

Bisimilarity of Distributionally Equivalent Markov Transition Systems

Ernst-Erich Doberkat^{1,2}

*Chair for Software Technology
Dortmund University of Technology
Dortmund, Germany*

Abstract

Markov transition systems for interpreting a simple negation free Hennessy-Milner logic are called distributionally equivalent iff for each formula the probability for its extension in one model is matched probabilistically in the other one. This extends in a natural way the notion of logical equivalence which is defined on the states of a transition system to its subprobability distributions. It is known that logical equivalence is equivalent to bisimilarity, i.e., the existence of a span of Borel maps that act as morphisms. We show that distributional equivalence is equivalent to bisimilarity as well, using a characterization of distributional equivalent transition systems through ergodic morphisms. As an aside, we relate bisimilar transition systems to those systems, for which cospans — taken in the category of measurable maps resp. in the Kleisli category associated with the Giry monad — exist.

Keywords: Modal logic, concurrency theory, Markov transition systems, Giry monad, stochastic relations.

1 Introduction and Motivation

Let \mathcal{R} be a non-deterministic or stochastic Kripke model for some modal logic, and assume that we observe \mathcal{R} statistically, yielding $\mu(\llbracket\phi\rrbracket_{\mathcal{R}})$ as the probability that \mathcal{R} is in a state that satisfies formula ϕ , provided the states' distribution is given by probability μ . Here $\llbracket\phi\rrbracket_{\mathcal{R}}$ is as usual the extension of formula ϕ , i.e., the set of worlds in which ϕ holds; we leave issues of measurability aside for the moment. Define for formula ϕ , a probability μ and a rational number q with $0 \leq q \leq 1$ the satisfaction relation \models_q through $\mathcal{R}, \mu \models_q \phi$ iff $\mu(\llbracket\phi\rrbracket_{\mathcal{R}}) \geq q$, thus distribution μ satisfies formula ϕ at least with threshold probability q . We get as a special case for the Dirac measure δ_r on state r that $\mathcal{R}, \delta_r \models_q \phi$ iff $\mathcal{R}, r \models \phi$, whenever $q > 0$. Now suppose that \mathcal{S} is another Kripke model with this property: for each probability μ on \mathcal{R} there

¹ Research funded in part by Deutsche Forschungsgemeinschaft, grant DO 263/8-3, *Algebraische Eigenschaften stochastischer Relationen*

² Email: ernst-erich.doberkat@udo.edu

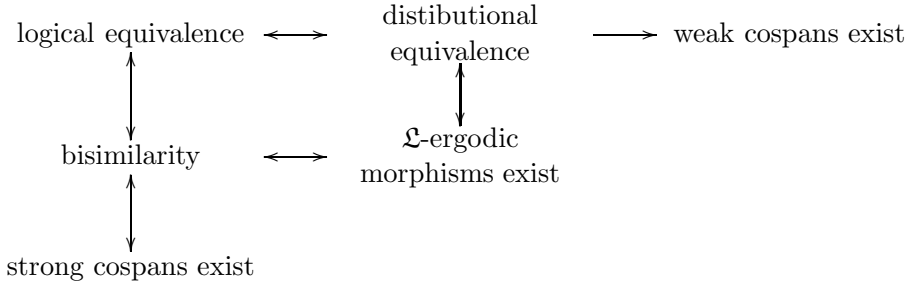


Fig. 1. Relations between Markov transition systems

exists a probability ν on \mathcal{S} such that $\mathcal{R}, \mu \models_q \phi$ iff $\mathcal{S}, \nu \models_q \phi$ holds for all rational q and all ϕ , and vice versa. Then \mathcal{R} and \mathcal{S} are considered to be distributionally equivalent with respect to the distributions on the respective worlds. This appears to be a fairly natural way of comparing the behavior of two non-deterministic or stochastic systems that are so large or so inaccessible that it is difficult to observe states directly, but in which knowledge about the distributions of states can be made available. Embedded systems come to mind as examples for such such systems.

We will investigate this notion of equivalence and put it into perspective with logical equivalence and with bisimilarity for the case that the Kripke models are Markov transition systems, so that they are governed by stochastic laws themselves. For this kind of models logical and behavioral equivalence (i.e., the existence of a cospan) has been investigated quite extensively, and it is well known that logical equivalence, bisimilarity and behavioral equivalence are the same for Markov transition systems. Consequently, e.g., bisimilarity is a suitable tool for the investigation of, say, behaviorally equivalent Kripke models. The question arises whether this is also true in a scenario which is governed by the consideration of distributions on the states rather than the states proper. It turns out, however, that the picture needs a finer brush for depicting the weak case. Among others, we will find some connections to the Kleisli morphisms that come with the Giry monad, which is the monad underlying the measure theoretic part of Probability Theory proposed by M. Giry [11] and investigated in [1,12,3,7], among others.

The problem of relating different ways of describing the behavior of a probabilistic system is interesting for a number of reasons: first, it gives further insight into the inner workings of the Giry monad. Second, it will show that logical equivalence and distributional equivalence are unexpectedly closely related, so are both forms of behavioral equivalence. This permits a better understanding of the underlying logic and shows that the description of the system's behavior through probability distributions is essentially no more powerful than a description that relates states directly to each other. Finally, the results will show that bisimilarity is a concept that, when discussed through distributions, does not yield results that are directly comparable to the results obtained for the familiar strong case. In fact, for reasons of space limitations we will focus on a discussion of bisimilarity and refer the reader for a full discussion to [8].

Organization.

Section 2 introduces the subprobability functor and the Giry monad with its Kleisli construction which is fundamental for the discussion. The logic, its interpretation through Kripke models and some important properties are discussed in Section 3. Bisimulations are defined in Section 4, and various forms of logical equivalence are related to bisimulations. This leads to the main result, see Figure 1. The concluding Section 5 wraps it all up and indicates where further work might be fruitful.

Acknowledgements.

The referees' comments and Christoph Schubert's suggestions have been appreciated.

2 The Giry Monad

We define the subprobability functor and recall some notions from topology, as far as we need them here. Based on this, we define Markov transition systems and their morphisms.

Polish and Analytic Spaces.

Given measurable spaces (X, \mathcal{A}) and (Y, \mathcal{B}) — thus \mathcal{A} and \mathcal{B} are σ -algebras on X resp. Y — a map $f : X \rightarrow Y$ is called *\mathcal{A} - \mathcal{B} -measurable* (or simply *measurable*, when the context is clear) whenever $f^{-1}[\mathcal{B}] \subseteq \mathcal{A}$, thus the inverse image $f^{-1}[B]$ of every member B of \mathcal{B} is a member of \mathcal{A} .

A *Polish space* X is a second countable topological space for which a complete metric exists. The Borel sets $\mathcal{B}(X)$ are the smallest σ -algebra on X which contain the open sets of X . Measurability refers always to the Borel sets, unless otherwise specified.

An *analytic space* is a Hausdorff topological space that is the image of a Polish space under a continuous, or, what amounts to the same, under a Borel map [13]. Hence it makes sense to talk about the Borel sets of an analytic space. The σ -algebra of an analytic space will be the Borel sets, unless otherwise specified, and will usually be omitted from the notation. Analytic spaces are closed under many important operations like countable products, direct and inverse images under Borel maps, factoring under smooth relations (see below), etc. These spaces include discrete spaces as well as the reals \mathbb{R} and in general Polish spaces. They have on one hand pleasant probabilistic properties, on the other hand they are versatile enough to incorporate nearly all interesting applications (in fact, it would be very difficult to find an application that operates on a space that is not analytic). Hence analytic spaces balance generality and applicability just in an adequate way.

The Giry monad.

$\mathbf{S}(X, \mathcal{A})$ denotes for a measurable space (X, \mathcal{A}) the set of all subprobability measures on \mathcal{A} ; this set is endowed with the weak*- σ -algebra \mathcal{A}^\bullet . This is the

smallest σ -algebra that renders the evaluation map $\mu \mapsto \mu(B)$ measurable for each set $B \in \mathcal{A}$. If X is Polish, then $\mathbf{S}(X)$ is a Polish space under the weak topology as well; this is the smallest topology making the maps $\mu \mapsto \int_X f d\mu$ continuous, where $f : X \rightarrow \mathbb{R}$ is bounded and continuous [13]. Then $\mathcal{B}(X)^\bullet = \mathcal{B}(\mathbf{S}(X))$ is folklore [7, Proposition 1.80].

\mathbf{S} is a functor on the category of measurable spaces with measurable maps, where an \mathcal{A} - \mathcal{B} -measurable map $f : X \rightarrow Y$ is assigned the map $\mathbf{S}(f) : \mathbf{S}(X, \mathcal{A}) \rightarrow \mathbf{S}(Y, \mathcal{B})$ through $\mathbf{S}(f)(\mu)(E) := \mu(f^{-1}[E])$. It is not difficult to see that $\mathbf{S}(f)$ is $\mathcal{B}(X)^\bullet$ - $\mathcal{B}(Y)^\bullet$ -measurable. If the Borel map $f : X \rightarrow Y$ is onto, so is $\mathbf{S}(f) : \mathbf{S}(X) \rightarrow \mathbf{S}(Y)$, provided X is Polish and Y is separable metric [7, Proposition 1.30]. If X is an analytic space, so is $\mathbf{S}(X)$ with $\mathcal{B}(\mathbf{S}(X)) = \mathcal{B}(X)^\bullet$. \mathbf{S} is the functorial part of the Girny monad [11]; since unit and multiplication are not of interest here, we do not define them here and refer for a discussion of this monad's properties to [7]. We call $K : (X, \mathcal{A}) \rightsquigarrow (Y, \mathcal{B})$ a *stochastic relation* between the measurable spaces (X, \mathcal{A}) and (Y, \mathcal{B}) iff $K : X \rightarrow \mathbf{S}(Y, \mathcal{B})$ is a measurable map; in Probability Theory, a stochastic relation would be called a sub-Markov kernel or a transition sub-probability. Stochastic relations are just the Kleisli morphisms of the Girny monad [7, Proposition 2.9]:

Proposition 2.1 *Let (X, \mathcal{A}) and (Y, \mathcal{B}) be measurable spaces. Then the following statements are equivalent:*

- i. $K : X \rightarrow \mathbf{S}(Y, \mathcal{B})$ is \mathcal{A} - \mathcal{B}^\bullet -measurable.
- ii. $K : X \rightarrow (\mathcal{B} \rightarrow [0, 1])$ is a map such that
 - a. $K(x)$ is a subprobability on \mathcal{B} for all $x \in X$.
 - b. $x \mapsto K(x)(E)$ is an \mathcal{A} -measurable map for each $E \in \mathcal{B}$.
- iii. K is a morphism in the Kleisli category associated with the Girny monad. \square

This characterization will be used silently throughout. Composition $L * K$ of morphisms $K : (X, \mathcal{A}) \rightsquigarrow (Y, \mathcal{B})$ and $L : (Y, \mathcal{B}) \rightsquigarrow (Z, \mathcal{C})$ in the Kleisli category associated with the Girny monad is what probabilists call the convolution of sub-Markov kernels, viz.,

$$(L * K)(x)(E) = \int_Y L(y)(E) K(x)(dy)$$

We will need the following statement on uniqueness; it hides a proof principle that is applied quite often in this area. The principle is akin to mathematical induction: one shows first that a property holds for a \cap -closed generator \mathcal{A}_0 of a σ -algebra, and in a second step one establishes that the sets for which a property holds is closed under complements and countable disjoint unions. Then one can conclude that all elements of $\sigma(\mathcal{A}_0)$ enjoy this property. Formally it is based on the well-known π - λ -Theorem of Measure Theory, for a discussion see [7, Theorem 1.1]. From the modal logic point of view, this observation makes us prefer conjunction over disjunction. We will return to this point below.

Lemma 2.2 *Let (X, \mathcal{A}) be a measurable space, $\mathcal{A}_0 \subseteq \mathcal{A}$ with $\mathcal{A} = \sigma(\mathcal{A}_0)$ such that \mathcal{A}_0 is closed under intersections. Let $\mathcal{D} \subseteq \mathcal{A}$ with $\mathcal{A}_0 \subseteq \mathcal{D}$ such that \mathcal{D} is closed under complements and countable disjoint unions. Then $\mathcal{A} = \mathcal{D}$. In particular, if $\mu_1, \mu_2 \in \mathbf{S}(X, \mathcal{A})$ are subprobabilities such that $\mu_1(A) = \mu_2(A)$ for all $A \in \mathcal{A}_0$, then $\mu_1 = \mu_2$. \square*

Markov transition systems.

Let G be an analytic space, a *Markov transition system* $\mathcal{G} = (G, (g_a)_{a \in \text{Act}})$ has for each action $a \in \text{Act}$ a stochastic relation $g_a : G \rightsquigarrow G$ modeling the transitions upon action a . So if \mathcal{G} is in state $x \in G$, the probability that the next state is a member of Borel set $A \in \mathcal{B}(G)$ after action $a \in \text{Act}$ is given by $g_a(x)(A)$; there are countable many actions. Note that we do not postulate that $g_a(x)(G) = 1$ for each state x , so that after some action the system may be in no state at all, e.g., if a computation does not terminate.

Congruences.

An equivalence relation ϱ on a measurable space (X, \mathcal{A}) is said to be *smooth* (or countably generated) iff there exists a sequence $(Q_n)_{n \in \mathbb{N}}$ of sets in \mathcal{A} such that $x \varrho x'$ iff $x \in Q_n \Leftrightarrow x' \in Q_n$ holds for all $n \in \mathbb{N}$. The sequence $(Q_n)_{n \in \mathbb{N}}$ is said to *determine* ϱ . We will construct an equivalence relation from the sets on which the formulas of the logic are valid. These sets then form a determining sequence, so that in this case two states are equivalent iff they cannot be separated by the logic.

Let X be analytic, then the factor space X/ϱ is analytic, whenever ϱ is smooth, provided the factor carries the largest σ -algebra that makes the factor map $\eta_\varrho : x \mapsto [x]_\varrho$ $\mathcal{B}(X)$ -measurable [13, Exercise 5.1.14]. Call a set $B \subseteq X$ ϱ -invariant iff $B = \bigcup \{[x]_\varrho \mid x \in B\}$, so that $x \in B$ and $x \varrho x'$ implies $x' \in B$. The ϱ -invariant \mathcal{A} -measurable sets

$$\Sigma(\mathcal{A}, \varrho) := \{C \in \mathcal{A} \mid C \text{ is } \varrho\text{-invariant}\}$$

form a σ -algebra for a measurable space (X, \mathcal{A}) . This σ -algebra has many interesting properties, see [7, Chapter 5.2] for a discussion.

A pair (α, β) of smooth equivalence relations α on X and β on Y (with both X and Y analytic) is called a *congruence* for the stochastic relation $K : X \rightsquigarrow Y$ iff $K(x)(B) = K(x')(B)$ whenever $x \alpha x'$ and $B \subseteq Y$ is a β -invariant Borel set. Thus if α cannot distinguish x from x' , and if β cannot discern the elements of B , then $K(x)$ and $K(x')$ both assign the same probability to B . Given a congruence (α, β) for $K : X \rightsquigarrow Y$, one constructs the factor relation $K_{(\alpha, \beta)} : X/\alpha \rightsquigarrow Y/\beta$ upon setting $K_{(\alpha, \beta)}([x]_\alpha)(G) := K(x)(\eta_\beta^{-1}[G]) = (\mathbf{S}(\eta_\beta) \circ K)(x)(G)$ for $x \in X, G \in \mathcal{B}(Y/\beta)$. As a rule, these factor spaces are usually analytic spaces, even if the base spaces are Polish, see [10, Example 2.7] for a discussion. Note that if the spaces X and Y coincide, and $\alpha = \beta$, we will write down a congruence only as α rather than (α, α) .

3 Hennessy-Milner Logic

Hennessy-Milner logic is introduced here as a simple negation free logic that contains just the bare minimum for a modal logic to be interesting. The equivalence relation that models the separation of states through the logic is introduced and investigated, in particular we have a look at the corresponding invariant sets. We then go on and lift this relation to the set of all subprobabilities on the state space of a transition system, showing that it entails some interesting and useful structural properties of the spaces involved.

The syntax of the Hennessy-Milner logic \mathcal{L} is given by

$$\phi ::= \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle_q \phi$$

Here $a \in \mathbf{Act}$ is an action, and q is a rational number. Hence \mathcal{L} is closed under conjunction and under the modal operator $\langle a \rangle_q$ for an action a , the index q indicating the degree of satisfaction after action a , as we will see. Conjunction serves not only as a representative of the Boolean operations that are usually included in a logic (here disjunction would do as well), it is also — with a view towards Lemma 2.2 — decisive for the development from a technical vantage point, and here disjunction would **not** do; this was first observed in [3]. Disjunction may be added, however, through a natural transformation; this technique is discussed at length in [9,10].

The set \mathbf{Act} of labels is assumed to be at most countably infinite. The semantics of $\langle a \rangle_q \phi$ says intuitively that we can make an a -move in a state x to a state that satisfies ϕ with probability at least q . \mathfrak{F} denotes the countable set of all formulas for this logic.

Let $\mathcal{G} = (G, (g_a)_{a \in \mathbf{Act}})$ be a Markov transition system. The logic \mathcal{L} will be interpreted in \mathcal{G} . Satisfaction of a state x for a formula ϕ is defined inductively, $\llbracket \phi \rrbracket_{\mathcal{G}}$ is defined as usual as the set of all states x that satisfy formula ϕ , i.e. for which $\mathcal{G}, x \models \phi$ holds. This is straightforward for \top and for formulas of the form $\phi_1 \wedge \phi_2$ upon defining $\llbracket \top \rrbracket_{\mathcal{G}} := G$ resp. $\llbracket \phi_1 \wedge \phi_2 \rrbracket_{\mathcal{G}} := \llbracket \phi_1 \rrbracket_{\mathcal{G}} \cap \llbracket \phi_2 \rrbracket_{\mathcal{G}}$. The more interesting case is making an a -move: $\mathcal{G}, x \models \langle a \rangle_q \phi$ holds iff $g_a(x)(\llbracket \phi \rrbracket_{\mathcal{G}}) \geq q$.

Note that the set $\{\llbracket \phi \rrbracket_{\mathcal{G}} \mid \phi \in \mathfrak{F}\}$ is closed under finite intersections. An easy induction on the structure of ϕ shows that the sets $\llbracket \phi \rrbracket_{\mathcal{G}}$ are well behaved:

Lemma 3.1 $\llbracket \phi \rrbracket_{\mathcal{G}}$ is a Borel subset of G for each formula ϕ . \square

Define the states x_1, x_2 to be equivalent iff they cannot be separated through a formula, thus

$$x_1 r_{\mathcal{G}} x_2 \text{ iff } \forall \phi \in \mathfrak{F} : [\mathcal{G}, x_1 \models \phi \Leftrightarrow \mathcal{G}, x_2 \models \phi]$$

The theory $\text{Th}_{\mathcal{G}}(x)$ of a state x is the set of all formulas that are valid in this state, thus two states cannot be separated by the logic iff their theories coincide, hence iff they are related through relation $r_{\mathcal{G}}$.

Because \mathfrak{F} is countable, the equivalence relation $r_{\mathcal{G}}$ is smooth, and the factor space $G/r_{\mathcal{G}}$ is analytic again. The relation $r_{\mathcal{G}}$ is a congruence for each $g_a : G \rightsquigarrow G$,

accordingly, the factor model $\mathcal{G}/\mathfrak{L} := (G/r_{\mathcal{G}}, (g_a, \mathfrak{L})_{a \in \text{Act}})$ is defined through

$$g_{a, \mathfrak{L}}([x]_{r_{\mathcal{G}}}) := (\mathbf{S}(\eta_{r_{\mathcal{G}}}) \circ g_a)(x).$$

Consequently, the probability $g_{a, \mathfrak{L}}([x]_{r_{\mathcal{G}}})(A)$ of making an a -move from class $[x]_{r_{\mathcal{G}}}$ to a Borel set $A \in \mathcal{B}(G/r_{\mathcal{G}})$ in the factor system \mathcal{G}/\mathfrak{L} equals the probability $g_a(x)(\eta_{r_{\mathcal{G}}}^{-1}[A])$ of making an a -move in the original system from state x into the invariant Borel set $\eta_{r_{\mathcal{G}}}^{-1}[A]$.

The transition law g_a is a stochastic relation when restricted to the $r_{\mathcal{G}}$ -invariant sets. Thus the inverse image of $x \mapsto g_a(x)(A)$ is a member of $\Sigma(G, r_{\mathcal{G}})$, whenever $A \subseteq G$ is a $r_{\mathcal{G}}$ -invariant Borel set. This holds in particular whenever $A = \llbracket \phi \rrbracket_{\mathcal{G}}$ for some formula $\phi \in \mathfrak{F}$. Denote this restriction by $|g_a|$.

Lemma 3.2 $|g_a| : (G, \Sigma(G, r_{\mathcal{G}})) \rightsquigarrow (G, \Sigma(G, r_{\mathcal{G}}))$ is a stochastic relation. \square

The equivalence relation

$$\mu_1 \overline{r_{\mathcal{G}}} \mu_2 \text{ iff } \forall A \in \Sigma(G, r_{\mathcal{G}}) : \mu_1(A) = \mu_2(A)$$

is defined on $\mathbf{S}(G)$. It is the *lifting* of relation $r_{\mathcal{G}}$ to $\mathbf{S}(G)$ [6, Section 3]. Given states $x_1, x_2 \in G$, the relation $x_1 r_{\mathcal{G}} x_2$ is equivalent to $\delta_{x_1} \overline{r_{\mathcal{G}}} \delta_{x_2}$ for the corresponding Dirac measures.

Define for rational q , a formula $\phi \in \mathfrak{F}$ and a subprobability $\mu \in \mathbf{S}(G)$ the threshold satisfiability relation \models_q through $\mathcal{G}, \mu \models_q \phi$ iff $\mu(\llbracket \phi \rrbracket_{\mathcal{G}}) \geq q$, so that ϕ is satisfied with at least probability q . Note that $\mathcal{G}, g_a(x) \models_q \phi$ iff $\mathcal{G}, x \models \langle a \rangle_q \phi$, hence \models_q may be used for modeling moves in \mathcal{G} . In fact, there are deeper connections which are exhibited through the Choquet representation for an integral, see [8].

We will introduce different kinds of morphisms now. Recall that $K * L$ denotes the Kleisli composition for stochastic relations K and L .

Definition 3.3 Let $\mathcal{G} = (G, (g_a)_{a \in \text{Act}})$ and $\mathcal{H} = (H, (h_a)_{a \in \text{Act}})$ be Markov transition systems.

- (i) A surjective map $s : \mathcal{G} \rightarrow \mathcal{H}$ is called a *strong morphism* (denoted through $s : G \rightarrow H$) iff $\mathbf{S}(s) \circ g_a = h_a \circ s$ holds for each $a \in \text{Act}$.
- (ii) A stochastic relation $\Phi : G \rightsquigarrow H$ is called a *weak morphism* ($\mathcal{G} \curvearrowright \mathcal{H}$) iff $\Phi * g_a = h_a * \Phi$ for each $a \in \text{Act}$ holds.

For a strong morphism $s : \mathcal{G} \rightarrow \mathcal{H}$ the probability $h_a(s(x))(G)$ for \mathcal{H} to hit a state in Borel set B from state $s(x)$ equals the probability $(\mathbf{S}(s) \circ g_a(x))(G) = g_a(x)(s^{-1}[G])$ for G to hit a state in its inverse image $s^{-1}[G]$ under s from state x . In contrast, a weak morphism $\Phi : \mathcal{G} \curvearrowright \mathcal{H}$ has the property that

$$(\Phi * g_a)(x)(B) = \int_G \Phi(t)(B) g_a(x)(dt) = \int_H h_a(y)(B) \Phi(x)(dy) = (h_a * \Phi)(x)(B)$$

holds for each action a , each state $x \in G$, and each Borel set $B \subseteq H$. A strong morphism $\gamma : \mathcal{G} \rightarrow \mathcal{H}$ hatches the weak morphism $\delta_{\gamma(\cdot)} : \mathcal{G} \curvearrowright \mathcal{H}$ [6, Lemma 4.4].

Weak morphisms with a particular affinity to the logic will be called \mathfrak{L} -ergodic (Arveson [2, p. 78] proposes the name *ergodic* for measures that take the values 0 or 1 on the invariant Borel sets for some equivalence relation; we extend this a bit to stochastic relations).

Definition 3.4 A stochastic relation $\Psi : (G, \Sigma(G, r_G)) \rightsquigarrow (H, \Sigma(H, r_H))$ is called \mathfrak{L} -ergodic ($\mathcal{G} \curvearrowright_{\mathfrak{L}} \mathcal{H}$) iff

- (i) $\Psi * |g_a| = |h_a| * \Psi$ for each $a \in \text{Act}$
- (ii) for all $\phi \in \mathfrak{F}$ the map $x \mapsto \Psi(x)(\llbracket \phi \rrbracket_{\mathcal{H}})$ takes only the values 0 and 1 with $\Psi(x)(\llbracket \phi \rrbracket_{\mathcal{H}}) = 1$ iff $\mathcal{G}, x \models \phi$.

If $\Psi : \mathcal{G} \curvearrowright_{\mathfrak{L}} \mathcal{H}$ is \mathfrak{L} -ergodic, we know that $(\Psi * |g_a|)(x)(B) = (|h_a| * \Psi)(x)(B)$ holds for each action a , each state $x \in G$, and each r_H -invariant Borel set B . Thus the compositional equation holds only on those Borel sets that are related to the logic. Hence \mathfrak{L} -ergodic morphisms are versions of weak morphisms that are in a specific way adapted to the logic, see [6, Section 5] for a discussion. In particular, we know for an \mathfrak{L} -ergodic morphism Φ that $\Phi(x)(\llbracket \phi \rrbracket_{\mathcal{H}}) = \Phi(x')(\llbracket \phi \rrbracket_{\mathcal{H}})$ for each formula $\phi \in \mathfrak{F}$, whenever $x r_G x'$. Consequently, states that cannot be separated through the logic will end up in states that satisfy the same formulas with identical probabilities. An ergodic \mathfrak{L} -morphism is further characterized through $\Phi(x)(\llbracket \phi \rrbracket_{\mathcal{H}}) = 1_{\llbracket \phi \rrbracket_{\mathcal{G}}}(x)$, 1_A denoting the indicator function of the set A . Putting it less formally, we know that, whenever $\mathcal{G}, x \models \phi$, the state's counterpart y in model \mathcal{H} will satisfy $\mathcal{H}, y \models \phi$ almost surely. Hence validity in model \mathcal{G} is very closely tied to validity in model \mathcal{H} , albeit on the level of distributions. Note that we do not need to impose any assumptions on the value of $\Phi(x)(B)$ whenever B is a r_H -invariant Borel subset of H other than the extension for some formula. This is so since standard arguments including the $\pi - \lambda$ -Theorem Lemma 2.2 show that $\Phi(x)(B) \in \{0, 1\}$ whenever $B \in \Sigma(H, r_H)$. Because (r_G, r_H) is a congruence for an ergodic relation $\Psi : \mathcal{G} \curvearrowright_{\mathfrak{L}} \mathcal{H}$, we may define the factor relation $\Psi_{\mathfrak{L}} : G/\mathfrak{L} \rightsquigarrow H/\mathfrak{L}$ through $\Psi_{\mathfrak{L}}([x]_{r_G}) := (\mathbf{S}(\eta_{r_H}) \circ \Psi)(x)$. This constitutes a weak morphism [6, Proposition 5.9]. Turning to logical equivalence, we define

Definition 3.5 Let \mathcal{G} and \mathcal{H} be Markov transition systems.

- (i) \mathcal{G} and \mathcal{H} are *logical equivalent* iff $\{\text{Th}_{\mathcal{G}}(x) \mid x \in G\} = \{\text{Th}_{\mathcal{H}}(y) \mid y \in H\}$.
- (ii) \mathcal{G} and \mathcal{H} are *distributionally equivalent* iff for each $\mu \in \mathbf{S}(G)$ there exists $\nu \in \mathbf{S}(H)$ such that $\mathcal{G}, \mu \models_q \phi \Leftrightarrow \mathcal{H}, \nu \models_q \phi$ for all q and all formulas $\phi \in \mathfrak{F}$, and vice versa.

Thus \mathcal{G} and \mathcal{H} are distributionally equivalent iff each distribution on one model is matched by a distribution on the other model:

Lemma 3.6 *The Markov transition systems \mathcal{G} and \mathcal{H} are distributionally equivalent iff these conditions hold*

- i. for all $\mu \in \mathbf{S}(G)$ there exists $\nu \in \mathbf{S}(H)$ with $\forall \phi \in \mathfrak{F} : \mu(\llbracket \phi \rrbracket_{\mathcal{G}}) = \nu(\llbracket \phi \rrbracket_{\mathcal{H}})$,
- ii. for all $\nu \in \mathbf{S}(H)$ there exists $\mu \in \mathbf{S}(G)$ with $\forall \phi \in \mathfrak{F} : \nu(\llbracket \phi \rrbracket_{\mathcal{H}}) = \mu(\llbracket \phi \rrbracket_{\mathcal{G}})$. \square

The definition for logical equivalence reads that given a state in one model there exists a state in the other model such that the respective theories coincide, so that given $x \in G$ there exists $y \in H$ with $\mathcal{G}, x \models \phi \Leftrightarrow \mathcal{H}, y \models \phi$ for each formula ϕ , and vice versa. Since $\mathcal{G}, x \models \phi$ is equivalent to $x \in \llbracket \phi \rrbracket_{\mathcal{G}}$ which in turn is equivalent to $\delta_x(\llbracket \phi \rrbracket_{\mathcal{G}}) = 1$ with δ_x as the Dirac measure on x , we can state that \mathcal{G} and \mathcal{H} are logical equivalent iff for each $x \in G$ there exists $y \in H$ such that $\forall \phi \in \mathfrak{F} : \delta_x(\llbracket \phi \rrbracket_{\mathcal{G}}) = \delta_y(\llbracket \phi \rrbracket_{\mathcal{H}})$ and vice versa. Thus distributional equivalence generalizes logical equivalence, and it can be shown that logically equivalent models are distributionally equivalent [8]. This is a characterization in terms of morphisms.

Proposition 3.7 *Let \mathcal{G} and \mathcal{H} be Markov transition systems. Then the following statements are equivalent*

- a. \mathcal{G} and \mathcal{H} are distributionally equivalent.
- b. There exist \mathfrak{L} -ergodic morphisms $\Lambda : \mathcal{G} \curvearrowright_{\mathfrak{L}} \mathcal{H}$ and $\Xi : \mathcal{H} \curvearrowright_{\mathfrak{L}} \mathcal{G}$. \square

4 Bisimulations

Fix the Markov transition systems $\mathcal{G} = (G, (g_a)_{a \in \text{Act}})$ and $\mathcal{H} = (H, (h_a)_{a \in \text{Act}})$. \mathcal{G} and \mathcal{H} are called *strongly bisimilar* whenever we can find a Markov transition system \mathcal{M} as well as strong morphisms $s : \mathcal{M} \rightarrow \mathcal{G}$ and $t : \mathcal{M} \rightarrow \mathcal{H}$. This means that the following diagram commutes for each action $a \in \text{Act}$.

$$\begin{array}{ccccc}
 G & \xleftarrow{s} & M & \xrightarrow{t} & H \\
 g_a \downarrow & & \downarrow m_a & & \downarrow h_a \\
 \mathbf{S}(G) & \xleftarrow{\mathbf{S}(s)} & \mathbf{S}(M) & \xrightarrow{\mathbf{S}(t)} & \mathbf{S}(H)
 \end{array}$$

\mathcal{M} is sometimes thought of as mediating between \mathcal{G} and \mathcal{H} . From now on we will omit the attribute *strong* from morphisms.

Each ergodic morphism induces a measurable map between the factor spaces.

Lemma 4.1 *Let $\Psi : \mathcal{G} \curvearrowright_{\mathfrak{L}} \mathcal{H}$ be \mathfrak{L} -ergodic. Then there exists a measurable map $\psi : G/r_{\mathcal{G}} \rightarrow H/r_{\mathcal{H}}$ such that $\Psi_{\mathfrak{L}}(t) = \delta_{\psi(t)}$ for all $t \in G/r_{\mathcal{G}}$.*

Proof. 0. If a subprobability $\mu \in \mathbf{S}(A)$ for an analytic space A takes only the values 0 and 1, and if $\mu(A) = 1$, then there exists a unique point $a \in A$ such that $\mu = \delta_a$. This is so because an analytic space is countably generated, so that [2, Lemma after Theorem 3.4.3] can be applied.

1. For each $x \in G$ the measure $\Psi([x]_{r_{\mathcal{G}}}) \in \mathbf{S}(H/r_{\mathcal{H}})$ is binary valued. Thus there exists by part 0. for each $x \in G$ some $y \in H$ such that $\Psi_{\mathfrak{L}}([x]_{r_{\mathcal{G}}}) = \delta_{[y]_{r_{\mathcal{H}}}}$. Since $[y]_{r_{\mathcal{H}}}$ is uniquely determined, the map $\psi([x]_{r_{\mathcal{G}}}) := [y]_{r_{\mathcal{H}}}$ is well defined.

2. The map ψ is Borel measurable. One observes first that for each $\phi \in \mathfrak{F}$ the

identity $\llbracket \phi \rrbracket_G = \eta_{r_G}^{-1} [\psi^{-1} [\eta_{r_H} [\llbracket \phi \rrbracket_H]]]$. In fact,

$$\begin{aligned} x \in \llbracket \phi \rrbracket_G &\Leftrightarrow \Psi(x)(\llbracket \phi \rrbracket_H) = 1 \\ &\stackrel{(\natural)}{\Leftrightarrow} \Psi_{\mathfrak{L}}([x]_{r_G})(\eta_{r_H} [\llbracket \phi \rrbracket_H]) = 1 \\ &\Leftrightarrow \psi([x]_{r_G}) \in \eta_{r_H} [\llbracket \phi \rrbracket_H] \\ &\Leftrightarrow x \in \eta_{r_G}^{-1} [\psi^{-1} [\eta_{r_H} [\llbracket \phi \rrbracket_H]]]. \end{aligned}$$

The equivalence (\natural) uses that by definition $\Psi(x)(D)$ equals $\Psi_{\mathfrak{L}}([x]_{r_G})(\eta_{r_H} [D])$, whenever $D \in \Sigma(H, r_H)$, and that for r_H -invariant D the equality $D = \eta_{r_H}^{-1} [\eta_{r_H} [D]]$ holds. Consequently, $\mathcal{T} := \{T \in \mathcal{B}(H/r_H) \mid \psi^{-1}[T] \in \mathcal{B}(G/r_G)\}$ includes $\{\eta_{r_H} [\llbracket \phi \rrbracket_H] \mid \phi \in \mathfrak{F}\}$. The latter set is closed under intersections due to the logic being closed under conjunctions, and the former is closed under complements and disjoint countable unions. Thus it follows from the π - λ -Theorem (Lemma 2.2) that $\psi^{-1}[T] \in \mathcal{B}(G/r_G)$ for all $T \in \mathcal{B}(H/r_H)$, so that $\mathcal{T} = \mathcal{B}(H/r_H)$. \square

The map ψ constructed in this way is nearly a morphism:

Corollary 4.2 *This diagram commutes, ψ being constructed according to Lemma 4.1:*

$$\begin{array}{ccc} G/r_G & \xrightarrow{\psi} & H/r_H \\ g_{a,\mathfrak{L}} \downarrow & & \downarrow h_{a,\mathfrak{L}} \\ \mathbf{S}(G/r_G) & \xrightarrow{\mathbf{S}(\psi)} & \mathbf{S}(H/r_H) \end{array}$$

Consequently, ψ preserves and reflects validity: the equivalence

$$\mathcal{G}/\mathfrak{L}, t \models \varphi \Leftrightarrow \mathcal{H}/\mathfrak{L}, \psi(t) \models \varphi$$

is true for every $t \in G/r_G$ and every formula $\varphi \in \mathfrak{F}$.

Proof. Because $\Psi_{\mathfrak{L}} : \mathcal{G}/\mathfrak{L} \curvearrowright \mathcal{H}/\mathfrak{L}$ is a weak morphism, we have for all $x \in G, D \in \mathcal{B}(H/r_H)$

$$\begin{aligned} (h_{a,\mathfrak{L}} \circ \psi)([x]_{r_G})(D) &= (h_{a,\mathfrak{L}} * \Psi_{\mathfrak{L}})([x]_{r_G})(D) \\ &= (\Psi_{\mathfrak{L}} * g_{a,\mathfrak{L}})([x]_{r_G})(D) \\ &= \int_{G/r_G} \delta_{\psi(z)}(D) g_{a,\mathfrak{L}}([x]_{r_G})(dz) \\ &= g_{a,\mathfrak{L}}([x]_{r_G})(\psi^{-1}[D]) \\ &= (\mathbf{S}(\psi) \circ g_{a,\mathfrak{L}})([x]_{r_G})(D). \square \end{aligned}$$

\square

If we would be able to convince ourselves that ψ is onto, then we would have constructed a morphism. Hence it remains to search for criteria under which ψ is surjective. The observation that the equivalence classes $[x]_{r_G}$ are exactly the atoms of the σ -algebra $\Sigma(G, r_G)$ is helpful for the proof.

Lemma 4.3 Define for $y \in H, \phi \in \mathfrak{F}$ the set

$$I(\phi, y) := \begin{cases} \llbracket \phi \rrbracket_{\mathcal{H}}, & \mathcal{H}, y \models \phi; \\ H \setminus \llbracket \phi \rrbracket_{\mathcal{H}}, & \text{otherwise,} \end{cases}$$

and put $J(\phi, y) := \{x \in G \mid \Psi(x)(I(\phi, y)) = 1\}$. Then $[y]_{r_G} = \bigcap_{\phi \in \mathfrak{F}} I(\phi, y)$, and if $\bigcap_{\phi \in \mathfrak{F}} J(\phi, y) \neq \emptyset$ for all $y \in H$, then ψ is onto. \square

This is now applied to distributionally equivalent transition systems.

Proposition 4.4 Assume that \mathcal{G} and \mathcal{H} are distributionally equivalent, then

- there exist strong morphisms $\mathcal{G}/\mathfrak{L} \rightarrow \mathcal{H}/\mathfrak{L}$ and $\mathcal{H}/\mathfrak{L} \rightarrow \mathcal{G}/\mathfrak{L}$,
- $g_a(x)(\llbracket \phi \rrbracket_{\mathcal{G}}) = h_a(y)(\llbracket \phi \rrbracket_{\mathcal{H}})$, holds for all $\phi \in \mathfrak{F}$, provided $x \in G$ and $y \in H$ satisfy exactly the same formulas.

Proof. 1. Construct the ergodic \mathfrak{L} -morphisms $\Psi : \mathcal{G} \curvearrowright_{\mathfrak{L}} \mathcal{H}$ and $\Xi : \mathcal{H} \curvearrowright_{\mathfrak{L}} \mathcal{G}$ according to Proposition 3.7, and let $\psi : G/r_G \rightarrow H/r_H$ as well as $\xi : H/r_H \rightarrow G/r_G$ be the corresponding point maps. Both ψ and ξ preserve validity by Corollary 4.2, so we infer from Lemma 4.3 that $J(\phi, y) = \eta_{r_G}^{-1} [\xi^{-1} [\eta_{r_H} [I(\phi, y)]]$. Since $I(\phi, y)$ is r_G -invariant, and because η_{r_H} interchanges with countable intersections on invariant sets, we conclude that $\bigcap_{\phi \in \mathfrak{F}} J(\phi, y) \neq \emptyset$, provided $\bigcap_{\phi \in \mathfrak{F}} I(\phi, y)$ is not empty. Thus ψ is onto, interchanging the rôles of \mathcal{G} and \mathcal{H} , the argument shows that ξ is surjective as well. Thus ψ and ξ are strong morphisms, since the crucial diagram commutes by Corollary 4.2.

2. Let $\phi \in \mathfrak{F}$ be an arbitrary formula, and assume that $x \in G, y \in \psi([x]_{r_G})$. We show that $g_a(x)(\llbracket \phi \rrbracket_{\mathcal{G}}) = h_a(y)(\llbracket \phi \rrbracket_{\mathcal{H}})$ holds:

$$\begin{aligned} g_a(x)(\llbracket \phi \rrbracket_{\mathcal{G}}) &= \int_G \Psi(x')(\llbracket \phi \rrbracket_{\mathcal{H}}) g_a(x)(dx') \\ &= (\Psi * g_a)(x)(\llbracket \phi \rrbracket_{\mathcal{H}}) \\ &= (h_a * \Psi)(x)(\llbracket \phi \rrbracket_{\mathcal{H}}) \\ &= h_{a, \mathfrak{L}}(\psi([x]_{r_G}))(\eta_{r_H} [\llbracket \phi \rrbracket_{\mathcal{H}}]) \\ &= h_a(y)(\llbracket \phi \rrbracket_{\mathcal{H}}), \end{aligned}$$

whenever $[y]_{r_H} = \psi([x]_{r_G})$, i.e., whenever $y \in \psi([x]_{r_G})$ according to the definition of the factor relation. This proves part b. \square

Proposition 4.5 Assume that \mathcal{G} and \mathcal{H} are distributionally equivalent, then they are bisimilar.

Proof. Denote by $\mathcal{T}_{\mathcal{G}} := \{\llbracket \phi \rrbracket_{\mathcal{G}} \mid \phi \in \mathfrak{F}\}$ the countable generator for the σ -algebra $\Sigma(G, r_G)$, similarly for $\mathcal{T}_{\mathcal{H}}$. Both generators are closed under finite intersection, since the logic is closed under finite conjunctions. Through Proposition 4.4, part b. and the representations

$$\llbracket \phi \rrbracket_{\mathcal{H}} = \bigcup \{\psi([x]_{r_G}) \mid x \in \llbracket \phi \rrbracket_{\mathcal{G}}\}, \llbracket \phi \rrbracket_{\mathcal{G}} = \bigcup \{\xi([y]_{r_H}) \mid y \in \llbracket \phi \rrbracket_{\mathcal{H}}\},$$

we infer bisimilarity from the general criterion [7, Proposition 5.39]. \square

Summing up, this has as a consequence the following general result.

Theorem 4.6 *These statements are equivalent for the Markov transition systems \mathcal{G} and \mathcal{H} .*

- i. \mathcal{G} and \mathcal{H} are logically equivalent.
- ii. \mathcal{G} and \mathcal{H} are distributionally equivalent.
- iii. \mathcal{G} and \mathcal{H} are bisimilar.
- iv. There exist \mathcal{L} -ergodic morphisms $\Psi : \mathcal{G} \curvearrowright_{\mathcal{L}} \mathcal{H}$ and $\Xi : \mathcal{H} \curvearrowright_{\mathcal{L}} \mathcal{G}$.

Proof. $\text{i} \Leftrightarrow \text{iii}$ is established in [5, Theorem 6.6], [8] establishes the implication $\text{i} \Leftrightarrow \text{ii}$, Proposition 4.5 shows $\text{ii} \Leftrightarrow \text{iii}$, and the equivalence $\text{ii} \Leftrightarrow \text{iv}$ is proven in Proposition 3.7. \square

It is known that logical equivalence and the existence of a cospan of strong morphisms is equivalent for Markov transition systems [7, Theorem 6.19], and it can be shown that distributional equivalence implies the existence of a cospan of weak morphisms [8]. Figure 1 summarizes the overall situation.

5 Conclusion and Further Work

This paper defines distributional equivalence with respect to a simple modal logic for Markov transition systems. It argues that this is a weak form of logical equivalence, and it relates these characterizations for the behavior of transition systems to bisimilarity. The main tool is a general criterion on bisimilarity for stochastic relations. This is made applicable through an investigation of ergodic morphisms between the transition systems. One of the referees suggested to use the final coalgebra and lift the morphism that maps each coalgebra to the final one; this might give the result. Alas — this is a difficult thing to do because in the category under consideration a coalgebra exists only for the probability functor [4]. Thus a direct argumentation has been used.

The discussion may be extended into two directions: First, an interplay of distributional equivalence and Kleisli morphisms for the Girmonad could be observed. The somewhat natural hypothesis that weak bisimilarity, distributional equivalence and the existence of a weak cospan are equivalent — carried over from the case of strong morphisms — does not seem to hold in this generality, as Figure 1 suggests. Hence it invites a careful investigation of bisimilarity in the weak case; a look at the results in [6] shows that entering the discussion of Kleisli morphisms adds considerable color. Second, the treatment of coalgebraic stochastic logic [9,10] should be extended to Kleisli morphisms for the Girmonad. Here some non-trivial technical obstacles need to be overcome.

References

- [1] S. Abramsky, R. Blute, and P. Panangaden. Nuclear and trace ideal in tensored $*$ -categories. *J. Pure Appl. Alg.*, 143(1 – 3):3 – 47, 1999.
- [2] W. Arveson. *An Invitation to C^* -Algebra*. Graduate Texts in Mathematics. Springer-Verlag, New York, 1976.
- [3] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation of labelled Markov-processes. *Information and Computation*, 179(2):163 – 193, 2002.
- [4] E.-E. Doberkat. Look: simple stochastic relations are just, well, simple. In J. Fiadeiro and J. Rutten, editors, *Proc. CALCO*, volume 3629 of *Lect. Notes Comp. Sci.*, pages 128 – 142, 2005.
- [5] E.-E. Doberkat. Stochastic relations: congruences, bisimulations and the Hennessy-Milner theorem. *SIAM J. Computing*, 35(3):590 – 626, 2006.
- [6] E.-E. Doberkat. Kleisli morphisms and randomized congruences for the Giry monad. *J. Pure Appl. Alg.*, 211:638–664, 2007.
- [7] E.-E. Doberkat. *Stochastic Relations. Foundations for Markov Transition Systems*. Chapman & Hall/CRC Press, Boca Raton, New York, 2007.
- [8] E.-E. Doberkat. Weak logical equivalence for Markov transition systems. Technical Report 167, Chair for Software-Technology, University of Dortmund, December 2007.
- [9] E.-E. Doberkat. Stochastic coalgebraic logic: Bisimilarity and behavioral equivalence. *Ann. Pure Appl. Logic*, in print, 2008.
- [10] E.-E. Doberkat and C. Schubert. Coalgebraic logic for stochastic right coalgebras. Technical report, Chair for Software-Technology, University of Dortmund, September 2007.
- [11] M. Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, number 915 in *J. Pure Appl. Alg.*, pages 68 – 85, Berlin, 1981. Springer-Verlag.
- [12] P. Panangaden. The category of Markov kernels. *ENTCS*, 22, 1999.
- [13] S. M. Srivastava. *A Course on Borel Sets*. Graduate Texts in Mathematics. Springer-Verlag, Berlin, 1998.