# Expressiveness of Probabilistic Modal Logics

**Nathanaël Fijalkow**[*1], **Bartek Klin**[†2], **and Prakash Panangaden**[‡3]

1 **University of Warwick, Warwick, UK**
nfijalkow@turing.ac.uk
2 **University of Warsaw, Warsaw, Poland**
klin@mimuw.edu.pl
3 **McGill University, Montreal, Canada**
prakash@cs.mcgill.ca

## Abstract

Labelled Markov processes are probabilistic versions of labelled transition systems. In general, the state space of a labelled Markov process may be a continuum. Logical characterizations of probabilistic bisimulation and simulation were given by Desharnais et al. These results hold for systems defined on *analytic* state spaces and assume that there are countably many labels in the case of bisimulation and finitely many labels in the case of simulation.

In this paper, we first revisit these results by giving simpler and more streamlined proofs. In particular, our proof for simulation has the same structure as the one for bisimulation, relying on a new result of a topological nature. This departs from the known proof for this result, which uses domain theory techniques and falls out of a theory of approximation of Labelled Markov processes.

Both our proofs assume the presence of countably many labels. We investigate the necessity of this assumption, and show that the logical characterization of bisimulation may fail when there are uncountably many labels. However, with a stronger assumption on the transition functions (continuity instead of just measurability), we can regain the logical characterization result, for arbitrarily many labels. These new results arose from a new game-theoretic way of understanding probabilistic simulation and bisimulation.

**1998 ACM Subject Classification** F.1.2 [Modes of Computation] Probabilistic Computation, F.4.1 Mathematical Logic

**Keywords and phrases** probabilistic modal logic, probabilistic bisimulation, probabilistic simulation

**Digital Object Identifier** 10.4230/LIPIcs.ICALP.2017.105

## 1 Introduction

It is now 40 years since the logical characterization of bisimulation was established by van Benthem [14] and by Hennessy and Milner [10] for nondeterministic transition systems. It was shown that two states (or processes) are bisimilar if and only if they satisfied the same formulas of a modal logic that has come to be called Hennessy-Milner logic. The probabilistic version was studied by Larsen and Skou [12] who defined probabilistic bisimulation for discrete probabilistic transition systems and established a logical characterization theorem for discrete systems with a strong finite-branching assumption.

* Supported by the Alan Turing Institute under the EPSRC grant EP/N510129/1.
† Supported by the Polish National Science Centre (NCN) grant 2012/07/E/ST6/03026.
‡ Supported by a grant from the Natural Sciences and Engineering Research Council of Canada.


44th International Colloquium on Automata, Languages, and Programming (ICALP 2017).
Editors: Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl;
Article No. 105; pp. 105:1–105:12

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The analysis of bisimulation was extended to continuous state spaces by Blute et al. [4] and a logical characterization was shown by Desharnais et al. [6, 7] who proved the result without using any negative constructs in the logic nor any kind of finite branching assumptions. Their proofs marked a departure from the usual combinatorial arguments and used some non-trivial results from measure theory, specifically using remarkable properties of analytic spaces; see [13] for an expository account.

The fact that the logical characterization result can be established with a purely positive logic was a surprise at the time. It opened the door to the possibility that there could be a logical characterization of *simulation*; we define this precisely below but the idea should be intuitively clear. A clever example, due to Josée Desharnais [8], showed that this cannot be done with the same logic as the one used for bisimulation; one needs to have disjunction in the logic. A logical characterization of simulation was proved [8] for transition systems with *finitely many* labels. The main contribution of [8] was approximation theory which included a domain-theoretic treatment; the logical characterization result fell out of the domain theory results. Desharnais [5] in her thesis gave a proof not using domain theory in the discrete case. What remained unknown until now is a proof that works for countably many labels, continuous state spaces and, if possible, not using domain theory. We provide such a result, extending the characterization for simulation to countably many labels with a proof very much analogous to the one given for bisimulation.

## 1.1 Results

1. We give a characterization of bisimulation and simulation in terms of Spoiler/Duplicator games. This is elementary but interesting: it was the main driver of the intuitions that led to the proofs of the present paper though, in the end, one does not actually need the games to establish the results.
2. The logical characterization of bisimulation has a proof which has a structure which can be boiled down to two main ingredients: Dynkin's $\pi$-$\lambda$ theorem and the Unique Structure Theorem for analytic spaces. For simulation, it turns out that a completely analogous proof exists. It is enough to replace the two main ingredients by new positive versions: a positive analogue of the monotone class theorem and a positive UST, both of which we prove. This simplifies the previous domain-theoretic proof and clarifies the picture. The small price to pay is to move from analytic spaces to Polish ones; moving back to analytic is future work.
3. Both proofs rely on the countability of the set of formulas. This is necessary, as an explicit counterexample shows. But if the transition structure is continuous, logical characterization results are regained for arbitrary sized sets of labels. As far as we know, this is the first result of this type for uncountable label spaces.

Both logical characterization proofs, for bisimulation and simulation, have a similar structure and can even be said to follow the same *top-level* strategy as the original Hennessy-Milner proof.

# 2 Probabilistic systems and logics

We review some definitions and concepts from measure theory and topology. We assume that the reader is familiar with concepts like: $\sigma$-algebra, measurable functions, measures, topology and continuity.

Given a topological space $X$ the $\sigma$-algebra induced by its open sets (or its closed sets) is called the *Borel algebra*; we will always work with Borel algebras of topological spaces. We call them Borel spaces.

A topological space is said to be *separable* if it has a countable dense subset. For metric spaces this is equivalent to having a countable base of open sets. A Polish space is the topological space underlying a complete separable metric space. Note that a space like $(0,1)$ which is not complete in its usual metric is nevertheless Polish, since it can be given a complete metric that produces the same topology. If $X, Y$ are Polish spaces and $f : X \to Y$ is a continuous function then the image $f(X) \subset Y$ is an *analytic* space. The class of analytic spaces is not altered if we allow $f$ to be measurable instead of continuous or if we take the image of a Borel set instead of all of $X$.

▶ **Definition 1.** A **Markov kernel** on a Borel space $(X, \Sigma)$ is a function $\tau : X \times \Sigma \to [0,1]$ such that for each fixed $x \in X$, the set function $\tau(x, \cdot)$ is a sub-probability measure, and for each fixed $C \in \Sigma$ the function $\tau(\cdot, C)$ is a measurable function.

One interprets $\tau(x, C)$ as the probability of the process starting in state $x$ making a transition into one of the states in $C$.

▶ **Definition 2.** A **labelled Markov process** (LMP) $\mathcal{S}$ with label set $\mathcal{A}$ is a structure $(X, \Sigma, \{\tau_a \mid a \in \mathcal{A}\})$, where $(X, \Sigma)$ is a Borel space and

$$\tau_a : X \times \Sigma \longrightarrow [0,1]$$

is a Markov kernel for each $a \in \mathcal{A}$.

A key concept is bisimulation. The following definition is adapted from Larsen and Skou [12] to deal with measurability issues.

▶ **Definition 3** (Bisimulation)**.** Let $\mathcal{S} = (X, \Sigma, \tau)$ be a labelled Markov process. An equivalence relation $R$ on $X$ is a **bisimulation** if whenever $xRy$, with $x, y \in X$, we have that for all $a \in \mathcal{A}$ and every $R$-closed measurable set $C \in \Sigma$, $\tau_a(x, C) = \tau_a(y, C)$. We say that $x$ and $y$ are bisimilar, denoted $x \approx y$, if there exists a bisimulation $R$ such that $xRy$.

The modal logic $\mathcal{L}_\wedge$ used in the logical characterization theorem of [7] is generated by the grammar:

$$\phi ::= \top \mid \phi \wedge \phi \mid \langle a \rangle_p \phi$$

where $p$ ranges over rational numbers between 0 and 1. A state $x$ satisfies the modal formula $\langle a \rangle_p \phi$ if there exists a measurable subset $C$ with every state in $C$ satisfying $\phi$ and $\tau_a(x, C) > p$. It is easy to show that the sets defined by formulas $[\![\phi]\!] := \{x \mid x \models \phi\}$ are all measurable. We write $x \equiv_\wedge y$ to say that $x$ and $y$ satisfy the same formulas in $\mathcal{L}_\wedge$.

The logical characterization theorem for probabilistic bisimulation is:

▶ **Theorem 4** ([7])**.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is analytic and $\mathcal{A}$ is countable, and for any $x, y \in X$, we have that $x \equiv_\wedge y$ if and only if $x \approx y$.* ◀

For a preorder $R$ on a set $X$, we say that $C \subseteq X$ is *$R$-closed* if $x \in C$ and $xRy$ implies $y \in C$, for all $x, y \in X$.

▶ **Definition 5** (Simulation)**.** Let $\mathcal{S} = (X, \Sigma, \tau)$ be a labelled Markov process. An preorder relation $R$ on $X$ is a **simulation** if whenever $xRy$, with $x, y \in X$, we have that for all $a \in \mathcal{A}$ and every $R$-closed measurable set $C \in \Sigma$, $\tau_a(x, C) \leq \tau_a(y, C)$. We say that $x$ is simulated by $y$, denoted $x \lesssim y$, if there exists a simulation $R$ such that $xRy$.

The logic $\mathcal{L}_{\vee\wedge}$ extends $\mathcal{L}_{\wedge}$ with disjunction:

$$\phi \;=\; \top \;\mid\; \phi \wedge \phi \;\mid\; \phi \vee \phi \;\mid\; \langle a \rangle_p \phi.$$

We write $x \leq_{\vee\wedge} y$ to say that every formula in $\mathcal{L}_{\vee\wedge}$ satisfied by $x$ is also satisfied by $y$.

The previous logical characterization theorem for probabilistic simulation is:

▶ **Theorem 6** ([8]). *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is analytic and $\mathcal{A}$ is finite, and for any $x, y \in X$, we have that $x \leq_{\vee\wedge} y$ if and only if $x \lesssim y$.*

Existing proofs of Theorems 4 and 6 span several pages each, and are markedly dissimilar. In particular, the latter relies on the machinery of domain theory. One of our main contributions is to provide new, short proofs of both results.

## 3 Probabilistic (bi)simulation games

The classical notion of bisimulation and simulation for nondeterministic processes has a simple and elegant characterization in terms of games. These games, played between two players named Spoiler (who tries to prove that some two states in a process are not bisimilar) and Duplicator (who claims the opposite), provide convenient intuitions about the essence of bisimilarity.

To the best of our knowledge, probabilistic bisimulation and simulations have not been characterized by games before. In this section we fill this gap; as we shall see, the relevant games have a pleasantly simple structure, even in the setting of continuous space processes.

We begin with the case of **bisimulation game**. As in the classical case, we consider a spoiler/duplicator game with two players. Duplicator's plays are pairs of states that she claims are bisimilar. Spoiler tries to show that a given pair of states proffered by Duplicator are not bisimilar. Let $\mathcal{S} = (X, \Sigma, \tau)$ be a labelled Markov process, and $x, y \in X$. The bisimulation game starting from the position $(x, y)$ alternates between moves of the following kinds:

- Spoiler chooses $a \in A$ and $C \in \Sigma$ such that $\tau_a(x, C) \neq \tau_a(y, C)$,
- Duplicator answers by choosing $x' \in C$ and $y' \notin C$ and the game continues from $(x', y')$.

A player who cannot make a move at any point loses. Duplicator wins if the game goes on forever.

Note that the only way for Spoiler to win is to choose $C = X$ at some point; otherwise Duplicator can always choose some $x'$ and $y'$ as prescribed. (The only other situation where Duplicator cannot proceed would be $C = \emptyset$, but that is not a legal move for Spoiler since always $\tau_a(x, \emptyset) = \tau_a(y, \emptyset) = 0$.) On the other hand, Duplicator can win either by forcing an infinite play or by reaching a position $(x, y)$ where $\tau_a(x, C) = \tau_a(y, C)$ for every $C \in \Sigma$.

The intuition behind the game should be clear. Spoiler tries to prove that states $x$ and $y$ are not bisimilar by showing a label $a$ and a set $C$, purportedly closed under bisimilarity, such that the probabilities of $a$-labelled transitions to $C$ are different for $x$ and $y$. This difference of probabilities is checked but not disputed by Duplicator, who instead claims that $C$, in fact, is not closed under bisimilarity. She does that by choosing $x' \in C$ and $y' \notin C$ and claiming that these two are bisimilar; the game then proceeds in the same fashion.

Before we formally prove the correctness of this game, let us spend a moment to consider what makes a game-theoretic characterization "elegant". In our opinion, the classical bisimulation game for nondeterministic processes is elegant because it allows one to characterize a global property of behaviours (bisimilarity) in terms of a game whose rules only depend on local considerations. Indeed, whether a move in the game is legal or not does not depend on

bisimilarity or other long-range properties, but merely on local observations about transition capabilities that cannot be disputed by either player.

We argue that this criterion of elegance is satisfied by our probabilistic game. One can imagine the players engaging in a brief experiment with the given Markov process after each move by Spoiler, to determine that the two transition probabilities involved are indeed different. By performing random $a$-transitions from $x$ and $y$ sufficiently many times, Spoiler can demonstrate to Duplicator, with an arbitrarily high confidence level, that the probabilities of getting to $C$ are different and so that the move to $C$ is legal for Spoiler. It is important to note, comparing the game to the definition of probabilistic bisimulation itself, that the legality of a Spoiler's move does not depend on the set $C$ being actually closed under bisimilarity; a game with such a condition would not be "elegant".

The question of *how many* random transitions are enough to convince Duplicator that a Spoiler's move is legal, and hence how much time it takes for Spoiler to win the game if $x$ and $y$ are not bisimilar, suggests a potentially interesting connection of the bisimulation game to the quantitative framework of metrics on labelled Markov processes [9]. We leave this for future work.

Back to formal development. Since all infinite plays are won by the same player (Duplicator), standard game-theoretic arguments prove that:

▶ **Fact 7.** *The bisimulation game is determined, i.e., from every position $(x, y)$ either Spoiler has a winning strategy or Duplicator does.* ◀

From this we infer:

▶ **Theorem 8.** *The states $x$ and $y$ are bisimilar if and only if Duplicator has a winning strategy from $(x, y)$.*

**Proof.** For the left-to-right implication, for bisimilar $x$ and $y$, we construct a winning strategy from $(x, y)$ for Duplicator. In this strategy, for any move $a$ and $C$ by Spoiler, Duplicator chooses some arbitrary $x' \in C$ and $y' \notin C$ such that $x'$ and $y'$ are bisimilar. This is always possible: since Spoiler's move was legal, and it originated from a pair of bisimilar states, $C$ cannot be closed under bisimilarity. This strategy is winning for Duplicator since it allows her response to any move by Spoiler, and Duplicator wins all infinite plays.

For the right-to-left implication, we shall show that the set $R$ of all pairs $(x, y)$ whence Duplicator has a winning strategy, is a bisimulation. To this end, first we need to show that $R$ is an equivalence relation. Reflexivity is immediate, since from a position $(x, x)$ Spoiler has no legal moves. For symmetry, given a winning strategy from $(x, y)$ Duplicator builds a strategy from $(y, x)$ trivially: she simply replies to any first move by Spoiler as if she would reply to a move from $(x, y)$, and then she follows the given strategy. For transitivity, assume winning strategies for Duplicator from $(x, y)$ and $(y, z)$. A winning strategy for $(x, z)$ works as follows: for a legal move $a$ and $C$ from Spoiler, it must be that $\tau_a(x, C) \neq \tau_a(y, C)$ or $\tau_a(y, C) \neq \tau_a(z, C)$. Depending on which of these inequalities holds, reply according to the strategy from $(x, y)$ or from $(y, z)$, and then follow that winning strategy.

Now assume towards contradiction that $R$ is not a bisimulation. This means that for some $x, y$ such that $xRy$, there exists a letter $a$ in $A$ and an $R$-closed subset $C$ of $X$ such that $\tau_a(x, C) \neq \tau_a(y, C)$. Consider $a$ and $C$ as a Spoiler's move from $(x, y)$. No matter what Duplicator chooses as $x' \in C$ and $y' \notin C$, since $C$ is $R$-closed we have that not $(x'Ry')$ and, by Fact 7, Spoiler has a winning strategy from $(x', y')$. This forms a winning strategy for Spoiler from $(x, y)$, contradicting the assumption that $xRy$. ◀

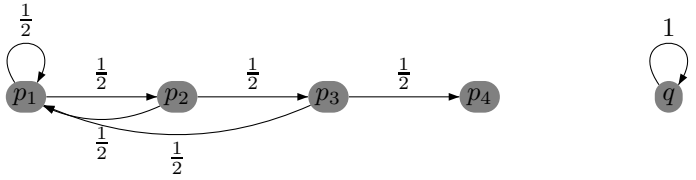


**Figure 1** It takes four steps for Spoiler to convince Duplicator that the state $p_1$ does not simulate $q$.

**Simulation game** is defined in a very similar fashion, alternating the following moves:

- Spoiler chooses $a \in A$ and $C \in \Sigma$ such that $\tau_a(x, C) > \tau_a(y, C)$,
- Duplicator answers by choosing $x' \in C$ and $y' \notin C$ and the game continues from $(x', y')$.

Again, a player who cannot make a move at any point loses, and Duplicator wins all infinite plays.

The intuition behind the game is as before, except now Spoiler maintains that his chosen sets $C$ are $\lesssim$-closed, and Duplicator contradicts that by choosing $x' \in C$ and $y' \notin C$ and maintaining that $x' \lesssim y'$. All other considerations remain virtually the same, up to and including:

▶ **Theorem 9.** *$x \lesssim y$ if and only if Duplicator has a winning strategy from $(x, y)$.*

▶ **Example 10.** We illustrate the simulation game on an example (see Fig. 1). In this Markov process there is only one label. From the state $q$, the process loops forever. On the other hand, from the state $p_1$, one can reach the deadlock state $p_4$ through the path to $p_2$ and $p_3$.

We examine the simulation game and how Spoiler can successfully prove to Duplicator that the state $p_1$ does not simulate $q$. We start the simulation game from $(q, p_1)$. A possible first move is $C = \{q, p_2\}$ since $\tau(q, C) = 1 > \tau(p_1, C) = \frac{1}{2}$, but it allows Duplicator to play $(q, p_1)$, back to the original position. A smarter move is $C = \{q, p_1\}$, to which Duplicator has several possible answers, all losing. For instance, if Duplicator plays $(q, p_4)$, Spoiler wins immediately by choosing $C = X$. Duplicator may survive more steps by playing $(q, p_2)$, then $(q, p_3)$, before the fatal $(q, p_4)$.

## 4 Logical characterization of bisimulation, revisited

In this section, we give a short proof for the logical characterization of bisimulation, which relies on two ingredients: the $\pi$-$\lambda$ theorem and the Unique Structure Theorem.

### 4.1 The $\pi$-$\lambda$ Theorem and the Unique Structure Theorem

A *$\pi$-system* is a family of subsets of a set $X$ closed under finite intersections. A *$\lambda$-system* is a family that contains $X$ and is closed under complement and countable disjoint unions. A *$\sigma$-algebra* is a family closed under complement, countable unions and countable intersections. For a family $\mathcal{E}$, let $\sigma(\mathcal{E})$ denote the least $\sigma$-algebra that contains $\mathcal{E}$.

▶ **Theorem 11** (Dynkin's $\pi$-$\lambda$ theorem, [3]). *For any $\pi$-system $\Pi$ and a $\lambda$-system $\Lambda$ on the same set $X$, if $\Pi \subseteq \Lambda$ then $\sigma(\Pi) \subseteq \Lambda$.*

Below, $\equiv_{\mathcal{E}}$ is the relation of equivalence up to $\mathcal{E}$, i.e., $x \equiv_{\mathcal{E}} y$ if and only if, for every $Y \in \mathcal{E}$, $x \in Y$ iff $y \in Y$.

▶ **Theorem 12** (Unique Structure Theorem, [1]). *In any analytic space $(X, \Sigma)$, for every countable family $\mathcal{E} \subseteq \Sigma$ such that $X \in \mathcal{E}$, every measurable, $\equiv_{\mathcal{E}}$-closed subset of $X$ is an element of $\sigma(\mathcal{E})$.*

## 4.2 Logical Characterization

▶ **Theorem 13.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is analytic and $\mathcal{A}$ is countable, $\equiv_\wedge$ is a bisimulation.*

**Proof.** Take some $x, y \in X$ and assume that there exists some $a \in \mathcal{A}$ such that $\tau_a(x, C) \neq \tau_a(y, C)$ for some $\equiv_\wedge$-closed set $C \in \Sigma$. We need to prove that $x \not\equiv_\wedge y$.

Denote $\delta = \tau_a(x, -)$ and $\gamma = \tau_a(y, -)$. If $\delta(X) > \gamma(X)$, then pick a rational number $p$ such that $\delta(X) > p > \gamma(X)$; it is easy so see that $x \models \langle a \rangle_p \top$ and $y \not\models \langle a \rangle_p \top$, therefore $x \not\equiv_\wedge y$. The same formula distinguishes $x$ and $y$ if $\delta(X) < \gamma(X)$.

If $\delta(X) = \gamma(X)$ then pick any $\equiv_\wedge$-closed $C \in \Sigma$ such that $\delta(C) \neq \gamma(C)$. Define

$$\Pi = \{[\![\phi]\!] \mid \phi \in \mathcal{L}_\wedge\} \qquad \text{and} \qquad \Lambda = \{Y \in \Sigma \mid \delta(Y) = \gamma(Y)\}.$$

It is easy to see that $\Pi$ is a $\pi$-system and $\Lambda$ is a $\lambda$-system (in particular, $\Lambda$ is closed under complement since $\delta(X) = \gamma(X)$). Clearly, $\equiv_\Pi = \equiv_\wedge$. Moreover, since there are only countably many formulas, $\Pi$ is countable and, by Theorem 12, $C \in \sigma(\Pi)$. Since by assumption $C \notin \Lambda$, we have $\sigma(\Pi) \not\subseteq \Lambda$, hence (by Theorem 11) $\Pi \not\subseteq \Lambda$. In other words, there exists an $\mathcal{L}_\wedge$ formula $\phi$ such that $\delta([\![\phi]\!]) \neq \gamma([\![\phi]\!])$.

Without loss of generality, assume $\delta([\![\phi]\!]) > \gamma([\![\phi]\!])$ and pick $p \in \mathbb{Q}$ such that $\delta([\![\phi]\!]) > p > \gamma([\![\phi]\!])$. We readily obtain $x \models \langle a \rangle_p \phi$ and $y \not\models \langle a \rangle_p \phi$, hence $x \not\equiv_\wedge y$ as requested. ◀

This easily implies Theorem 4, repeated here:

▶ **Corollary 14.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is analytic and $\mathcal{A}$ is countable, and for any $x, y \in X$, we have that $x \equiv_\wedge y$ if and only if $x \approx y$.*

**Proof.** The right-to-left implication is an easy induction on the structure of formulas. The left-to-right implication is immediate by Theorem 13. ◀

# 5 Logical characterization of simulation, revisited

Our proof of the logical characterization of simulation is completely analogous to the one for bisimulation. It is enough to replace the two main ingredients (Theorems 11 and 12) by new ones.

## 5.1 The Positive Monotone Class Theorem and the Positive Unique Structure Theorem

A *lattice of sets* is a family of subsets of a set $X$ closed under finite unions and intersections.[1] A *monotone class* is a family closed under unions of increasing chains and under intersections of decreasing chains. A *$\sigma$-lattice of sets* is a family of sets closed under countable unions and countable intersections. For a family $\mathcal{E}$, let $L(\mathcal{E})$ denote the least $\sigma$-lattice of sets that contains $\mathcal{E}$.

[1] A lattice of sets is sometimes called *ring of sets*. However, in measure theory ring of sets means something else (a family closed under union and set difference), so we choose a different name.

▶ **Theorem 15** (Positive Monotone Class Theorem). *For any lattice of sets $\mathcal{E}$ and any monotone class $\mathcal{M}$ on the same set $X$, if $\mathcal{E} \subseteq \mathcal{M}$ then $L(\mathcal{E}) \subseteq \mathcal{M}$.*

This result is similar and easier to prove than Theorem 11, and it should be treated as folklore.

Below, $\sqsubseteq_{\mathcal{E}}$ is the preorder determined by $\mathcal{E}$, i.e., $x \sqsubseteq_{\mathcal{E}} y$ if and only if, for every $Y \in \mathcal{E}$, $x \in Y$ implies $y \in Y$.

▶ **Theorem 16** (Positive Unique Structure Theorem). *In any Polish space $(X, \Sigma)$, for every countable family $\mathcal{E} \subseteq \Sigma$, every nonempty, different from $X$, measurable and $\sqsubseteq_{\mathcal{E}}$-closed subset of $X$ is an element of $L(\mathcal{E})$.*

This result strengthens Theorem 12, albeit on the restricted domain of Polish spaces. (Extending it to analytic spaces is future work.) Its proof is also more involved, using ideas similar to the proof of Lusin's Separation Theorem for analytic sets (see [11]). The proof was pointed out to us by Roman Pol.

## 5.2 The logical characterization

▶ **Theorem 17.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is Polish and $\mathcal{A}$ is countable, $\leq_{\wedge}$ is a simulation.*

**Proof.** Take some $x, y \in X$ and assume that there exists some $a \in A$ such that $\tau_a(x, C) > \tau_a(y, C)$ for some $\leq_{\wedge}$-closed set $C \in \Sigma$. We need to prove that $x \not\leq_{\wedge} y$.

Denote $\delta = \tau_a(x, -)$ and $\gamma = \tau_a(y, -)$. Pick any $\leq_{\wedge}$-closed $C \in \Sigma$ such that $\delta(C) > \gamma(C)$. Then $C$ cannot be empty, since $\delta(\emptyset) = \gamma(\emptyset) = 0$. If $C = X$, pick a rational number $p$ such that $\delta(X) > p > \gamma(X)$; it is easy so see that $x \models \langle a \rangle_p \top$ and $y \not\models \langle a \rangle_p \top$, therefore $x \not\leq_{\wedge} y$.

If $C \neq X$, define

$$\mathcal{E} = \{[\![\phi]\!] \mid \phi \in \mathcal{L}_{\wedge}\} \qquad \text{and} \qquad \mathcal{M} = \{Y \in \Sigma \mid \delta(Y) \leq \gamma(Y)\}.$$

It is easy to see that $\mathcal{E}$ is a lattice of sets and (by continuity of measure) $\mathcal{M}$ is a monotone class. Clearly, $\sqsubseteq_{\mathcal{E}} = \leq_{\wedge}$. Moreover, since there are only countably many formulas, $\mathcal{E}$ is countable hence, by Theorem 16, $C \in L(\mathcal{E})$. Since by assumption $C \notin \mathcal{M}$, we have $L(\mathcal{E}) \not\subseteq \mathcal{M}$, hence (by Theorem 15) $\mathcal{E} \not\subseteq \mathcal{M}$. In other words, there exists a formula $\phi$ such that $\delta([\![\phi]\!]) > \gamma([\![\phi]\!])$. Pick $p \in \mathbb{Q}$ such that $\delta([\![\phi]\!]) > p > \gamma([\![\phi]\!])$. We readily obtain $x \models \langle a \rangle_p \phi$ and $y \not\models \langle a \rangle_p \phi$, hence $x \not\leq_{\wedge} y$ as requested. ◀

Compared to Theorem 6, the following easy consequence is restricted to Polish spaces but generalized to countable sets of labels.

▶ **Corollary 18.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is Polish and $\mathcal{A}$ is countable, and for any $x, y \in X$, we have that $x \leq_{\wedge} y$ if and only if $x \lesssim y$.*

**Proof.** The right-to-left implication is an easy induction on the structure of formulas. The left-to-right implication is immediate by Theorem 17. ◀

# 6 The case of uncountably many labels

Our proofs of the logical characterizations for simulation and bisimulation rely on the assumption that the set of formulas (and, equivalently, the set of transition labels) is countable. In this section we investigate the necessity of this assumption. We first observe that indeed if there are uncountably many labels, then the logical characterization fails in general. However, we show that if the transition structure is continuous, then the logical characterization holds again, without any assumption on the set of labels.

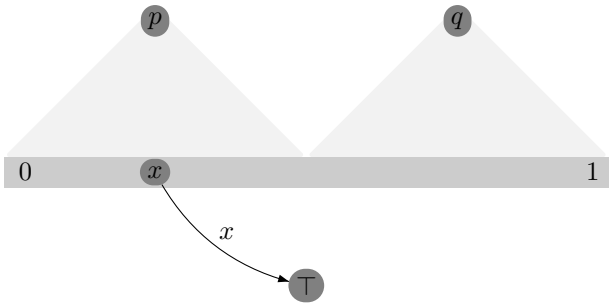


**Figure 2** The two states $p$ and $q$ do not simulate each other, but they satisfy the same formulas of $\mathcal{L}_{\curlywedge}$.

## 6.1 A counterexample

In the classical logical characterization of (bi)similarity for nondeterministic labelled transition systems [10], one can restrict to a logic with finite conjunction and disjunction only if the systems in question satisfy a finite branching property called image finiteness: each state can have only finitely many successors for any given transition label. Since [6, 7] it has been known that this restriction does not apply to probabilistic systems, where a finitary logic is enough to characterize bisimilarity on systems with arbitrary (probabilistic) branching.

On the other hand, in the classical nondeterministic setting, once image finiteness is ensured, the size of the set of transition labels matters very little. Even if infinitely many, or even uncountably many labels are permitted, a finitary logic (with a correspondingly large set of modal operators) is enough to characterize (bi)similarity for nondeterministic transition systems labelled with them.

We now show that this is not the case for labelled Markov processes with continuous state spaces. Specifically, we show an example where the set of labels is uncountable and the logical characterization fails, even though the space of states is a particularly simple, compact Polish space.

Denote $X = \{p, q, \top\} \cup [0, 1]$. We equip $X$ with the smallest $\sigma$-algebra that makes all Borel sets of $[0, 1]$ as well as the singletons $\{p\}, \{q\}$ and $\{\top\}$ measurable. Denote by $\mu$ the Lebesgue[2] probability measure on $[0, 1]$.

Consider a set of actions $\mathcal{A} = [0, 1]$. Define functions $\tau_a : X \times \Sigma \to [0, 1]$ for each $a \in \mathcal{A}$ as follows:

$$\begin{array}{rcl} \tau_a(p, C) & = & \mu(C \cap [0, \frac{1}{2}]) \\ \tau_a(q, C) & = & \mu(C \cap [\frac{1}{2}, 1]) \\ \tau_a(x, \top) & = & \begin{cases} 1 & \text{if } x = a \\ 0 & \text{otherwise} \end{cases} \end{array}$$

The following proposition easily implies that logical characterizations both for bisimulation and for simulation fail for this labelled Markov process.

▶ **Proposition 19.** *Neither $p$ nor $q$ simulates the other, but they satisfy the same formulas of $\mathcal{L}_{\curlywedge}$.*

[2] We mean the usual measure on $[0, 1]$ which assigns to intervals their length. However this is usually extended to the Lebesgue $\sigma$-algebra, i.e., the one obtained by completing the Borel $\sigma$-algebra with respect to this measure. We are just using this measure on the Borel sets.

**Proof.** We prove that neither $p$ nor $q$ simulates the other. First, for any $x, y$ in $[0,1]$, if $x \neq y$ then neither of these simulates the other. Indeed, from $x$, the action $a = x$ leads to $\top$ with probability 1 and leads nowhere from $y$. It follows that every subset of $[0,1]$ is $\lesssim$-closed; in particular this applies to $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. This implies that neither $p$ nor $q$ simulates the other, because $\tau_a(p, [0, \frac{1}{2}]) = 1$ and $\tau_a(q, [0, \frac{1}{2}]) = 0$, and vice-versa $\tau_a(p, [\frac{1}{2}, 1]) = 0$ and $\tau_a(q, [\frac{1}{2}, 1]) = 1$.

To see that $p$ and $q$ satisfy the same formulas, we observe that for every finite subset $\mathcal{B} \subseteq \mathcal{A}$, $p$ and $q$ do simulate each other (indeed, they are even bisimilar) in the system restricted to labels from $\mathcal{B}$. The claim easily follows from this, since every formula of $\mathcal{L}_{\vee\wedge}$ uses finitely many labels.

So for a finite $\mathcal{B} \subseteq \mathcal{A}$, define a relation $R$ on $X$ to be the least equivalence relation such that $pRq$ and $xRy$ for each $x, y \in [0,1] \setminus \mathcal{B}$. We claim that $R$ is a bisimulation on the system restricted to labels with $\mathcal{B}$. The only nontrivial case is the pair $pRq$: every $R$-closed set $C \subseteq [0,1]$ is either finite or co-finite, from which it easily follows that $\tau_a(p, C) = \tau_a(q, C)$. ◀

Intuitively, the core of the problem here is the highly non-continuous nature of transitions from $[0,1]$, allowing one to observe specific states from that uncountable space. Indeed, as we show in the following section, the problem disappears and the logical characterizations hold if we assume that the transition function $\tau_a(\cdot, C)$ is continuous for each $a$ and $C$.

## 6.2 Logical characterizations for continuous transition functions

Given a labelled Markov process $(X, \Sigma, \tau)$ with labels from a set $\mathcal{A}$, we denote by $(X, \Sigma, \tau_{|\mathcal{B}})$ the same system restricted to labels from $\mathcal{B} \subseteq \mathcal{A}$.

▶ **Theorem 20.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is Polish and such that for all $a \in \mathcal{A}, C \in \Sigma$, the function $\tau_a(\cdot, C)$ is continuous, there exists a countable set $\mathcal{B}$ such that the bisimilarity relation $\approx$ on $(X, \Sigma, \tau_{|\mathcal{B}})$ coincides with that on $(X, \Sigma, \tau)$.*

**Proof.** We will use the fact that, under the above assumptions, $X^2$ is also a Polish space for the product topology, hence it satisfies the *hereditary Lindelöf property*: any open cover of a subset of $X^2$ has a countable subcover.

By definition, the bisimilarity relation $\approx$ on $(X, \Sigma, \tau)$ is the largest bisimulation. It is standard to define it as the greatest fixpoint of a certain operator on binary relations on $X$. For us it will be convenient to speak in terms of complements, and we consider the *complement* of $\approx$ as the *least* fixpoint of the operator:

$$\Phi(R) = \left\{ (x,y) \in X^2 \;\middle|\; \exists a \in \mathcal{A}, \exists C \in \Sigma \ (X^2 \setminus R)\text{-closed, s.t. } \tau_a(x, C) \neq \tau_a(y, C) \right\}$$

Thanks to Tarski's fixed point theorem, this is obtained by defining a sequence $(W_\alpha)_\alpha$ of subsets of $X^2$ indexed by ordinals $\alpha$: for $\alpha + 1$ a successor ordinal and $\beta$ a limit ordinal, define:

$$\begin{array}{lcl} W_0 & = & \emptyset \\ W_{\alpha+1} & = & \left\{ (x,y) \in X^2 \;\middle|\; \exists a \in \mathcal{A}, \exists C \in \Sigma \ (X^2 \setminus W_\alpha)\text{-closed, s.t. } \tau_a(x, C) \neq \tau_a(y, C) \right\} \\ W_\beta & = & \bigcup_{\alpha < \beta} W_\alpha. \end{array}$$

The complement of $\approx$ on $(X, \Sigma, \tau)$ is the union of all $W_\alpha$ for all ordinals $\alpha$. More specifically, $(W_\alpha)_\alpha$ form an increasing sequence that reaches a fixpoint at some ordinal $\gamma$ not larger than the cardinality of $\mathcal{P}(X^2)$.

Note that all $W_\alpha$ are open sets in $X^2$. This is proved by ordinal induction: for a successor ordinal, $W_{\alpha+1}$ is a union of sets of the form

$$\{(x,y) \in X^2 \mid \tau_a(x,C) \neq \tau_a(y,C)\}$$

for some $a$ and $C$. Such a set is open, since it is the inverse image of the (open) inequality relation on $[0,1]$ along the continuous function $\tau_a(\cdot, C)$.

For each ordinal $\alpha$ we construct a countable subset $\mathcal{B}_\alpha \subseteq \mathcal{A}$ such that $W_\alpha$ calculated on $(X, \Sigma, \tau_{|\mathcal{B}_\alpha})$ coincides with $W_\alpha$ calculated on $(X, \Sigma, \tau)$.

For sucessor ordinals, rewrite the definition of $W_{\alpha+1}$ as:

$$W_{\alpha+1} = \bigcup_{a \in \mathcal{A}} \{(x,y) \in X^2 \mid \exists C \in \Sigma\ (X^2 \setminus W_\alpha)\text{-closed, s.t. } \tau_a(x,C) \neq \tau_a(y,C)\}.$$

This is a union of open sets. Since $X^2$ is hereditary Lindelöf, one can extract a countable subcover of this union, indexed by some set $\mathcal{B} \subseteq \mathcal{A}$. It is then enough to take $\mathcal{B}_{\alpha+1} = \mathcal{B}_\alpha \cup \mathcal{B}$.

For limit ordinals, extract a countable subcover of the union $W_\beta = \bigcup_{\alpha<\beta} W_\alpha$ and take $\mathcal{B}_\beta$ to be the union of the $\mathcal{B}_\alpha$'s defined for $\alpha$'s from that subcover.

Now the countable set $\mathcal{B}_\gamma$, where $\gamma$ is the ordinal for which $W_\gamma$ reaches the least fixpoint of $\Phi$, satisfies the desired property. ◀

The same result holds for simulation:

▶ **Theorem 21.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is Polish and such that for all $a \in \mathcal{A}, C \in \Sigma$, the function $\tau_a(\cdot, C)$ is continuous, there exists a countable set $\mathcal{B}$ such that the similarity preorder $\lesssim$ on $(X, \Sigma, \tau_{|\mathcal{B}})$ coincides with that on $(X, \Sigma, \tau)$.*

**Proof.** Completely analogous to the proof of Theorem 20, but with the operator

$$\Phi(R) = \{(x,y) \in X^2 \mid \exists a \in \mathcal{A}, \exists C \in \Sigma\ (X^2 \setminus R)\text{-closed, s.t. } \tau_a(x,C) > \tau_a(y,C)\}$$

instead. In particular the fact that each $W_\alpha$ is open, still holds. ◀

The following immediately follows from Theorems 20 and 21 in the light of Corollaries 14 and 18.

▶ **Corollary 22.** *For any labelled Markov process $(X, \Sigma, \tau)$ where $(X, \Sigma)$ is Polish and such that for all $a \in \mathcal{A}, C \in \Sigma$, the function $\tau_a(\cdot, C)$ is continuous, for any $x, y \in X$,*

- *$x \equiv_\wedge y$ if and only if $x \approx y$,*
- *$x \leq_\curlywedge y$ if and only if $x \lesssim y$.*

## 7 Conclusions

The results of this paper suggest that we have arrived at a deeper understanding of the interplay of modal logic and probabilistic transition structure. Variations of the logic can also be used for logical characterization of bisimulation, for example, with the modal construct and just disjunction instead of just conjunction, as studied in [2]. The arguments are minor variations of the proofs given in Section 3sec:bisim. The earlier proof of logical characterization of simulation [8] emerged as a by-product of the theory of approximation; the proof of the present paper is direct. It is particularly pleasing that the two logical characterization proofs have the same general shape and also resemble the overall strategy of the Hennessy-Milner proof.

The game characterization, though elementary, is both pleasing and intriguing. As suggested earlier, there might be interesting links to metrics and the number of moves it takes for Spoiler to win a game. The connection between metrics and bisimulation is well understood but it is possible that via the game one might gain a more quantitative understanding of the numerical significance of the metric.

**Acknowledgments.** We are very much indebted to Roman Pol, who showed us the proof of Theorem 16 which had eluded us for a long time.

We would like to thank the Simons Institute for hosting the program *Logical Structures in Computation* during the Fall of 2016 where we were able to work together in a congenial atmosphere. We are grateful to Martin Otto and Thomas Colcombet for stimulating conversations in Berkeley.

## References

1 W. Arveson. *An Invitation to $C^*$-Algebra*. Springer-Verlag, 1976.
2 M. Bernardo and M. Miculan. Disjunctive probabilistic modal logic is enough for bisimilarity on reactive probabilistic systems. In *ICTCS*, volume 1720, pages 203–220, 2016.
3 P. Billingsley. *Probability and Measure*. Wiley-Interscience, 1995.
4 R. Blute, J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. In *LICS*, 1997.
5 J. Desharnais. *Labelled Markov Processes*. PhD thesis, McGill University, 1999.
6 J. Desharnais, A. Edalat, and P. Panangaden. A logical characterization of bisimulation for labelled Markov processes. In *LICS*, pages 478–489, 1998.
7 J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labeled Markov processes. *Information and Computation*, 179(2):163–193, 2002.
8 J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labeled Markov processes. *Information and Computation*, 184(1):160–200, 2003.
9 J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. A metric for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, June 2004.
10 M. Hennessy and R. Milner. On observing nondeterminism and concurrency. In *ICALP*, volume 85, pages 299–309, 1980.
11 A. S. Kechris. *Classical Descriptive Set Theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, 1995.
12 K. G. Larsen and A. Skou. Bisimulation through probablistic testing. *Information and Computation*, 94:1–28, 1991.
13 P. Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.
14 J. van Benthem. *Modal correspondence theory*. PhD thesis, University of Amsterdam, 1976.