

Belief bisimulation for hidden Markov models

logical characterisation and decision algorithm

David N. Jansen¹, Flemming Nielson² and Lijun Zhang²

¹ Radboud Universiteit Nijmegen, Netherlands

² Technical University of Denmark

Abstract. This paper establishes connections between logical equivalences and bisimulation relations for hidden Markov models (HMM). Both standard and belief state bisimulations are considered. We also present decision algorithms for the bisimilarities. For standard bisimilarity, an extension of the usual partition refinement algorithm is enough. Belief bisimilarity, being a relation on the continuous space of belief states, cannot be described directly. Instead, we show how to generate a linear equation system in time cubic in the number of states.

1 Introduction

Probabilistic models like Markov chains allow to describe processes whose behaviour is governed by probabilistic distributions. Together with extensions with nondeterministic choices, reward structures and continuous time, they are widely used in networked and distributed systems. During the last twenty years, efficient model-checking algorithms of Markov chains and their extensions have been extensively studied, allowing for performance evaluation and formal reasoning.

Markov chains are fully observable, in the sense that at any time, an observer can determine the exact state and infer the probability to be in a specific state at later times. This may be too restrictive in many applications: Intuitively, the underlying state space of a Markov chain may contain fine-grained information, which is not always visible from the outside. For instance, a meteorologist might use a Markov chain with states for several kinds of snow [18] to model the weather behaviour. Non-expert observers only see whether it is snowing or not, implying that the states of the Markov chain are not fully visible to them.

Hidden Markov models (HMM) [14] enhance Markov chains with *observations*. These reveal partial information about the state, while the actual state remains unknown. Given the sequence of produced observations, we may infer a probability distribution over the states, a so-called *belief state*.

HMMs have received much attention in the area of speech recognition [9], communication channel modelling [15], and biological systems [5]. Recently, they have also been used to analyse stochastic dynamic systems [16]. A typical problem is to find the most probable state after a given observation sequence and perhaps other constraints. For example, in speech recognition, a sequence of sound recordings is given, and the sentence that has probably been pronounced is sought.

As for Markov chains, model checking and other algorithms depend on the size of the HMM, which is usually very large. Bisimulation equivalences have been shown to be an effective way to amend the state space problem for Markov chains [10]. In contrast, behavioral equivalences for HMMs have only been introduced recently by Castro et al. [4]. It is, however, not clear whether such equivalences agree with the logical properties in HMMs. To pave the way for efficient algorithms using reduction techniques, we study various bisimulation equivalences and characterise them logically with variants of the logic POCTL* (probabilistic observation-CTL*) introduced in [19].

Contributions. Our main contribution is the *logical characterisation* for three variants of bisimulation for HMMs, and their corresponding *decision algorithms*. For standard state-based bisimulation, we show that the logic POCTL* is sound and complete. Since Markov chains are special instances of HMMs, this result conservatively extends the logical characterisation for Markov chains [2]. More interesting are the strong and weak belief bisimulations defined by [4]. (We shall follow [4] and call the two equivalence relations strong and weak belief bisimulation, although this differs from the usual distinction between strong and weak bisimulation.) We show that these relations are too coarse for POCTL*: the nested probabilistic operator, conjunction and some forms of the until operator can distinguish belief bisimilar states. We introduce two sublogics SBBL* and WBBL*, which correspond to strong and weak belief bisimilarity, respectively. The key difference between SBBL* and WBBL* is that the latter cannot describe requirements on the most probable state after a certain sequence of observations.

We also present decision algorithms for the bisimilarities. For standard bisimilarity, an extension of the usual partition refinement algorithm [13] is enough. Belief bisimilarity is a relation over distributions and cannot be computed with partition refinement. Instead, we extend the approach in [7]: we generate a linear equation system that is satisfied by two belief states iff they are bisimilar. The time to construct the system is in $\mathcal{O}(|S|^3)$ for weak and in $\mathcal{O}(|S|^3 \cdot |\Omega|)$ for strong belief bisimilarity, where $|S|$ is the number of states and $|\Omega|$ the number of observations. Since the bisimulation for labelled Markov chains considered in [7] can be regarded as a special case of strong belief bisimulation, our results apply also in that setting. This produces another logical characterisation. More interestingly, our algorithm improves their complexity $\mathcal{O}(|S|^4)$.

We believe that our results are of practical relevance. We have identified the properties corresponding to the bisimulation relations considered, so the model checker can choose the appropriate relation and reduce the size of the HMM under consideration, using our efficient decision algorithm. Such characterisation and decision algorithms will make it possible to analyse HMMs of larger size.

Organisation of the paper. In Section 2 we recall the definition of HMM, belief states, probabilistic measures on them and the logic POCTL*. Section 3 discusses the three different notions of bisimulation for HMMs. The corresponding logical characterisations are presented in Section 4. The decision algorithm is presented in Section 5. We discuss related work in Section 6.

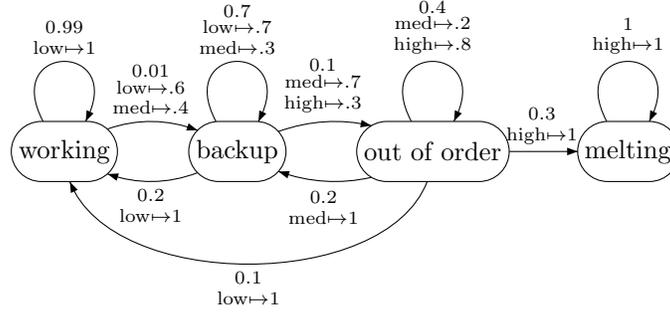


Fig. 1. A hidden Markov model for a cooling system

2 Hidden Markov models and the logic POCTL*

In this section we recall the definition of hidden Markov models (HMM) [4] and some related notions. On that basis, we can define the logic POCTL*.

2.1 Hidden Markov models

Definition 1. A hidden Markov model is a sextuple $M = (S, P, L, \Omega, O, \alpha)$, where S is a finite set of states; $P : S \times S \rightarrow [0, 1]$ is a probabilistic transition relation satisfying $\sum_{s' \in S} P(s, s') = 1$ for every $s \in S$; $L : S \times AP \rightarrow \{0, 1\}$ describes the truth values of atomic propositions; Ω is a finite set of observations; the partial function $O : S \times S \rightarrow \text{Dist}(\Omega)$ assigns a probability distribution over the observations to each transition in $P^{-1}((0, 1])$; $\alpha : S \rightarrow [0, 1]$ is the initial distribution.

Note that we assign observations to transitions. Many other definitions assign observations to states, but in that case, the observations would be almost the same as atomic propositions. Our choice is inspired by [4].

Example 1. In Fig. 1, a simple HMM that describes a small part of a nuclear power plant is depicted. It describes the state of the cooling system and how much information about this state can be obtained based on the incomplete information provided by temperature sensors, a situation that may occur in a partially broken power plant. For example, if the temperature sensor produces a “high” reading, it is not completely clear whether the power plant is melting down, renouncing all hope of repair, or it is only in state “out of order”, so a repair should be attempted.

2.2 Belief states

In this and the following sections, we assume we are given a fixed set of atomic propositions AP and a hidden Markov model $M = (S, P, L, \Omega, O, \alpha)$.

In a hidden Markov model, only the observation can be seen, and a standard problem is to guess the real state of the HMM based on the observations. We can summarize the history of observations in a *belief state* (or *information state*) [14].

Definition 2. *A belief state is a probability distribution over S . Moreover, we let $\mathbf{1}_s$ be the characteristic belief state for $s \in S$ defined by: $\mathbf{1}_s(s) = 1$.*

A belief state is not really a state of the HMM. Rather, it is a way to describe what we know about the state. The set of all belief states is called the *belief space* and is denoted by \mathcal{B} . The labelling function can easily be extended to belief states by: $L(b, a) := \sum_{s \in S} b(s) \cdot L(s, a)$. Intuitively, $L(b, a)$ gives the probability of satisfying a in belief state b .

The belief state b_n at time $n \geq 0$, i. e. the distribution over S at time n given the observation history $\omega_0, \dots, \omega_{n-1}$, captures all information about the past. We can inductively calculate the next belief state b_{n+1} based on the previous belief state b_n and the current observation ω_n . More details will be given after introducing probability spaces for HMMs.

2.3 Paths in HMM and probability spaces over paths

Given $M = (S, P, L, \Omega, O, \alpha)$ (as fixed above), we first introduce some notation. A path σ of M is a sequence $s_0, \omega_0, s_1, \omega_1, \dots \in (S \times \Omega)^\omega$ where $P(s_i, s_{i+1}) > 0$ and $O(s_i, s_{i+1})(\omega_i) > 0$ for all $i \in \mathbb{N}$. For $i \in \mathbb{N}$, let $\sigma(i)_s = s_i$ denote the $(i+1)$ th state of σ , and $\sigma(i)_o = \omega_i$ denote the $(i+1)$ st observation of σ . Let $\sigma(i \dots)$ denote the suffix path of σ starting with $\sigma(i)_s$, i. e., $s_i, \omega_i, s_{i+1}, \omega_{i+1}, \dots$.

Let $Path^M$ denote the set of all paths in M , and $Path^M(s)$ denote the set of paths in M that start in s . The superscript M is omitted whenever it is clear from the context. We define a probability space on paths of M using the standard cylinder construction. For a finite state–observation sequence $s_0, \omega_0, s_1, \omega_1, \dots, s_n$, its induced *basic cylinder set* is $\mathcal{C}(s_0, \omega_0, s_1, \omega_1, \dots, s_n) := \{\sigma \in Path \mid \forall i \leq n : \sigma(i)_s = s_i \wedge \forall j < n : \sigma(j)_o = \omega_j\}$. This set consists of all paths σ starting with $s_0, \omega_0, s_1, \omega_1, \dots, s_n$. Let \mathcal{Cyl} contain all basic cylinder sets for all finite state–observation sequences. Given a finite sequence $C_0, \mathcal{Y}_0, C_1, \mathcal{Y}_1, \dots, C_n$ of state sets and observation sets, we define the *cylinder set* to be the (disjoint) union of the basic cylinder sets with state–observation sequences picked from the sequence of sets:

$$\mathcal{C}(C_0, \mathcal{Y}_0, \dots, C_n) := \bigcup_{s_0 \in C_0} \bigcup_{\omega_0 \in \mathcal{Y}_0} \dots \bigcup_{s_n \in C_n} \mathcal{C}(s_0, \omega_0, \dots, s_n)$$

Given a belief state b , we define the premeasure $Prob_b$ on \mathcal{Cyl} by induction on n as: $Prob_b(\mathcal{C}(s_0)) = b(s_0)$ and, for $n > 0$, $Prob_b(\mathcal{C}(s_0, \omega_0, \dots, s_n))$ equals: $P(s_{n-1}, s_n)O(s_{n-1}, s_n)(\omega_{n-1}) \cdot Prob_b(\mathcal{C}(s_0, \omega_0, \dots, s_{n-1}))$. By induction, we get:

$$Prob_b(\mathcal{C}(s_0, \omega_0, \dots, s_n)) = b(s_0) \prod_{i=1}^n O(s_{i-1}, s_i)(\omega_{i-1}) P(s_{i-1}, s_i)$$

The above premeasure can be extended uniquely (Carathéodory's theorem, see e. g. [17, page 272]) to a measure on the σ -algebra generated by \mathcal{Cyl} . We introduce a few shorthand notations, which will be used frequently later on:

- $Prob_b(\omega, s') := \sum_{s \in S} O(s, s')(\omega)P(s, s')b(s)$ is the probability to get observation ω and end in some state s' . So, $Prob_b(\omega, s') = Prob_b(\bigcup_{s \in S} \mathcal{C}(s, \omega, s'))$. For a set of states $A \subseteq S$, let $Prob_b(\omega, A) := \sum_{s' \in A} Prob_b(\omega, s')$.
- $Prob_b(\omega) := Prob_b(\omega, S)$ is the probability to get observation ω in belief state b . For a set of observations \mathcal{Y} , let $Prob_b(\mathcal{Y}) := \sum_{\omega \in \mathcal{Y}} Prob_b(\omega)$.
- $\tau(b, \omega)(s') := \frac{Prob_b(\omega, s')}{Prob_b(\omega)}$. Then, $\tau(b, \omega)$ is the resulting belief state under the condition that we take a transition from belief state b and that we get observation ω .
- $Prob_b(b') := \sum_{\omega \in \Omega} Prob_b(\omega) \cdot \mathbf{1}_{b'=\tau(b, \omega)}$ is the probability of getting to b' in the next step, starting from b . Here $\mathbf{1}_{b'=\tau(b, \omega)}$ equals 1 if $b' = \tau(b, \omega)$, and 0 otherwise. For a set of belief states B , we define $Prob_b(B) := \sum_{b' \in B} Prob_b(b')$.

For belief state $b = \mathbf{1}_s$, we sometimes write s when clear from the context. The updating of belief state described above can now be written by: $b_{n+1} = \tau(b_n, \omega_n)$.

2.4 Syntax of POCTL*

In our article [19], we defined a logic POCTL* to describe properties of HMMs. In POCTL*, we distinguish state formulas (denoted Φ), path formulas (denoted φ), and belief state formulas (denoted ε). Its syntax is:

$$\begin{aligned} \Phi &::= \mathbf{true} \mid a \mid \neg\Phi \mid \Phi \wedge \Phi \mid \varepsilon \\ \varphi &::= \Phi \mid \neg\varphi \mid \varphi \wedge \varphi \mid X_{\mathcal{Y}}\varphi \mid \varphi \mathcal{U}^{\leq n} \varphi \\ \varepsilon &::= \neg\varepsilon \mid \varepsilon \wedge \varepsilon \mid \mathcal{P}_{\bowtie p}(\varphi) \end{aligned}$$

where a is an atomic proposition, \mathcal{Y} is a set of observations, n is a natural number or ∞ , \bowtie is a comparison operator $\in \{<, \leq, \geq, >\}$, and p is a probability bound $\in [0, 1]$.³

The disjunction \vee is defined as usual as an abbreviation. If $\mathcal{Y} = \Omega$, we will sometimes suppress the index of a next-state operator: $X\varphi := X_{\Omega}\varphi$. The future-operator $\diamond^{\leq n}\varphi$ abbreviates $\mathbf{true} \mathcal{U}^{\leq n} \varphi$.

The semantics of Φ and φ is mostly defined in the same way as for CTL over states and paths, respectively [19, 1]. A few examples: $s \models \varepsilon$ iff $\mathbf{1}_s \models \varepsilon$, $\sigma \models X_{\mathcal{Y}}\varphi$ iff $\sigma(0)_o \in \mathcal{Y}$ and $\sigma(1\dots) \models \varphi$, and $b \models \mathcal{P}_{\bowtie p}(\varphi)$ iff $Prob_b\{\sigma \mid \sigma \models \varphi\} \bowtie p$. POCTL* can be applied to the typical problem (based on the sequence of observations and perhaps other constraints, find a probable state) by verifying a formula like $\mathcal{P}_{\geq 0.25}(X_{\omega_1} X_{\omega_2} X_{\omega_3} \mathbf{true})$.

³ Some formulas, e. g. state formula $\neg\varepsilon$, have two derivations: either use the negation of state formulas or the negation of belief state formulas. However, this will not pose problems because the two are semantically equivalent.

3 Bisimulation notions for HMMs

In this section we define various bisimulations for HMMs. First, one can simply extend standard bisimulation of Markov chains [12] to the HMM setting:

Definition 3. Let $R \subseteq S \times S$ be an equivalence relation on the states of M . R is a strong bisimulation if it respects the following conditions for every $(s, t) \in R$:

1. For all atomic propositions $a \in AP$, we have $s \models a$ iff $t \models a$.
2. For all observations $\omega \in \Omega$, we have $Prob_{\mathbf{1}_s}(\omega) = Prob_{\mathbf{1}_t}(\omega)$.
3. For all observations $\omega \in \Omega$, and all R -equivalence classes $C \in S/R$, we have $\tau(s, \omega)(C) = \tau(t, \omega)(C)$.

Two states $s, t \in S$ are strongly bisimilar if there exists a strong bisimulation R with $s R t$. We denote this as $s \sim t$. Bisimilarity can be extended to paths: two paths $\sigma, \rho \in Path$ are strongly bisimilar if $\sigma(i)_o = \rho(i)_o$ and there exists a strong bisimulation R such that $\sigma(i)_s R \rho(i)_s$ for all $i \in \mathbb{N}$.

Note that Conditions 2 and 3 can be subsumed to: For all observations $\omega \in \Omega$ and all R -equivalence classes $C \in S/R$, we have $Prob_s(\omega, C) = Prob_t(\omega, C)$. Since probabilities agree on bisimilar states, we sometimes denote $Prob_s$ by $Prob_{[s]_R}$. The definition conservatively extends bisimilarity on Markov chains: if $|\Omega| = 1$, HMM bisimilarity reduces to standard bisimilarity for Markov chains [12].

The state-based bisimulation defined above does not take into account that states in HMMs are hidden, i. e., only indirectly observable. Recently, Castro et al. [4] introduced two new notions of bisimulation relations, not on the states of the HMM, but on the belief states, i. e., on distributions over states. We recall their definitions and adapt them to our fully probabilistic setting.

Definition 4. Let $R \subseteq \mathcal{B} \times \mathcal{B}$ be an equivalence relation on the belief states. R is a strong belief bisimulation if it respects the following conditions for every $(b, c) \in R$:

1. For all atomic propositions $a \in AP$, we have $L(b, a) = L(c, a)$.
2. For all observations $\omega \in \Omega$, we have $Prob_b(\omega) = Prob_c(\omega)$.
3. For all observations $\omega \in \Omega$, we have $\tau(b, \omega) R \tau(c, \omega)$.

Two belief states $b, c \in \mathcal{B}$ are strongly belief bisimilar if there exists a strong belief bisimulation R with $b R c$. This is denoted $b \sim_{sb} c$.

The first condition requires that b and c have the same labelling. The second condition states that the probability of observing ω is the same from b or c . The new condition is the third one, stating that the updated belief states with respect to ω must also be in the relation R . It is weaker than the third condition of state-based bisimulation: The following example illustrates the difference.

Example 2. Consider the HMM depicted in Fig. 2. Assume $L(s_3) \neq L(s_4)$, and other states have the same labelling. First, $s_1 \not\sim t_1$, independent of the observations. The reason is that s_2 cannot be bisimilar with either t_2 or t_3 . Now let $b = \mathbf{1}_{s_1}$ and $c = \mathbf{1}_{t_1}$. It is easy to verify that $b \sim_{sb} c$.

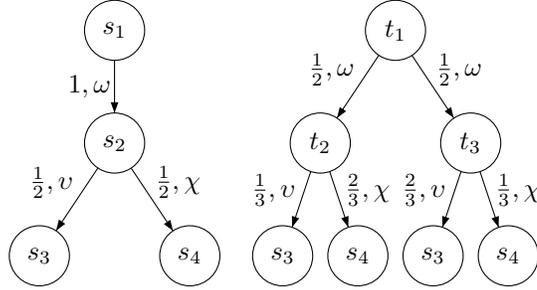


Fig. 2. State-based strong bisimulation and strong belief bisimulation differ.

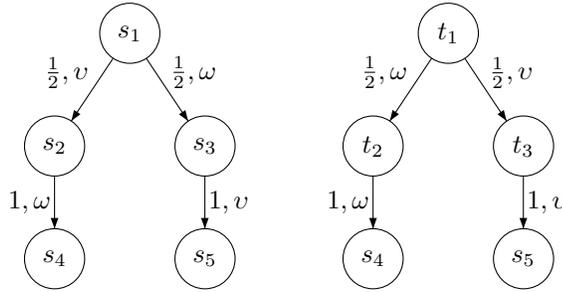


Fig. 3. Strong and weak belief bisimilarity differ.

Now we recall weak belief bisimulation for HMMs, based on [4]:

Definition 5. Let $R \subseteq \mathcal{B} \times \mathcal{B}$ be an equivalence relation on the belief states. R is a weak belief bisimulation if it respects the following conditions for every $(b, c) \in R$:

1. For all atomic propositions $a \in AP$, we have $L(b, a) = L(c, a)$.
2. For all observations $\omega \in \Omega$, we have $Prob_b(\omega) = Prob_c(\omega)$.
3. For all R -equivalence classes $B \in \mathcal{B}/R$, we have $Prob_b(B) = Prob_c(B)$.

Two belief states $b, c \in \mathcal{B}$ are weakly belief bisimilar if there exists a weak belief bisimulation R with $b R c$. This is denoted $b \sim_{wb} c$.

Indeed, it holds that $\sim_{sb} \subset \sim_{wb}$, where the inclusion is strict [4]. Intuitively, while strong belief bisimulation requires that the updated belief states must be in the relation, in weak belief bisimulation we require only that the updated belief states evolve with the same probability to each $B \in \mathcal{B}/R$.

The example in Fig. 3, taken from [4], illustrates the difference: $\mathbf{1}_{s_1}$ and $\mathbf{1}_{t_1}$ are not strongly belief bisimilar, but they are weakly belief bisimilar.

4 Characterising bisimilarity

This section presents the logical characterisation results for the three bisimilarities for HMMs. We first show that state-based bisimilarity agrees with the logical equivalence induced by POCTL*. Then, we shall identify two sublogics of POCTL* to characterise strong and weak belief bisimilarities, respectively.

4.1 Strong bisimilarity

We show that the equivalence induced by POCTL* agrees with state-based bisimilarity. As a preparation, we introduce bisimulation-closed sets of paths.

Definition 6. *A set of paths is bisimulation-closed if it is a (disjoint) union of equivalence classes induced by strong bisimilarity on paths.*

Lemma 1. *Assume that s is strongly bisimilar to t . Then, for all bisimulation-closed sets of paths Π , we have that $Prob_s(\Pi) = Prob_t(\Pi)$.*

Proof. It is enough to show equality for a \cap -closed generator of the σ -algebra of all bisimulation-closed sets of paths. Therefore, assume w.l.o.g. that Π is a cylinder set $\mathcal{C}(C_0, \omega_0, C_1, \omega_1, \dots, C_n)$, where the C_i are bisimulation equivalence classes, and assume that $s \in C_0$. Bisimilarity implies $t \in C_0$. Clearly,

$$Prob_s(\Pi) = Prob_{C_0}(\omega_0, C_1) \cdot Prob_{C_1}(\omega_1, C_2) \cdots Prob_{C_{n-1}}(\omega_{n-1}, C_n) = Prob_t(\Pi)$$

where $Prob_{C_i} = Prob_{s_i}$ for some $s_i \in C_i$; as C_i is a bisimulation equivalence class, $Prob_{C_i}$ is well-defined. The intersection of two such cylinder sets is either the smaller of the two or empty.

The following theorem shows that the equivalence induced by the logic POCTL* agrees with strong bisimulation:

Theorem 1. *The logic POCTL* characterises strong bisimilarity, i. e., two states are strongly bisimilar iff they satisfy the same POCTL* state formulas, and two paths are (statewise) strongly bisimilar iff they satisfy the same POCTL* path formulas.*

The proof is mostly based on the proof of Theorem 10.67 of [1], adapted to the setting of HMMs – for details see Appendix A. The completeness proof does not rely on the until operator being part of the logic; therefore, the sublogic of POCTL* without until formulas is sufficient to characterise state-based strong bisimilarity. Thus, it conservatively extends the result for Markov chains [1].

4.2 Strong belief bisimilarity

In this section we will present a logical characterisation of strong belief bisimilarity. First, in Subsection 4.2, we will discuss that several operators of POCTL* are too discriminative with respect to belief bisimilarity. Then, we define the logic SBBL*, which characterises strong belief bisimilarity.

POCTL* is too discriminative. In the example of Fig. 4, we illustrate why we shall have to remove a few operators to characterise strong belief bisimilarity. Every transition in the HMM produces the same observation.

- The nested probabilistic operator $\varepsilon_1 := \mathcal{P}_{\geq 0.5}(\mathcal{P}_{\geq 1}(X a_3))$. Consider belief state b_1 defined by $b_1(s_2) = b_1(s_4) = 0.5$, and b_2 defined by $b_2(s_1) = b_2(s_3) = 0.5$. It follows that $b_1 \sim_{\text{sb}} b_2$, but $b_1 \models \varepsilon_1$, while $b_2 \not\models \varepsilon_1$. The distinguishing power of ε_1 comes from the fact that s_2 (in the support of b_1) satisfies the inner probabilistic formula, whereas no state in the support of b_1 does so.
- The conjunction $\varepsilon_2 := \mathcal{P}_{\geq 0.5}(a_1 \wedge a_2)$. For the belief states b_1 and b_2 defined above, it holds then $b_2 \models \varepsilon_2$ but $b_1 \not\models \varepsilon_2$.
- The conjunction after the path operator $\varepsilon_3 := \mathcal{P}_{\geq 0.5}(X(a_1 \wedge a_2))$, and belief states $\mathbf{1}_{s_7} \sim_{\text{sb}} \mathbf{1}_{s_8}$. We again have $\mathbf{1}_{s_7} \not\models \varepsilon_3$ but $\mathbf{1}_{s_8} \models \varepsilon_3$.
- The until formula $(X a_1) \mathcal{U}^{\leq \infty} a_2$ is satisfied by paths in $\mathcal{C}(s_8, \omega, s_3)$, but not by any path starting in s_7 . Therefore, $\mathbf{1}_{s_7} \models \mathcal{P}_{=0}((X a_1) \mathcal{U}^{\leq \infty} a_2)$, but $\mathbf{1}_{s_8}$ does not satisfy this formula.
- The nested until formula $\neg a_1 \mathcal{U}^{\leq \infty}(a_2 \mathcal{U}^{\leq \infty} a_3)$ holds on paths in $\mathcal{C}(s_8, \omega, s_3, \omega, s_6)$, so similarly $\mathbf{1}_{s_7} \models \mathcal{P}_{=0}(\neg a_1 \mathcal{U}^{\leq \infty}(a_2 \mathcal{U}^{\leq \infty} a_3))$.

The logic SBBL*. Based on the discussion above, we present a sublogic of POCTL* to characterise strong belief bisimilarity. We call this logic SBBL*:

$$\begin{aligned} \Phi &::= \mathbf{true} \mid a \mid \neg \Phi \\ \varphi &::= \Phi \mid X_{\mathcal{Y}} \varphi \\ \varepsilon &::= \neg \varepsilon \mid \varepsilon \wedge \varepsilon \mid \mathcal{P}_{\bowtie p}(\varphi) \\ &\quad \mid \mathcal{P}_{\bowtie p}(\Phi \mathcal{U}^{\leq n} \Phi) \end{aligned}$$

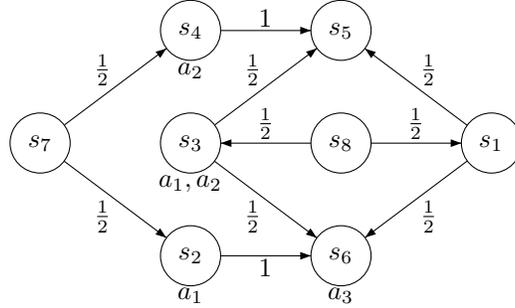


Fig. 4. A hidden Markov model

Theorem 2. *The logic SBBL* characterises strong belief bisimilarity, i. e., two belief states are strongly belief bisimilar iff they satisfy the same SBBL* belief state formulas.*

Proof. We prove *soundness* by induction over the structure of the formulas; in contrast to Theorem 1, the induction runs only over the belief state formulas.

We assume given two belief states $b \sim_{\text{sb}} c$ and a belief state SBBL*-formula ε ; we prove that $b \models \varepsilon$ iff $c \models \varepsilon$. For symmetry reasons, it is enough to prove one direction, so assume that $b \models \varepsilon$; then it remains to be proven that $c \models \varepsilon$.

- $\varepsilon = \neg \varepsilon'$ and $\varepsilon = \varepsilon_1 \wedge \varepsilon_2$. These two cases are simple consequences of the induction hypothesis.
- $\varepsilon = \mathcal{P}_{\bowtie p}(\mathbf{true})$ or $\mathcal{P}_{\bowtie p}(\neg \mathbf{true})$. Trivial.
- $\varepsilon = \mathcal{P}_{\bowtie p}(a)$ or $\mathcal{P}_{\bowtie p}(\neg a)$. A simple consequence of Condition 1 of Def. 4.
- $\varepsilon = \mathcal{P}_{\bowtie p}(\varphi)$, where $\varphi = X_{\mathcal{Y}_1} X_{\mathcal{Y}_2} \dots X_{\mathcal{Y}_k} \Phi$. Let Π be the set of paths satisfying φ . So, $\Pi = \{\sigma \mid \sigma(0)_o \in \mathcal{Y}_1 \wedge \sigma(1)_o \in \mathcal{Y}_2 \dots \sigma(k-1)_o \in \mathcal{Y}_k \wedge \sigma(k \dots) \models \Phi'\}$.

Note that $Prob_b(\Pi)$ is a product of factors of the form $Prob_b(\omega_1)$ for $\omega_1 \in \mathcal{Y}_1$, $Prob_{\tau(b,\omega_1)}(\omega_2)$ for $\omega_2 \in \mathcal{Y}_2$, all constructed using $Prob_{(\cdot)}$ and $\tau(\cdot, \cdot)$. Similarly, $Prob_c(\Pi)$ can be described using $Prob_c(\omega_1)$, $Prob_{\tau(c,\omega_1)}(\omega_2)$ etc. All these terms for b and c are equal, because $\tau(b,\omega_1) \sim_{sb} \tau(c,\omega_1)$ for all ω_1 (Condition 3 of Def. 4), $Prob_b(\omega_1) = Prob_c(\omega_1)$ (Condition 2 of Def. 4), etc.

– $\varepsilon = \mathcal{P}_{\bowtie p}(\Phi_1 \mathcal{U}^{\leq n} \Phi_2)$. First assume that $n < \infty$. We evaluate this property on a modified HMM M' . It has the same states, labels and observations as M , but $(\Phi_2 \vee \neg \Phi_1)$ -states are made absorbing. This does not change the truth values of Φ_1 or Φ_2 . Further, once a path has reached a $(\Phi_2 \vee \neg \Phi_1)$ -state, it has become clear whether it satisfies φ . So modifying transitions out of these states does not change the truth value of ε . On M' , the formula ε is equivalent to $\mathcal{P}_{\bowtie p}(X X \cdots X \Phi_2)$ (n next-operators); then, the argumentation for the next-operator can be used to complete the proof.

Now, if $n = \infty$, note that the sequence $(Prob_b(\Phi_1 \mathcal{U}^{\leq i} \Phi_2))_{i \in \mathbb{N}}$ is a nondecreasing sequence in a compact interval, so it does have a limit, which is $Prob_b(\Phi_1 \mathcal{U}^{\leq \infty} \Phi_2)$. The corresponding sequence for $Prob_c$ consists of the same elements, so it must have the same (unique) limit.

This finishes the proof of soundness. To show *completeness*, we define the equivalence relation on belief states $R := \{(b, c) \mid \forall \text{ SBBL}^*\text{-belief state formulas } \varepsilon : b \models \varepsilon \text{ iff } c \models \varepsilon\}$. We have to show that this relation is a strong belief bisimulation. Assume given two belief states b and c such that $b R c$.

- Condition 1. One sees easily that $L(b, a) = \sup \{r \mid b \models \mathcal{P}_{\geq r}(a)\}$. Obviously, $\{r \mid b \models \mathcal{P}_{\geq r}(a)\} = \{r \mid c \models \mathcal{P}_{\geq r}(a)\}$; therefore $L(b, a) = L(c, a)$.
- Condition 2. The same reasoning with $\sup \{r \mid b \models \mathcal{P}_{\geq r}(X_\omega \mathbf{true})\} = Prob_b(\omega)$.
- Condition 3. Assume given any $\omega \in \Omega$. We prove that $b' := \tau(b, \omega) R \tau(c, \omega) =: c'$. Assume given a belief state formula ε such that $b' \models \varepsilon$; if we can prove that $c' \models \varepsilon$, then we get the desired result.

First assume that ε has the special form $\mathcal{P}_{\bowtie p}(\varphi)$. Then, $b \models \mathcal{P}_{\bowtie p, Prob_b(\omega)}(X_\omega \varphi)$, as $Prob_b(X_\omega \varphi) = Prob_b(\omega) \cdot Prob_{b'}(\varphi)$. From the definition of R , we know that $c \models$ the same formula, and therefore $c' \models \mathcal{P}_{\bowtie p}(\varphi)$.

Now assume that ε is constructed from the special form above using negation and conjunction, then a trivial induction over the structure of ε shows $c' \models \varepsilon$.

Again, from the completeness proof we see that the sublogic of SBBL* without until formulas is sufficient to characterise strong belief bisimilarity.

4.3 Weak belief bisimilarity

In this section we present logical characterisation results for weak belief bisimilarity. We restrict SBBL* further to the following logic, named WBBL*:

$$\begin{aligned} \Phi &::= \mathbf{true} \mid a \mid \neg \Phi \\ \varphi &::= \Phi \mid X_\Gamma \mathbf{true} \mid X \varphi \\ \varepsilon &::= \neg \varepsilon \mid \varepsilon \wedge \varepsilon \mid \mathcal{P}_{\bowtie p}(\varphi) \mid \mathcal{P}_{\bowtie p}(\Phi \mathcal{U}^{\leq n} \Phi) \end{aligned}$$

Essentially, the operator $X_{\mathcal{Y}}\varphi$ in SBBL* is replaced by two subformulas $X_{\mathcal{Y}}\mathbf{true}$ and $X\varphi$. Note that properties like $\mathcal{P}_{\geq 0.25}(X_{\omega_1}X_{\omega_2}X_{\omega_3}\mathbf{true})$ are not in WBBL*, so it cannot be used to describe to solve the corresponding standard problem. The following theorem shows the main result:

Theorem 3. *The logic WBBL* characterises weak belief bisimilarity, i. e., two belief states are weakly belief bisimilar iff they satisfy the same WBBL*-belief state formulas.*

Proof. We proceed as in the previous two cases. To prove *soundness*, assume given two belief states b and c that are weakly belief bisimilar and a belief state WBBL*-formula ε such that $b \models \varepsilon$. We have to prove that $c \models \varepsilon$.

- $\varepsilon = \neg\varepsilon'$, $\varepsilon = \varepsilon_1 \wedge \varepsilon_2$, $\varepsilon = \mathcal{P}_{\bowtie p}(\mathbf{true})$, $\mathcal{P}_{\bowtie p}(\neg\mathbf{true})$, $\mathcal{P}_{\bowtie p}(a)$, $\mathcal{P}_{\bowtie p}(\neg a)$, or $\mathcal{P}_{\bowtie p}(\Phi \mathcal{U}^{\leq n} \Psi)$. These cases are handled as in Theorem 2.
- $\varepsilon = \mathcal{P}_{\bowtie p}(X_{\mathcal{Y}}\mathbf{true})$: The set Π of paths that satisfy $X_{\mathcal{Y}}\mathbf{true}$ has probability $Prob_b(\Pi) = Prob_b(\mathcal{Y})$. From Condition 2 of Def. 5, it follows that this is equal to $Prob_c(\mathcal{Y}) = Prob_c(\Pi)$.
- $\varepsilon = \mathcal{P}_{\bowtie p}(X\varphi)$. From the induction hypothesis, we can conclude that $b' \sim_{wb} c'$ implies $b' \models \mathcal{P}_{> p}(\varphi)$ iff $c' \models \mathcal{P}_{> p}(\varphi)$, so for every weak belief bisimilarity class $B \in \mathcal{B}/\sim_{wb}$, $Prob_B(\varphi)$ is well-defined. Therefore, $Prob_b(X\varphi) = \sum_{B \in \mathcal{B}} Prob_B(\varphi) \cdot Prob_b(B)$, and $Prob_c(X\varphi) = \sum_{B \in \mathcal{B}} Prob_B(\varphi) \cdot Prob_c(B)$. The right-hand sides are equal because of Condition 3 of Def. 5.

To show *completeness*, we define the equivalence relation on belief states

$$R := \{(b, c) \mid \forall \text{ WBBL*}-\text{belief state formulas } \varepsilon : b \models \varepsilon \text{ iff } c \models \varepsilon\}$$

We have to show that this relation is a weak belief bisimulation. Assume given two belief states b and c such that $b R c$.

- Conditions 1 and 2 are handled as in Theorem 2.
- Condition 3. Assume given any R -equivalence class B . We prove $Prob_b(B) = Prob_c(B)$ by regarding the satisfaction sets $Sat(\varepsilon)$ for all WBBL*-belief state formulas with rational probability bounds, i. e., every subformula $\mathcal{P}_{\bowtie p}(\cdot)$ has $p \in \mathbb{Q}$. Let $Sat_{\mathbb{Q}}$ contain all such satisfaction sets, and let \mathcal{F} be the σ -algebra generated from $Sat_{\mathbb{Q}}$. Then, $B \in \mathcal{F}$, because B is a countable intersection of elements of $Sat_{\mathbb{Q}}$. Note that $Sat_{\mathbb{Q}}$ is \cap -closed. Therefore, if two premeasures agree on $Sat_{\mathbb{Q}}$, then their extensions to measures on \mathcal{F} also agree. From the definition of R , it follows easily that $Prob_b(\varphi) = Prob_c(\varphi)$ for any belief states $b R c$, because $Prob_b(\varphi) = \sup\{q \in \mathbb{Q} \mid b \models \mathcal{P}_{> q}(\varphi)\}$. So, $Prob_b$ and $Prob_c$ agree on $Sat_{\mathbb{Q}}$.

5 Decision algorithms

In this section we present decision algorithms for the three different bisimilarities. The state-based strong bisimilarity is the easiest one, as it can be computed

by a simple extension of the usual partition refinement algorithm [13, 6, 10]. The complexity is linear in the number of transitions and observations and logarithmic in the number of states. We do not go further into that matter, as details can be found in [3].

As the belief states are probability distributions, the set of belief states is uncountable. Therefore, one cannot describe the belief state bisimulation quotient as a partition of the state space as for standard bisimilarity or ordinary lumping of Markov chains. Another approach has been proposed by [7]: two belief states b and c are belief bisimilar if they are a solution to a specific equation system over $b(s)$ and $c(s)$, for all $s \in S$. We adapt their algorithm to our setting and show an improved time bound. The equation system is constructed as follows.

Let $\{s_1, s_2, \dots, s_n\}$ be an order of the states. We denote $b(s_j)$ as b_j and $c(s_j)$ as c_j ; these variables will be the unknowns in the system. We construct the equation system iteratively. We start with the system

$$\bigwedge_{a \in AP} \sum_{s_i | = a} b_i - c_i = 0 \wedge \bigwedge_{\omega \in \Omega} \sum_{s_i \in S} Prob_{s_i}(\omega) \cdot (b_i - c_i) = 0$$

The base case is the same for strong and weak belief bisimilarity: the first conjunction corresponds to the condition on the labelling, and the second one to the condition that the probability of observing $\omega \in \Omega$ agrees with b and c . Considering b and c and row vectors, this equation system can be written as

$$(A_1 \quad -A_1) \cdot (b, c)^T = \mathbf{0}$$

where A_1 is an $(|AP| + |\Omega|) \times n$ -matrix. We assume that A_1 is brought to upper triangular form (i. e., a matrix with zeroes below the main diagonal) immediately, and the equations that turn out to be linearly dependent are removed. Let k_1 be the number of rows in A_1 (after the triangular transformation), i. e., $k_1 \leq |AP| + |\Omega|$. There can be at most n linearly independent equations of this form (since A_1 has n columns); this property will be used to ensure termination. If $k_1 = n$, we stop immediately.

5.1 Deciding weak belief bisimilarity

Now we describe the iteration step for weak belief bisimilarity – corresponding to the third condition of weak belief bisimulation in Def. 5. In the i th iteration step, we assume given an equation system of the form

$$\begin{pmatrix} A_1 & -A_1 \\ \vdots & \vdots \\ A_i & -A_i \end{pmatrix} \cdot (b, c)^T = \mathbf{0} \tag{1}$$

with at most $n - 1$ equations (n equations cannot occur because the algorithm would have terminated earlier in that case), all of them linearly independent, in upper triangular form. From this, we construct an extended equation system of the same form, but possibly with more equations:

$$\begin{pmatrix} A_1 & -A_1 \\ \vdots & \vdots \\ A_i & -A_i \\ A_{i+1} & -A_{i+1} \end{pmatrix} \cdot (b, c)^T = \mathbf{0} \quad (2)$$

If it does not have more equations, we have reached a fixpoint. In that case, or if the new system has n equations, we can stop after the i th iteration step.

To find A_{i+1} , we first add new equations to the system: the new equations are produced from equations in (1) by replacing the variable b_j with $\sum_{\omega \in \Omega} Prob_b(\omega, s_j)$ and replacing the c_j with $\sum_{\omega \in \Omega} Prob_c(\omega, s_j)$. It is enough to add the new equations for the rows of A_i , as equations for A_1, A_2, \dots, A_{i-1} have been added earlier. This adds at most k_i equations, where k_i is the number of rows in A_i . Then, we bring the matrix in equation 2 with all these new equations into upper triangular form, to find out which ones are linearly dependent. As A_1, \dots, A_i are already in upper triangular form, we only have to do calculations with A_{i+1} . Finally, we drop the linearly dependent equations, giving us $k_{i+1} \leq k_i$ additional equations.

Time complexity. The algorithm generates an equation system in upper triangular form. It basically interleaves (i) steps where k_i new equations are generated, corresponding to A_{i+1} , with (ii) steps where these new equations are brought into upper triangular form, and the linear dependent ones are removed. Some equations then turn out to be linearly dependent; this will happen exactly $|AP| + |\Omega|$ times in total, because we started with this number of equations. To see this, remember that every single row in A_1 (i. e., a single equation) is transferred to A_2, A_3, \dots by the variable substitution described above. In one of those transfers, the generated equation turns out to be linearly dependent, and from that iteration on, it is dropped completely. Therefore, at most $n + |AP| + |\Omega|$ equations over $2n$ variables are generated. The most costly step is to turn the equation system into upper triangular form. For all equations together, this takes time $\in \mathcal{O}(n^2(n + |AP| + |\Omega|))$. Notably, this improves the time bound $\mathcal{O}(n^4)$ in [7] by a factor of n , as in most cases $|AP| + |\Omega| \ll n$.

5.2 Deciding strong belief bisimilarity

The argumentation for strong belief bisimilarity is almost the same as for weak belief bisimilarity; only in the iteration step, one set of equations for each observation $\omega \in \Omega$ is generated – corresponding to the third condition of strong belief bisimulation in Def. 4. In the i th iteration step, assume that we start with an equation system of the form in (1). We similarly add new equations to it, but now, for every $\omega \in \Omega$, we add a set of equations where we replace b_j by $Prob_b(\omega, s_j)$ and c_j by $Prob_c(\omega, s_j)$. So, A_{i+1} consists of rows of the form $A_i \cdot P \cdot O(\dots, \dots)(\omega)$. It adds at most $k_i |\Omega|$ equations to the system.

We similarly bring all these equations into upper triangular form and eliminate the linearly dependent ones.

Time complexity. The final equation systems contains at most n equations. In the worst case, from every of these equations, we generated $|\Omega|$ new equations in some iteration, brought them into upper triangular form and found them (almost) all linearly dependent. So, at most $n|\Omega|$ equations over $2n$ variables have been generated. Turning them into upper triangular form takes time $\in \mathcal{O}(n^3|\Omega|)$.

6 Related work

The three bisimulation relations we have considered in this paper are based on existing definitions in the literature for Markov chains and their extensions. The state-based strong bisimulation was considered earlier in [3] and is a simple extension of the bisimulation for Markov chains [12], by incorporating the notion of observations in HMM. Our logic POCTL* [19] is an extension of the logic PCTL* introduced in [8]. Moreover, the logical characterisation result for state-based bisimulation is also a conservative extension of the logical characterisation result for Markov chains presented e. g. in [2].

The strong and weak belief bisimulations we have used were taken from [4], where they are defined for a general model with nondeterministic choices. The new concept here is to match distributions with distributions, instead of states with states as in the classical setting. This notion of equivalence has also been studied in [7], where bisimulation between distributions is defined for labelled Markov chains: strong belief bisimulation can be considered as an extension of the definition in [7] with the observation function attached to the transitions. In HMMs where all transitions generate the same trivial observation, it agrees with the definition in [7]. Thus, inspired by the work in [7], we have presented an algorithm for deciding strong belief bisimulation. As we have noted, our time bound improves theirs. Because of the mentioned connection to [7], our logical characterisation also carries over to the setting of labelled Markov chains.

Finally, we want to mention the recent related paper [11] in which the algorithm in [7] was – independently – improved to cubic as well: they have a similar observation as our paper by keeping the basis in a canonical orthogonal set. Moreover, they have proposed a randomized algorithm with quadratic complexity which could be applied in our setting as well.

Acknowledgements. This work was partly done while David N. Jansen was visiting MT-LAB at the Technical University of Denmark; he was partially supported by the NWO/DFG Bilateral Research Programme ROCKS, the EU FP7 under grant number ICT-214755 (Quasimodo) IDEA4CPS and by MT-LAB, a VKR Centre of Excellence.

References

1. Baier, C., Katoen, J.P.: Principles of model checking. MIT Press, Cambridge, MA (2008)

2. Baier, C., Katoen, J.P., Hermanns, H., Wolf, V.: Comparative branching-time semantics for Markov chains. *Inf. Comput.* 200(2), 149–214 (2005)
3. Bicego, M., Dovier, A., Murino, V.: Designing the minimal structure of hidden Markov model by bisimulation. In: Figueiredo, M., Zerubia, J., Jain, A.K. (eds.) *Energy Minimization Methods in Computer Vision and Pattern Recognition: ... EMMCVPR*. LNCS, vol. 2134, pp. 75–90. Springer, Berlin (2001)
4. Castro, P.S., Panangaden, P., Precup, D.: Equivalence relations in fully and partially observable Markov decision processes. In: Boutilier, C. (ed.) *IJCAI-09, proc. of the twenty-first intl. joint conference on artificial intelligence*. pp. 1653–1658. AAAI Press, Menlo Park, CA (2009)
5. Christiansen, H., Have, C.T., Lassen, O.T., Petit, M.: Inference with constrained hidden Markov models in PRISM. *TPLP* 10(4-6), 449–464 (2010)
6. Derisavi, S., Hermanns, H., Sanders, W.H.: Optimal state-space lumping in Markov chains. *Inf. proc. lett.* 87(6), 309–315 (2003)
7. Doyen, L., Henzinger, T.A., Raskin, J.F.: Equivalence of labeled Markov chains. *Int. J. Foundations of Computer Science* 19(3), 549–563 (2008)
8. Hansson, H., Jonsson, B.: A logic for reasoning about time and reliability. *Formal Asp. Comput.* 6(5), 512–535 (1994)
9. Jurafsky, D., Martin, J.H.: *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, Upper Saddle River, N.J. (2000)
10. Katoen, J.P., Kemna, T., Zapreev, I., Jansen, D.N.: Bisimulation minimisation mostly speeds up probabilistic model checking. In: Grumberg, O., Huth, M. (eds.) *Tools and algorithms for the construction and analysis of systems: ... TACAS*. LNCS, vol. 4424, pp. 87–101. Springer, Berlin (2007)
11. Kiefer, S., Murawski, A.S., Ouaknine, J., Wachter, B., Worrell, J.: Language equivalence for probabilistic automata. In: Gopalakrishnan, G., Qadeer, S. (eds.) *Computer Aided Verification: ... CAV*. LNCS, vol. 6806, pp. 526–540. Springer (2011)
12. Larsen, K.G., Skou, A.: Bisimulation through probabilistic testing. *Inf. Comput.* 94(1), 1–28 (1991)
13. Paige, R., Tarjan, R.E.: Three partition refinement algorithms. *SIAM J. Comput.* 16(6), 973–989 (1987)
14. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77(2), 257–286 (1989)
15. Salamatian, K., Vaton, S.: Hidden Markov modeling for network communication channels. In: *Proc. 2001 ACM SIGMETRICS intl. conf. on measurement and modeling of computer systems*. vol. 29, pp. 92–101. ACM, New York (2001)
16. Sistla, A.P., Zefran, M., Feng, Y.: Monitorability of stochastic dynamical systems. In: Gopalakrishnan, G., Qadeer, S. (eds.) *Computer Aided Verification: ... CAV*. LNCS, vol. 6806, pp. 720–736. Springer (2011)
17. Stein, E.M., Shakarchi, R.: *Real analysis: measure theory, integration, and Hilbert spaces*, Princeton lectures in analysis, vol. III. Princeton Univ. Pr., Princeton, NJ (2005)
18. Woodward, A., Penn, R.: *The wrong kind of snow*. Hodder & Stoughton, London (2007)
19. Zhang, L., Hermanns, H., Jansen, D.N.: Logic and model checking for hidden Markov models. In: Wang, F. (ed.) *Formal techniques for networked and distributed systems, FORTE 2005*. LNCS, vol. 3731, pp. 98–112. Springer, Berlin (2005)

A Proof of Theorem 1

Proof. We first prove *soundness* by simultaneous induction over the structure of state and path formulas.

For state formulas, we assume given two states s and t that are strongly bisimilar and a state formula Φ ; we prove that $s \models \Phi$ iff $t \models \Phi$. For symmetry reasons, it is enough to prove one direction of the equivalence, so assume that $s \models \Phi$; then it remains to be proven that $t \models \Phi$.

- $\Phi = \mathbf{true}$. Trivial.
- $\Phi = a$. A direct consequence of Condition 1 of Def. 3.
- $\Phi = \neg\Phi'$ and $\Phi = \Phi_1 \wedge \Phi_2$. These two cases are simple consequences of the induction hypothesis.
- $\Phi = \varepsilon$. The only interesting subcase is $\Phi = \mathcal{P}_{\bowtie p}(\varphi)$. Let Π be the set of paths that satisfy φ . (We sometimes denote $Prob_s(\Pi)$ as $Prob_s(\varphi)$.) We have to show that $Prob_s(\varphi) = Prob_t(\varphi)$. The induction hypothesis applied to φ implies that Π is bisimulation-closed. Therefore, by lemma 1, we can conclude that $Prob_s(\varphi) = Prob_s(\Pi) = Prob_t(\Pi) = Prob_t(\varphi)$.

Now let's look at path formulas. Assume given two bisimilar paths σ and ρ and a path formula φ . We prove that $\sigma \models \varphi$ iff $\rho \models \varphi$. (This also implies that the satisfaction set of a path formula is bisimulation-closed.) Again, it is enough to prove one direction; so assume that $\sigma \models \varphi$.

- $\varphi = \Phi$, $\varphi = \neg\varphi'$ and $\varphi = \varphi_1 \wedge \varphi_2$. These three cases are simple consequences of the induction hypothesis.
- $\varphi = X_{\mathcal{Y}} \varphi'$. The induction hypothesis for φ' and the path fragments $\sigma(1\dots)$ and $\rho(1\dots)$ (which are bisimilar to each other) gives us: $\sigma \models X_{\mathcal{Y}} \varphi'$ implies $\sigma(0)_o \in \mathcal{Y}$ and $\sigma(1\dots) \models \varphi'$, this implies $\rho(0)_o \in \mathcal{Y}$ and $\rho(1\dots) \models \varphi'$, and this again implies $\rho \models X_{\mathcal{Y}} \varphi'$.
- $\varphi = \varphi_1 \mathcal{U}^{\leq n} \varphi_2$. Let $j \leq n$ be such that $\sigma(j\dots) \models \varphi_2$. Then, by the induction hypothesis, $\rho(j\dots) \models \varphi_2$. Similarly, for every $i < j$, $\sigma(i\dots) \models \varphi_1$ implies $\rho(i\dots) \models \varphi_1$. Therefore, $\rho \models \varphi_1 \mathcal{U}^{\leq n} \varphi_2$.

This completes the soundness proof. To show *completeness*, we define the equivalence relation on states

$$R := \{(s, t) \mid \forall \text{ state formulas } \Phi : s \models \Phi \text{ iff } t \models \Phi\}$$

and show that it is a strong bisimulation relation. Assume given two states $s R t$.

- Condition 1: $s \models a$ iff $t \models a$. Trivial.
- Condition 2. Let ω be any observation. Let $p := Prob_s(\omega)$. Obviously, $s \models \mathcal{P}_{=p}(X_{\omega} \mathbf{true})$. Therefore, also $t \models \mathcal{P}_{=p}(X_{\omega} \mathbf{true})$, and from this follows that $p = Prob_t(\omega)$.
- Condition 3. Let ω be any observation, and let $C \in S/R$ be any equivalence class. We have to show that $\tau(s, \omega)(C) = \tau(t, \omega)(C)$. As we already know that $Prob_s(\omega) = Prob_t(\omega)$, it is enough to show that $p := Prob_s(\omega, C)$ is equal to $Prob_t(\omega, C)$.

It is possible to find a formula Φ_C that is satisfied exactly by the states in C . (If $s_1 \in C$ and $s_2 \in S \setminus C$, then there exists a formula $\Phi_{s_1 s_2}$ such that $s_1 \models \Phi_{s_1 s_2}$ and $s_2 \not\models \Phi_{s_1 s_2}$. As S is finite, a finite conjunction of such formulas can serve as Φ_C .) Now $s \models \mathcal{P}_{=p}(X_\omega \Phi_C)$, so $t \models \mathcal{P}_{=p}(X_\omega \Phi_C)$, so $p = \text{Prob}_t(\omega, C)$.

It now remains to be proven that paths satisfying the same path formulas are also bisimilar. We define the following equivalence relation on paths:

$$R := \{(\sigma, \rho) \mid \forall \text{ path formulas } \varphi : \sigma \models \varphi \text{ iff } \rho \models \varphi\}$$

and show that it is a strong bisimulation relation, i. e., if two paths are in R , then they are pointwise bisimilar. We prove this by reductio ad absurdum: Assume given two paths σ and ρ and an index $n \in \mathbb{N}$ such that $\sigma(n) \not\sim \rho(n)$.

Then, based on the proof above for state formulas, there exists a state formula Φ such that $\sigma(n) \models \Phi$ and $\rho(n) \not\models \Phi$. As a consequence, $\sigma \models X X \cdots X \Phi$ (the X operator is repeated n times), while ρ does not. Contradiction!