \neq \emptyset \) 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 \( a \in \text{init}(t_1) \).

Let \( \text{supp}_\phi = \text{supp}(\Delta_{1,a}) \cup \text{supp}(\Delta_{2,a}) \), which can be partitioned into \( \text{supp}_\phi = \{ t' \in \text{supp}_\phi \mid \Delta_{1,a}(t') \neq \Delta_{2,a}(t') \} \).
We recall that $\Phi$ such that, for all arbitrary node belonging to $\text{supp}|_N$ disjunctions in at least one probabilistic lower bound in a formula without for each formula without disjunctions in $\Phi_{t_2}$ denote by $\text{supp}$.

Suppose that $\varphi$ enjoys the property over the entire set $\{t_1, t_2, \ldots, t_{n+1}\}$, we denote by $t'$ the node in $\text{supp}|_N$ that, by the induction hypothesis, enjoys the property over that subset. There are two cases:

- If $\Phi_{t'(n'+1)}$ is not a $(\leq, <)$-variant of $\Phi_{t'}$ (if, then $t'$ enjoys the property over the entire set $\text{supp}|_N$.

- Suppose that $\Phi_{t'(n'+1)}$ is a $(\leq, <)$-variant of $\Phi_{t'}$, which implies that $\Phi_{t'}$ cannot be a $(\leq, <)$-variant of $\Phi_{t'(n'+1)}$. From the fact that, for all $t'' \in \text{supp}|_N \setminus \{t', t'(n'+1)\}$, $\Phi_{t''}$ (if $t''$ is not a $(\leq, <)$-variant of $\Phi_{t'}$, then $\Phi_{t'}$ contains at least a formula without disjunctions that is not a variant of any formula without disjunctions in $\Phi_{t'}$, or all formulas without disjunctions in $\Phi_{t'}$ are identical to formulas without disjunctions in $\Phi_{t'}$, hence this holds true with respect to $\Phi_{t'(n'+1)}$, too, given that $\Phi_{t'(n'+1)}$ is a $(\leq, <)$-variant of $\Phi_{t'}$. As a consequence, $t_{n'+1}$ enjoys the property over the entire set $\text{supp}|_N$.

Within the set of all the nodes in $\text{supp}|_N$ enjoying the property above, we select one with a minimal $\Phi_{t'}$-set, which we denote by $t_{\min}$. Suppose that $\Delta_{t', \text{supp}}(t_{\min}) > \max_{t'}\Delta_{t', \text{supp}}(t_{\min})$ and let $t'_{2,j}$ be an arbitrary node belonging to $\text{supp}|_N$, then such a formula can be taken as $\Phi_{t'_2}$ given the maximality of the probabilistic lower bounds of any basic formula in $\Phi_{t'_2}$. We recall that $\Phi_{t'_2}$ (if $t'_2$ is not a $(\leq, <)$-variant of any formula without disjunctions in $\Phi_{t'_2}$).

- If at least one formula without disjunctions in $\Phi_{t'_2}$ is not a variant of any formula without disjunctions in $\Phi_{t'_2}$, then such a formula can be taken as $\Phi_{t'_2}$ given the maximality of the probabilistic lower bounds of any basic formula in $\Phi_{t'_2}$.

- If all basic formulas in $\Phi_{t'_2}$ are identical to basic formulas in $\Phi_{t'_2}$, then $\Phi_{t'_2}$ must contain some more formulas (with disjunctions) not in $\Phi_{t'_2}$ 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 $\Phi_{t'_2}$ can be taken as $\Phi_{t'_2}$.

Letting $\text{supp}|_N = \{t' \in \text{supp}|_N : t' \not\in \bigcup_{i,j,k} \text{spbl}(\{a_i\}_{\Delta_{t', \text{supp}}(t'_{2,j})}) \cap \text{supp}_{a_i,j,k}(\{t_{\min}\}) \}$ where $\cup$ and $\cap$ are multiset parentheses, $K_{i,j}$ is the index set for $\Phi_{t'_2}$, and function $\text{spbl}$ merges all formulas possibly differing 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 $\Delta_{t'}$).

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 PML$\lambda$ distinguishing formula at the beginning of this subsection is the maximality of the $\Phi_{t'}$-set.

Example 4.15. In Fig.4(b), the multiset giving rise to $\Phi_{t_3}$ contains two occurrences of $(a_0.1)_1(b_1)$ and two occurrences of $(a_0.1.0)_1(b_1)$, which are merged into $(a_0.0.0.0)_1(b_1)$ by function $\text{spbl}$. Likewise, the multiset behind $\Phi_{t_4}$ contains formulas $(a_0.1.1)_1(b_1)$, $(a_0.3.1)_1(b_1)$, and $(a_0.2.2)_1(b_1)$, which are merged into $(a_0.0.0)_1(b_1)$.
EXAMPLE 4.16. For the nodes $t_1$ and $t_2$ in Fig. 1(a), we have:

$$\Phi_{t_1}(t_1) = \{(a_1), (a_0, 0.5) (b_1), (a_0, 0.5) (c_1), (a_0, 0.5) (b_1) \land (c_1)\}$$

$$\Phi_{t_2}(t_2) = \{(a_1), (a_0, 0.5) (b_1), (a_0, 0.5) (c_1)\}$$

The conjunction $(a_0, 0.5) (b_1) \land (a_0, 0.5) (c_1)$ is useless for discriminating between $t_2$ and $t_2$ — it is part of neither $\Phi_{t_1}(t_1)$ nor $\Phi_{t_2}(t_2)$ — while $(a_0, 0.5) (b_1) \land (c_1)$ is the only distinguishing formula and belongs only to $\Phi_{t_1}(t_1)$, 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 $t_2 \xrightarrow{a} \Delta_{1,0}^a$ and $t_2 \xrightarrow{a} \Delta_{2,0}$ where $\Delta_{1,0}(t_2) = 0.5 = \Delta_{2,0}(t_2)$ and $\Delta_{1,0}(t_2) = \Delta_{1,0}(t_2) = 0 = \Delta_{2,0}(t_2)$, we have:

$$\Phi_{t_1}(t_2) = \{(b_1), (c_1)\}$$

$$\Phi_{t_2}(t_2) = \{(b_1)\}$$

If we focus on $t_1'$ because $\Delta_{1,0}(t_1') > \Delta_{2,0}(t_1)$ and its $\Phi_{t_1}$-set is maximal, then $t_1' = (a_0) \not\equiv t_2$ with $(a_0) \in \Phi_{t_1}(t_1') \setminus \Phi_{t_2}(t_2)$ as well as $t_1' = (b_1) \not\equiv t_2$ with $(b_1) \in \Phi_{t_1}(t_1') \setminus \Phi_{t_2}(t_2)$. As a consequence, $t_1' = (a_0, 0.5) (b_1) \land (c_1) \not\equiv t_2$ where the value 0.5 decorating the $c$-diamond stems from $\Delta_{1,0}(t_1')$.

As far as the other two variables occurring in the PML$_\Lambda$ distinguishing formula at the beginning of this subsection are concerned, $\not\equiv$ only contains any derivative of the second node different from $t'$ to which the two distributions assign two different probabilities, while $\not\equiv$ is the probability of reaching derivatives having the same probability in both distributions that satisfy $\bigwedge_{j \in J} \phi_{d,j}$. Therefore, when selecting $t'$, we have to leave out all the derivatives whose $\Phi_{t_1}$-sets have $(\leq, <)$-variants.

THEOREM 4.17. Let $(T, A, \xrightarrow{r})$ be in RPT$_T$ and $t_1, t_2 \in T$. Then $t_1 = t_2$ if and only if $t_1 = \phi \iff t_2 = \phi$ for all $\phi \in$ PML$_\Lambda$. Moreover, if $t_1 \not\equiv t_2$, then there exists $\phi \in$ PML$_\Lambda$ distinguishing $t_1$ from $t_2$ such that $\text{depth}(\phi) \leq \max(\text{height}(t_1), \text{height}(t_2))$.

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

- We select $t_{\text{max}}$ as one of the nodes with maximal $\Phi_{t_1}$-set in $\text{supp}_{\not\equiv} \phi$ having no $(\leq, <)$-variants.

- It holds that $t_{\text{max}} = \phi_{d, j} \not\equiv t_{\text{max}}$ for all $t_{\text{max}} \in \text{supp}_{\not\equiv}$ because $\phi_{d, j} \in \Phi_{t_1}(t_{\text{max}})$.

- Letting $\text{supp}_{\not\equiv, \text{max}} = \{t' \in \text{supp}_{\not\equiv} = \{t' = (a_0, 0.5) (b_1) \land (c_1)\} \not\equiv t_2$ because $\Delta_{1,0}(t_{\text{max}}) + p_{\text{min}} > \Delta_{2,0}(t_{\text{max}}) + p_{\text{min}}$ and the maximum probabilistic lower bound for which $t_2$ satisfies a formula of that form cannot exceed $\Delta_{2,0}(t_{\text{max}}) + p_{\text{min}}$.

- The PML$_\Lambda$ distinguishing formula is in $\Phi_{t_1}(t_1)$ due to $\not\equiv$.

5. Conclusions

In this paper, we have studied modal logic characterizations of bisimilarity over reactive probabilistic processes. Starting from previous work by Larsen and Skou [9] (who provided a characterization based on an extension of Hennessy-Milner modal logic where diamonds are decorated with probabilistic lower bounds) and by Desharnais, Edalat, and Panangaden [8] (who showed that negation is not necessary), we have proved that conjunction can be replaced by disjunction without having to reintroduce negation. Thus, in our probabilistic setting conjunction and disjunction are interchangeable to characterize bisimulation equivalence, which they are both necessary in the case of simulation preorder [4]. As a side result, using the same proof technique we have provided new (and simpler) proofs of the expressivity of the logics PML$_\not\equiv$, and PML$_\equiv$.

The intuition behind our result is that from a conjunctive distinguishing formula it is often possible to derive a disjunctive one by suitably increasing some probabilistic lower bounds contained in the former. On the model side, this corresponds to summing up the probabilities of reaching certain states that are in the support of a target distribution. More generally, the final coalgebra of the behavioral functor $\mathcal{B}$ can be endowed with a topology, obtained from the base of clopen sets $\{[p, 1) \mid 0 < p < 1\}$ for the interval $[0, 1]$, which induces the diamond modality $\Diamond$. In this view, the final coalgebra can be seen as a coalgebra over trees of clopen sets. Thus, each state of the final coalgebra is characterized by the countable set of formulas obtained by doing finite visits of its tree; in fact, this set is approximated by the $\Phi_{t_1}$-set introduced in Def. 4.17.

As far as the application of our result is concerned, the disjunctive modal logic characterization of bisimilarity over reactive probabilistic processes may simplify the proof of the second conjecture contained in [2]. This work investigates the discriminating power of three different testing equivalences respectively based on reactive probabilistic tests, fully nondeterministic tests, and nondeterministic and probabilistic tests. The above-mentioned conjecture refers to the fact that numerous examples indicate that testing equivalence based on nondeterministic and probabilistic tests may have the same discriminating power as bisimilarity over reactive probabilistic processes. Given two processes not related by $\sim_{\text{PRP}}$, the idea of the tentative proof of the conjecture is to build a distinguishing nondeterministic and probabilistic test from a distinguishing PML$_\Lambda$ formula. The new disjunctive modal logic characterization based on PML$_\Lambda$ may help to overcome one of the main difficulties with carrying out such a proof, i.e., that choices within tests fit well together with disjunction rather than conjunction.

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] 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.

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

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

[5] 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.

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

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

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

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

[10] J.J.M.M. Rutten. Universal coalgebra: A theory of systems. Theoretical Computer Science, 249:3–90, 2000.

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

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