

Approximating Bisimilarity for Markov Processes¹

Chunlai Zhou²

Computer Science Department
School of Information
Renmin University of China
Beijing, CHINA

Abstract

In this paper we investigate bisimilarity for *general* Markov processes through the correspondence between sub- σ -algebras and equivalence relations. In particular, we study bisimulations from the perspective of fixed-point theory. Given a Markov process $M = \langle \Omega, \Sigma, \tau \rangle$, we characterize its state bisimilarity as the *greatest* fixed point of a composition of two natural set operators between equivalence relations on Ω and sub- σ -algebras of Σ . Moreover, we employ a Smith-Volterra-Cantor-set-construction to obtain an example to show that state bisimilarity is beyond ω iterations of these two operators alternately from event bisimilarity and hence the composite operator is not continuous. This process of iteration illustrates the gap between event bisimilarity (or logical equivalence) and state bisimilarity, and hence provides insights about the Hennessy-Milner property for general Markov processes. At the end of this paper, we also study approximation of Markov processes related to filtration.

Keywords: Markov processes, state bisimilarity, event bisimilarity, fixed point, Hennessy-Milner logic

1 Introduction

Markov processes with continuous state spaces are important mathematical models in different physical sciences such as physics, biology, finance and computer sciences. The dynamics of the processes is governed by the present state rather than by the past history of the processes. With the ever-growing computer technology, we need to develop a theory of computational grip of this kind of important structures. If one is interested in computing them, we must build a machinery to approximate Markov processes with continuous state space and also make sure that the approximating

¹ I want to thank Professor Ernst-Erich Doberkat, Ingo Battenfeld, Yuan Feng and Pedro Sánchez Terraf for helpful discussions at different stages of this research. This work is partly supported by NSF of China (Grant Number: 60905036), by the Key Project for basic research from the Ministry of Science and Technology of China (Grant Number: 2012CB316205), by the Basic Research Funds in Renmin University from the Central Government (Grant No. 201230005) and by German Academic Exchange Service (DAAD Codenumber: A/10/00803).

² Email: chunlai.zhou@gmail.com

processes *preserve* all the essential properties especially the dynamic aspects of the original processes.

The limit of the approximating processes is usually not the original approximated process but instead the *quotient* process with respect to some *bisimilarity*. There are two totally different notions of bisimilarities for Markov processes in the literature. The first one is called *state bisimilarity*. Intuitively, two states are state bisimilar in the process if they match transition probabilities for the same moves. And the other one is called *event bisimilarity*. Two states are event bisimilar if they are indistinguishable by any sub- σ -algebra of events that respects the dynamics in the process. For any *general* Markov process, event bisimilarity coincides with logical equivalence and is a superset of state bisimilarity [4]. For Markov processes on analytical spaces or Polish spaces [5][7], these three kinds of equivalences are the same, which is the well-known Hennessy-Milner property. However, a general Markov process does not necessarily satisfy the Hennessy-Milner property [12].

Conceptually, there is a *mismatch* between approximating Markov processes and bisimilarities in the literature. Most approaches to approximate Markov processes [6][15][3][13] employ similar *syntactic* machineries. The limit of the approximating Markov processes is the quotient Markov process with respect to event bisimilarity (or logical equivalence). However, it is the quotient Markov process with respect to *state bisimilarity* that preserves the dynamics of the original Markov process.

In order to understand better the approximation of Markov processes, we study in this paper approximating bisimilarity for *general* Markov processes. There are two approaches for approximating bisimilarities: bottom-up and top-down. The approximation according to the bottom-up approach is essentially syntactic and consists of a sequence of n -bisimilarities, which corresponds to logical equivalence up to depth n . So this approach is about event bisimilarity and is in spirit closely related to those of approximating Markov process in the literature. The second and top-down approach is semantical and studies state bisimilarity from the perspective of fixed-point theory. Given a Markov process $M = \langle \Omega, \Sigma, \tau \rangle$, we characterize its state bisimilarity as the *greatest* fixed point of a composition \mathbf{O} of two natural set operators between equivalence relations on Ω and sub- σ -algebras of Σ (Section 4). Not only may state bisimilarity be obtained from the universal relation on Ω by iterating α times the composite operator \mathbf{O} for some ordinal α , but also it can be reached top-down from event bisimilarity by iterating β times \mathbf{O} for some ordinal β . This top-down approach is actually reflected in many algorithms of computing bisimilarity in the literature [5] [6]. In this paper, we employ the above ordinal β to *measure* the gap from event bisimilarity to state bisimilarity. Sánchez Terraf [12] constructed an example and showed that the gap there is at least one. In this paper, we employ a Smith-Volterra-Cantor set (so-called *fat* Cantor set) to build an example and show that the gap is beyond the limit ordinal ω . This implies that the operator \mathbf{O} is not continuous and the gap between state and event bisimilarities is *very big*. Also the example illustrates the gap between the above two approaches for approximating bisimilarities: bottom-up and top-down.

At the end of the paper, we present a general theory about filtration as an ap-

proach of approximating Markov processes and discuss its relations to the above approaches to approximate bisimilarities. In particular we provide another characterization of the Hennessy-Milner property through *filtration*. Essentially, a filtration of a Markov process M' through a sublanguage \mathcal{L}' of the whole language \mathcal{L} for Markov processes is its quotient that respects the satisfiability of all formulas in \mathcal{L}' . We show (Theorem 5.4) that a Markov process satisfies the Hennessy-Milner property iff it has *only one* filtration through the language \mathcal{L} .

2 Preliminaries

Let \mathcal{A} be a (Boolean) algebra on a set X , i.e. a non-empty collection of subsets of X closed under complements and binary unions. \mathcal{A} is a σ -algebra if it is also closed under countable unions. If \mathcal{A} is a σ -algebra, then $\mathcal{X} = \langle X, \mathcal{A} \rangle$ is a *measurable space* and the elements of \mathcal{A} are usually called *events* or *measurable subsets of X* . We write $\sigma(\mathcal{A}_0)$ for the smallest σ -algebra containing a given set \mathcal{A}_0 of subsets of \mathcal{A} . When $\sigma(\mathcal{A}_0) = \mathcal{A}$, we usually say that \mathcal{A}_0 *generates* \mathcal{A} . A *measurable function* $f : \langle X, \mathcal{A} \rangle \rightarrow \langle X', \mathcal{A}' \rangle$ is a function $f : X \rightarrow X'$ such that, for any $A' \in \mathcal{A}'$, $f^{-1}(A') \in \mathcal{A}$ where $\langle X', \mathcal{A}' \rangle$ is also a measurable space. A set function $\mu : \mathcal{A} \rightarrow [0, \infty]$ on \mathcal{A} in X is *finitely additive* if $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$ whenever A_1 and A_2 are disjoint elements of \mathcal{A} . μ is called a (countably additive) *measure* if it satisfies the following conditions:

- (i) $\mu(\emptyset) = 0$;
- (ii) $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ where $\{A_i\}_{i=1}^{\infty}$ is a pairwise disjoint sequence of events of \mathcal{A} .

The second property is usually called the *countable additivity*. The measure μ is *finite* or *infinite* as $\mu(X) < \infty$ or $\mu(X) = \infty$. If $\mu(X) \leq 1$, then μ is called a *subprobability measure*. If $\mu(X) = 1$, then μ is called a *probability measure*. A metric space $\langle X, \rho \rangle$ is *complete* if any Cauchy sequence has a limit in X , and ρ is called a *complete metric*. A topological space $\langle X, \tau \rangle$ is called *separable* if it has a countable dense subset. A *Polish space* $\langle X, \tau \rangle$ is a separable topological space which is metrizable through a complete metric. The *Borel σ -algebra* $\mathcal{B}(X, \tau)$ for the topology τ is the smallest σ -algebra that contains τ . An *analytical space* is the image of a Polish space under a continuous function from one Polish space to another. The interested reader may refer to [1] for the basics about measure theory.

A *transition (sub)probability function* T on a measurable space $\mathcal{X} = \langle X, \mathcal{A} \rangle$ is a function from $X \times \mathcal{A}$ to $[0, 1]$ satisfying the following two conditions:

- for each $x \in X$, $T(x, \cdot)$ is a (sub)probability measure, and
- for each $A \in \mathcal{A}$, $T(\cdot, A)$ is a measurable function.

T is also called a *Markov kernel*. A *Markov process* M is a structure $\langle X, \mathcal{A}, T \rangle$, where $\langle X, \mathcal{A} \rangle$ is measurable space and T is a subprobability transition function. A function $f : \langle X, \mathcal{A}, T \rangle \rightarrow \langle X', \mathcal{A}', T' \rangle$ is a *zigzag morphism* if it is surjective, measurable, and the following equality holds:

$$T(x, f^{-1}(A')) = T'(f(x), A'), \text{ for any } x \in X, A' \in \mathcal{A}'.$$

The two Markov processes $\langle X, \mathcal{A}, T \rangle$ and $\langle X', \mathcal{A}', T' \rangle$ are probabilistically *bisimilar* if there is a Markov process $\langle X'', \mathcal{A}'', T'' \rangle$ with two *surjective* zigzag morphisms $h' : X'' \rightarrow X'$, and $h : X'' \rightarrow X$.

One important result about Markov processes is that there is a Hennessy-Milner logic to characterize the above probabilistic bisimulation. A *formula* ϕ of the logic is formed by the following syntax:

$$\phi := \top \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid L_r\phi(r \in \mathbb{Q} \cap [0, 1])$$

where \mathbb{Q} is the field of rationals. \mathcal{L} denotes the language of this simple syntax. The *depth* $dp(\phi)$ of formulas ϕ is defined inductively as in modal logic. \mathcal{L}_n denotes the sublanguage of \mathcal{L} of formulas of depth $\leq n$ ($n = 0, 1, 2, \dots$). The interpretation of formulas in the Markov process $M = \langle S, \mathcal{A}, T \rangle$ is straightforward except the following crucial clause:

$$M, w \models L_r\phi \text{ iff } T(w)(\llbracket \phi \rrbracket_M) \geq r, \text{ where } \llbracket \phi \rrbracket_M := \{w \in S : M, w \models \phi\}.$$

A formula ϕ is called satisfied at w in M if $M, w \models \phi$. Two states in S are called *logical equivalent* if they satisfy the same set of formulas in \mathcal{L} . The following is the well-known theorem about the Hennessy-Milner property or expressivity of Markov processes [5][7][8].

Theorem 2.1 *Let $\langle S, \mathcal{A}, T \rangle$ be a Markov process in which S is a Polish space and \mathcal{A} is a Borel σ -algebra. Two states are probabilistically bisimilar iff they satisfy the same set of formulas of \mathcal{L} .*

Before moving to the main part, we first fix some notations. Let $\langle \Omega, \Sigma \rangle$ be a measurable space and R be an equivalence relation on Ω . For $E \subseteq \Omega$, E/R denotes the set $\{[s]_R : s \in E\}$ and, for $\Sigma' \subseteq \Sigma$, $\Sigma'/R = \{E'/R : E' \in \Sigma'\}$. It is easy to see that $\langle \Omega/R, \Sigma/R \rangle$ is also a measurable space. Conversely, for $B \subseteq \Omega/R$, B_\cup denotes $\bigcup B$ and, for $\Sigma'' \subseteq \Sigma/R$, Σ''_\cup denotes the set $\{B_\cup : B \in \Sigma''\}$. In particular, $2_\cup^{\Omega/R}$ denotes the set $\{A \in 2^\Omega : A = \bigcup C \text{ for some } C \in 2^{\Omega/R}\}$ ³. For the equivalence relation R on Ω , elements of Σ are called *R-closed* if they are also unions of R -equivalence classes. $\Sigma(R)$ denotes the sub- σ -algebra of R -closed events in Σ , i.e., $\Sigma \cap 2_\cup^{\Omega/R}$. Let $\Sigma(\cdot)$ denote this mapping from equivalence relations on Ω to sub- σ -algebras of Σ . Conversely, for any sub- σ -algebra Σ' of Σ , $\mathbf{R}(\Sigma')$ denotes the equivalence relation:

$$s\mathbf{R}(\Sigma')s' \text{ if, for any } A' \in \Sigma' (s \in A' \Leftrightarrow s' \in A').$$

Let $\mathbf{R}(\cdot)$ denote this mapping from sub- σ -algebras of Σ to equivalence relations on Ω . It is easy to check that, given the space (Ω, Σ) , these two maps $\Sigma(\cdot)$ and $\mathbf{R}(\cdot)$ form a Galois connection [4]:

- (i) for any sub- σ -algebra \mathcal{B} of Σ , $\mathcal{B} \subseteq \Sigma(\mathbf{R}(\mathcal{B}))$;
- (ii) for any equivalence relation R' on S , $R' \subseteq \mathbf{R}(\Sigma(R'))$.

³ In topology, it is usually denoted as $(\Omega/R)^\cup$.

For a Markov process $M = \langle \Omega, \Sigma, \tau \rangle$ on $\langle \Omega, \Sigma \rangle$, we define a relation $\mathbf{R}^T(M)$ as follows: any s and t in Ω ,

$$(s, t) \in \mathbf{R}^T(M) \text{ whenever } \tau(s, E) = \tau(t, E) \text{ for all } E \in \Sigma.$$

Whenever τ is clear, we also write $\mathbf{R}^T(M)$ as $\mathbf{R}^T(\Sigma)$. Let $\mathbf{R}^T(\cdot)$ denote this mapping from sub- σ -algebras of Σ to equivalence relations on Ω . $\mathbf{R}^T(\cdot)$ and $\Sigma(\cdot)$ don't generally form a Galois connection (Example 3.3). Let Σ' be a sub- σ -algebra of Σ . We say that Σ' is *stable* with respect to $M = \langle \Omega, \Sigma, \tau \rangle$ if, for all $E \in \Sigma'$, $r \in [0, 1]$,

$$\{w \in \Omega : \tau(w, E) > r\} \in \Sigma'.$$

It is easy to see that Σ' is stable iff $\tau(\cdot, E)$ is Σ' -measurable for each $E \in \Sigma'$, i.e., $\langle \Omega, \Sigma', \tau \rangle$ is a Markov process [4].

Lemma 2.2 *Let R be an equivalence relation on Ω and Σ' be a sub- σ -algebra of Σ such that Σ' is stable.*

- (i) *$\langle \Omega, \Sigma(R), \tau \rangle$ is a Markov process if and only if $\tau(\cdot, E)$ is constant on R -classes for all $E \in \Sigma(R)$.*
- (ii) *$\mathbf{R}(\Sigma') \subseteq \mathbf{R}^T(\Sigma')$.*

Note that, generally, Part 2 does not hold if Σ' is not stable.

Lemma 2.3 *Let $\langle \Omega, \Sigma, \tau \rangle$ be a Markov process, Σ_1 and Σ_2 be two sub- σ -algebras of Σ , and R_1 and R_2 be two equivalence relations on Ω .*

- (i) *If $\Sigma_2 \subseteq \Sigma_1$, then $\mathbf{R}^T(\Sigma_2) \supseteq \mathbf{R}^T(\Sigma_1)$ and $\mathbf{R}(\Sigma_2) \supseteq \mathbf{R}(\Sigma_1)$.*
- (ii) *If $R_1 \subseteq R_2$, then $\Sigma(R_1) \supseteq \Sigma(R_2)$.*
- (iii) *If $\Sigma_2 \subseteq \Sigma_1$, then $\Sigma(\mathbf{R}^T(\Sigma_2)) \subseteq \Sigma(\mathbf{R}^T(\Sigma_1))$.*
- (iv) *If $R_1 \subseteq R_2$, then $\mathbf{R}^T(\Sigma(R_1)) \subseteq \mathbf{R}^T(\Sigma(R_2))$.*

3 Fixed-point characterization of state bisimilarity

In the following sections, we consider a given Markov process $M = \langle S, \mathcal{A}, \tau \rangle$ and study relationships between sub- σ -algebras of \mathcal{A} and equivalence relations on S .

Definition 3.1 An equivalence relation R on Markov process $M := \langle S, \mathcal{A}, \tau \rangle$ is called a *state bisimulation* if $R \subseteq \mathbf{R}^T(\mathcal{A}(R))$, namely,

$$\text{for any } s, t \in S, sRt \text{ implies that } \tau(s, E) = \tau(t, E) \text{ for every } E \in \mathcal{A}(R).$$

In other words, R is a state bisimulation if it is a *post-fixpoint* of the composite operator $\mathbf{R}^T(\mathcal{A}(\cdot))$. From Part (1) of Lemma 2.2, we know that R is a state bisimulation iff $\langle S, \mathcal{A}(R), \tau \rangle$ is a Markov process. Two states s and t in S are *state bisimilar* if there is a state bisimulation R such that $(s, t) \in R$. An equivalence relation R' on M is called an *event bisimulation* if it is defined through a Markov process with a sub- σ -algebra \mathcal{A}' in the sense that

- $R' = \mathbf{R}(\mathcal{A}')$, i.e., for any $s, t \in S$, $sR't$ iff s and t are indistinguishable in \mathcal{A}' ;
- $\langle S, \mathcal{A}', \tau \rangle$ is a Markov process.

Two states s and t are *event bisimilar* if there is an event bisimulation R' such that $(s, t) \in R'$.

The classes of both state and event bisimulations are closed under the following operation: for arbitrary index set I ,

- $\bigvee_{i \in I} R_i := (\bigcup_{i \in I} R_i)^*$ where $(\bigcup_{i \in I} R_i)^*$ denotes the transitive closure of the relation $\bigcup_{i \in I} R_i$.

Thus state bisimilarity is the union of all state bisimulations, and event bisimilarity that of state bisimulations. \approx_M and \sim_M denote state and event bisimilarities on M , respectively. When the context is clear, we usually drop the subscript M .

Originally, Danos et. al. [4] would like to present event bisimulation as a *weakening* of state bisimulation. However, from the following example (adapted from Example 4.11 in [4]), we know that a state bisimulation R is not in general an event bisimulation although a closely-related *bigger* state bisimulation $\mathbf{R}(\mathcal{A}(R))$ is indeed an event bisimulation (part 4 of the following proposition, which is from [4]).

Proposition 3.2 *Let R be a state bisimulation.*

- (i) $R \subseteq \mathbf{R}(\mathcal{A}(R))$;
- (ii) *If Λ is a sub- σ -algebra of \mathcal{A} , $\mathbf{R}(\Lambda) = \mathbf{R}(\mathcal{A}(\mathbf{R}(\Lambda)))$ and $\Lambda \subseteq \mathcal{A}(\mathbf{R}(\Lambda))$;*
- (iii) R is an event bisimulation *iff* $R = \mathbf{R}(\mathcal{A}(R))$;
- (iv) $\mathbf{R}(\mathcal{A}(R))$ is both a state bisimulation and an event bisimulation.

Example 3.3 Let $S = [0, 1]$ and \mathcal{B} be the σ -algebra of Borel sets on S . A non-Lebesgue-measurable subset N of $[0, 1]$ and its complement N^c as equivalence classes define an equivalence relation R on S . Note that $\mathcal{B}(R) = \{\emptyset, S\}$. Now we define a Markov kernel τ on $\langle S, \mathcal{B} \rangle$ such that R is a state bisimulation but not an event bisimulation. Let λ be the usual Lebesgue measure. Assume that s_1 and s_2 are two points in N and t in N^c . Define τ' as

$$\tau'(s, E) = \lambda(E) \text{ for any } s \in S \text{ and } E \in \mathcal{B}.$$

It is easy to see that τ' is a Markov kernel on $\langle S, \mathcal{B} \rangle$. Now we obtain τ from τ' by modifying the measures only at s_1 and s_2 as follows:

$$\tau(s_1, \{s\}) = 1 = \tau(s_2, \{t\}).$$

Such τ is also a Markov kernel on $\langle S, \mathcal{B} \rangle$. So $\langle S, \mathcal{B}(R), \tau \rangle$ is a Markov process and hence R is a state bisimulation. Since $R \subsetneq \{(s, s') : s, s' \in S\} = \mathbf{R}(\mathcal{B}(R))$, R is not an event bisimulation according to Proposition 3.2. Moreover, $\mathcal{B}(\mathbf{R}^T(\mathcal{B})) \subsetneq \mathcal{B}$ and hence $\mathbf{R}^T(\cdot)$ and $\mathcal{B}(\cdot)$ don't form a Galois connection.

In the remainder of this section, we will investigate state (event) bisimulation from the perspective of *fixed-point* theory. From the above Proposition 3.2, we know that, if an equivalence relation R on S is both a state bisimulation and an event bisimulation, it is a fixed point of the operator $\mathbf{R}(\mathcal{A}(\cdot))$ on the class of equivalence relations on S .

Theorem 3.4 *Both state bisimilarity and event bisimilarity are fixed points of the composite operator $\mathbf{R}(\mathcal{A}(\cdot))$. So state bisimilarity \approx is also an event bisimulation and hence $\approx \subseteq \sim$.*

Proof. From Proposition 3.2, we know that $\approx \subseteq \mathbf{R}(\mathcal{A}(\approx))$ and $\mathbf{R}(\mathcal{A}(\approx))$ is a state bisimulation. Since \approx is the union of all state bisimulations, $\approx = \mathbf{R}(\mathcal{A}(\approx))$. The proof for event bisimilarity is similar. \square

However, event bisimilarity \sim is not the *greatest* fixed point of the operator, since the universal relation $S \times S$ is also a fixed point. In the next section, we will show that generally the above containment in Theorem 3.4 is strict.

Theorem 3.5 *The state bisimilarity \approx is the greatest fixed point of the composite operator $\mathbf{R}^T(\mathcal{A}(\cdot))$.*

Proof. We know from Theorem 3.4 that, for state bisimilarity \approx , $\approx = \mathbf{R}(\mathcal{A}(\approx))$. Since $\langle S, \mathcal{A}(\approx), \tau \rangle$ is a Markov process, $\approx = \mathbf{R}(\mathcal{A}(\approx)) \subseteq \mathbf{R}^T(\mathcal{A}(\approx))$ (according to Lemma 3.2). Let R' denote $\mathbf{R}^T(\mathcal{A}(\approx))$. It follows that $\mathcal{A}(R') \subseteq \mathcal{A}(\approx)$ and hence $\tau(\cdot, E)$ is constant on R' -classes for all $E \in \mathcal{A}(\approx)$ and hence is constant on R' -classes for all $E \in \mathcal{A}(R')$. It follows from Lemma 2.2 that R' is a state bisimulation. Since we have shown $\approx \subseteq R'$ and \approx is the greatest state bisimulation, the state bisimilarity \approx is the same as R' . In other words, \approx is also the fixed point of the operator $\mathbf{R}^T(\mathcal{A}(\cdot))$. It is also the greatest fixed point. Indeed, each fixed point R of $\mathbf{R}^T(\mathcal{A}(\cdot))$ is also a state bisimulation and hence is contained in the state bisimilarity \approx . \square

One may also appeal *directly* to the well-known Tarski-Knaster Theorem (Chapter 1 of [11]) to show that state bisimilarity is the greatest fixed point of the composite operator $\mathbf{R}^T(\mathcal{A}(\cdot))$. Desharnais et.al. [6] also studied state bisimilarity from the perspective of fixed point but did not consider its relationship with other bisimilarities. The main purpose of our above presentation of state bisimilarity by detouring to transition bisimilarity is to characterize both the relationships among different bisimulations and the gaps among them through the operator $\mathbf{R}^T(\mathcal{A}(\cdot))$.

4 Gap between state and event bisimilarities

For simplicity, we use \mathbf{O} to denote the composite operator $\mathbf{R}^T(\mathcal{A}(\cdot))$. For a relation R , we construct by transfinite induction a chain of equivalence relations on $M = \langle S, \mathcal{A}, \tau \rangle$ as follows:

- $\mathbf{O}^{\alpha+1}(R) = \mathbf{O}(\mathbf{O}^\alpha(R))$;
- $\mathbf{O}^\lambda(R) = \bigcup_{\alpha < \lambda} \mathbf{O}^\alpha(R)$ if λ is a limit ordinal.

According to Lemma 2.3, \mathbf{O} is monotonic.

Theorem 4.1 *For the above operator \mathbf{O} ,*

- (i) *The greatest fixed point exists and is $\mathbf{O}^\alpha(R_u)$ for some ordinal α . So state bisimilarity can be obtained by iterating the operator \mathbf{O} α times from the uni-*

versal relation R_u for some ordinal α whose cardinality is no larger than that of S .

- (ii) state bisimilarity \approx can be obtained from event bisimilarity \sim by iterating α times \mathbf{O} for some ordinal α ; in other words, $\approx = \mathbf{O}^\alpha(\sim)$.

Proof. The first part follows trivially from Tarski-Knaster’s fixed point Theorem and the second from Theorem 3.5 and Lemma 2.3. \square

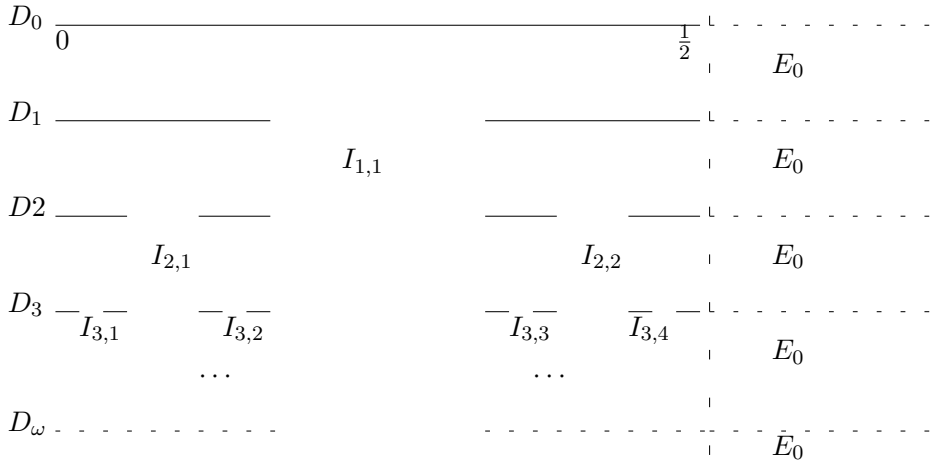
The above theorem tells us that the ordinal α in the equation $\approx = \mathbf{O}^\alpha(\sim)$ may be employed to “measure” the gap between state bisimilarity and event bisimilarity. In the following, we employ a *Smith-Volterra-Cantor set* (or simply SVC set) to construct an example to show that state bisimilarity can not be obtained by iterating ω times the operator \mathbf{O} from event bisimilarity. This example illustrates the gaps between these two bisimilarities and further between the two approaches for approximating bisimilarity: top-down and bottom-up. Also this example shows that \mathbf{O} is not downward continuous (Corollary 4.7). But, if $\langle S, \mathcal{A} \rangle$ is analytical or discrete, then state bisimilarity and event bisimilarity coincide and the operator \mathbf{O} is continuous ([4] and [9]).

Example 4.2 (SVC-set-construction) We define a sequence of partitions $\Pi_i (i \geq 0)$ and corresponding equivalence relations $R_i (i \geq 0)$ of $S = [0, 1]$ inductively as follows. For an interval I of S , let $\mathcal{B}(I)$ denote the σ -algebra of Borel sets in I , $\mathcal{C}(I)$ the countable subclass that generates the σ -algebra and $\mathcal{M}(I)$ the Lebesgue completion. There is a non-Lebesgue-measurable subset E_0 of $[\frac{1}{2}, 1]$. Let \mathcal{B}_0 denote $\sigma(\mathcal{C}[\frac{1}{2}, 1]) \cup \{[0, \frac{1}{2}], E_0\}$. The construction will proceed in steps. At the first step, let $I_{1,1}$ denote the open interval $(\frac{1}{2} \cdot \frac{3}{8}, \frac{1}{2} \cdot \frac{5}{8})$. Thus $I_{1,1}$ is the open middle of the interval $I_0 := [0, \frac{1}{2}]$ of length $\frac{1}{2} \cdot \frac{1}{2^{2 \cdot 1}}$. The second step involves performing the first step on each of the two remaining closed intervals of $I_0 \setminus I_{1,1}$. That is, we produce two open intervals $I_{2,1}$ and $I_{2,2}$, each being the open middle with length $\frac{1}{2} \cdot \frac{1}{2^{2 \cdot 2}}$ of one of the two intervals comprising $I_0 \setminus I_{1,1}$. At the i -th step we produce 2^{i-1} open intervals, $I_{i,1}, I_{i,2}, \dots, I_{i,2^{i-1}}$, each of length $\frac{1}{2} \cdot \frac{1}{2^{2^i}}$. The $(i+1)$ -th step consists of producing open middles of length $\frac{1}{2} \cdot \frac{1}{2^{2^{i+1}}}$ of each of the intervals of

$$I_0 \setminus \bigcup_{j=1}^i \bigcup_{k=1}^{2^{j-1}} I_{j,k}.$$

D_0 denotes $[0, \frac{1}{2}]$. For any natural number i , let D_i denote the set $I_0 \setminus \bigcup_{j=1}^i \bigcup_{k=1}^{2^{j-1}} I_{j,k}$. With D_ω denoting the SVC set with respect to $[0, \frac{1}{2}]$, we define its complement by

$$I_0 \setminus D_\omega = \bigcup_{j=1}^\infty \bigcup_{k=1}^{2^{j-1}} I_{j,k}$$



There are some facts about this SVC set D_ω that we need for the following construction.

- (i) D_ω is a closed set and has a positive measure $\frac{1}{4}$. In fact, the total sum of the lengths of the deleted open intervals is $\sum_{i=1}^{\infty} \frac{1}{2} \frac{1}{2^{i+1}} = \frac{1}{4}$.
- (ii) D_ω contains a non-Lebesgue-measurable subset $D_{\omega+1}$, since D_ω has a positive measure.
- (iii) D_ω is totally disconnected, i.e., all connected components are singletons. That is to say, each connected component in D_ω is a singleton.

Set $R_0 := \{(x, x) : x \in (\frac{1}{2}, 1]\} \cup \{(x, y) : x, y \in [0, \frac{1}{2}]\}$. Next we define another equivalence relation R_1 on $[0, 1]$ which refines R_0 by simulating the trisection process in the construction of the SVC set D_ω .

$$\begin{aligned}
 R_1 := & \{(x, x) : x \in (\frac{1}{2}, 1]\} \\
 & \cup \{(x, y) : x, y \in I_{1,1}\} \\
 & \cup \{(x, y) : x, y \in I_0 \setminus I_{1,1}\}
 \end{aligned}$$

More generally, we define, for $i \geq 1$,

$$\begin{aligned}
 R_i := & \{(x, x) : x \in (\frac{1}{2}, 1]\} \\
 & \cup \bigcup_{j=1}^i \{(x, y) : x, y \in \bigcup_{k=1}^{2^{j-1}} I_{j,k}\} \\
 & \cup \{(x, y) : x, y \in I_0 \setminus \bigcup_{j=1}^i \bigcup_{k=1}^{2^{j-1}} I_{j,k}\}
 \end{aligned}$$

R_ω denotes the intersection of all R_i 's, i.e., $R_\omega = \bigcap_i R_i$. Actually R_ω can be expressed as follows:

$$\begin{aligned}
R_\omega &= \{(x, x) : x \in (\frac{1}{2}, 1]\} \\
&\cup \bigcup_{j=1}^{\infty} \{(x, y) : x, y \in \bigcup_{k=1}^{2^{j-1}} I_{j,k}\} \\
&\cup \{(x, y) : x, y \in I_0 \setminus \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{2^{j-1}} I_{j,k}\}
\end{aligned}$$

Note that $I_0 \setminus \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{2^{j-1}} I_{j,k}$ is precisely the SVC set D_ω with respect to $[0, \frac{1}{2}]$. Define $\mathcal{B}_n := \sigma(\mathcal{C}[\frac{1}{2}, 1] \cup \{[0, \frac{1}{2}], E_0\} \cup \{D_j : j \preceq n\})$ for $n \preceq \omega + 1$ where \preceq is the ordinal relation.

Let $\mathcal{C}_{\omega+1}$ denote $\mathcal{C}[\frac{1}{2}, 1] \cup \{[0, \frac{1}{2}], E_0\} \cup \{D_i : i = 1, 2, \dots\} \cup \{D_{\omega+1}\}$, \mathcal{B}_{-i} the σ -algebra $\sigma(\mathcal{C}_{\omega+1} \setminus \{D_i\})$ ($i = 1, 2, \dots$) and $\mathcal{B}_{-(\omega+1)}$ the σ -algebra $\sigma(\mathcal{C}_{\omega+1} \setminus \{D_{\omega+1}\})$. Note that all the events in $\mathcal{C}_{\omega+1}$ are Lebesgue-measurable except $D_{\omega+1}$ and E_0 . So $\mathcal{B}_{\omega+1} = \sigma(\mathcal{C}_{\omega+1})$ and is countably generated. The following Extension Theorem is the most important “weapon” that we will use to construct our Markov kernel τ .

Proposition 4.3 (Theorem 1.12.14 in [2]) Assume that

- (i) μ is a finite nonnegative measure on the measurable space $\langle \Omega, \Sigma \rangle$; and
- (ii) A is a subset of Ω such that $\mu_*(A) < \mu^*(A)$ where μ_* and μ^* are the inner and outer measures of μ , respectively.

Then, for any r such that $\mu_*(A) \leq r \leq \mu^*(A)$, there is a countably additive measure μ' on the σ -algebra $\sigma(\Sigma \cup \{A\})$ such that $\mu'(A) = r$ and $\mu' = \mu$ on Σ .

By appealing to the above theorem, we obtain a measure λ_ω on \mathcal{B}_ω such that λ_ω is an extension of the Lebesgue measure on the sub- σ -algebra generated by $\mathcal{C} \setminus \{E_0, D_{\omega+1}\}$. It is easy to see that, since $D_{\omega+1}$ is a non-Lebesgue-measurable subset of $[0, \frac{1}{2}]$, $(\lambda_\omega)_*(D_{\omega+1}) < (\lambda_\omega)^*(D_{\omega+1})$. According to the above Extension Theorem, for any r such that $(\lambda_\omega)_*(D_{\omega+1}) \leq r \leq (\lambda_\omega)^*(D_{\omega+1})$, there is a countably additive extension $\lambda_{\omega+1}^r$ such that $\lambda_{\omega+1}^r = \lambda_\omega$ on \mathcal{B}_ω and $\lambda_{\omega+1}^r(D_{\omega+1}) = r$. Let $I_{\omega+1} = \{r : (\lambda_\omega)_*(D_{\omega+1}) \leq r \leq (\lambda_\omega)^*(D_{\omega+1})\}$. There is an injective and increasing f from C to the set $I_{\omega+1}$. For each $x \in C$, if $f(x) = r$, then we also use $\lambda_{\omega+1}^{f(x)}$ to denote $\lambda_{\omega+1}^r$. Especially, we simply use $\lambda_{\omega+1}$ to denote the “last” such extension $\lambda_{\omega+1}^{f(\frac{1}{2})}$. Note that $\lambda_{\omega+1}^{f(x)}$ is a measure on $\mathcal{B}_{\omega+1}$ for all $x \in C$.

It is easy to check that, for each D_i ($i = 1, 2, \dots$),

$$(\lambda_{\omega+1} \upharpoonright_{\mathcal{B}_{-i}})_*(D_i) = \lambda(D_{i+1}) < \lambda(D_{i-1}) = (\lambda_{\omega+1} \upharpoonright_{\mathcal{B}_{-i}})^*(D_i)$$

Similarly, according to Theorem 4.3, there is a measure λ_{-i} on $\mathcal{B}_{\omega+1}$ such that $\lambda_{-i}(D_i) \neq \lambda_{\omega+1}(D_i)$ and $\lambda_{-i} = \lambda_{\omega+1}$ on \mathcal{B}_{-i} .

Now we define a Markov kernel on the measurable space $\langle S, \mathcal{B}_{\omega+1} \rangle$.

$$\tau(x, E) = \begin{cases} x \cdot \lambda_{\omega+1}(E) & \text{if } x \in [\frac{1}{2}, 1], \\ \frac{1}{2} \cdot \lambda_{\omega+1}^{f(x)}(E) & \text{if } x \in C \setminus \{\frac{1}{2}\}, \\ \frac{1}{2} \cdot \lambda_{-i}(E) & \text{if } x \in \bigcup_{k=1}^{2^{i-1}} I_{i,k} \end{cases}$$

Lemma 4.4 *The above defined $M := \langle S, \mathcal{B}_{\omega+1}, \tau \rangle$ is a Markov process.*

Proof. The crucial part is to show that $\tau(\cdot, D_{\omega+1})$ is $\mathcal{B}_{\omega+1}$ -measurable. This follows from the fact that f is injective and increasing. \square

Lemma 4.5 $\llbracket \mathcal{L} \rrbracket_M := \{ \llbracket \phi \rrbracket_M : \phi \in \mathcal{L} \} \subseteq \{ E : E = E_1 \cup [0, \frac{1}{2}] \text{ for some } E_1 \in \mathcal{B}[\frac{1}{2}, 1] \}$. And the logical equivalence or event bisimilarity \sim_M is

$$R_0 = \{ (x, x) : x \in (\frac{1}{2}, 1] \} \cup \{ (x, y) : x, y \in [0, \frac{1}{2}] \}$$

Theorem 4.6 (Main Theorem) *For simplicity, let \mathcal{A} denote the reference σ -algebra $\mathcal{B}_{\omega+1}$. For the above sequences of σ -algebra \mathcal{B}_i and of equivalence relations R_i , they satisfy the interrelations illustrated as follows:*

$$\begin{array}{ccccccc} \mathcal{B}_{\omega+1} & \supseteq & \sigma(\bigcup_i \mathcal{B}_i) & \cdots & \supseteq & \mathcal{B}_1 & \supseteq & \mathcal{B}_0 & \supseteq & \sigma(\llbracket \mathcal{L} \rrbracket_M) \\ & \searrow \mathcal{A}(\cdot) & \downarrow R^T(\cdot) & & \searrow \mathcal{A}(\cdot) & \downarrow R^T(\cdot) & \searrow \mathcal{A}(\cdot) & \downarrow R^T(\cdot) & \searrow \mathcal{A}(\cdot) & \downarrow R^T(\cdot) \\ \cdots & \subseteq & R_\omega & \cdots & \subseteq & R_2 & \subseteq & R_1 & \subseteq & R_0 (= \sim_M) \end{array}$$

Corollary 4.7 $\mathcal{A}(R_\omega) = \mathcal{B}_{\omega+1}$ and $R^T(\mathcal{B}_{\omega+1}) \subsetneq R_\omega$. So $\bigcap_i \mathbf{O}(R_i) \supsetneq \mathbf{O}(\bigcap_i R_i)$ and hence \mathbf{O} is not downward continuous.

Proof. The first part is straightforward. The second one follows from the fact

$$R^T(\mathcal{B}_{\omega+1}) = \{ (x, x) : x \in [\frac{1}{2}, 1] \cup C \} \cup \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{2^{j-1}} I_{j,k}$$

and hence $R^T(\mathcal{B}_{\omega+1}) \subsetneq R_\omega$. \square

5 Filtration and Hennessy-Milner property

In this section, we simulate Goldblatt's work in [10] to develop a general theory about the relationship among bisimilarity, filtration and Hennessy-Milner property by providing another characterization of the Hennessy-Milner property through filtration (Theorem 5.4). The following proposition from [4] tells us that event bisimilarity is characterized by the simple logic \mathcal{L} .

Theorem 5.1 *For the Markov process $M = \langle S, \mathcal{A}, \tau \rangle$,*

- (i) $\langle S, \sigma(\llbracket \mathcal{L} \rrbracket_M), \tau \rangle$ is a Markov process;
- (ii) $R(\mathcal{A}(\sim)) = \sim$;
- (iii) $\sigma(\llbracket \mathcal{L} \rrbracket_M)$ is the smallest stable sub- σ -algebra \mathcal{A}' that defines \sim , i.e., $R(\mathcal{A}') = \sim$.

From the above proposition, we know that $\mathcal{A}(\sim)$ is the biggest σ -algebra that defines \sim but is generally not the biggest stable σ -algebra that defines \sim because

otherwise $\sim = \approx$ (Theorem 5.4).

However, the biggest stable σ -algebra that defines \sim always exists. Let $\mathbf{F} = \{\mathcal{B} : \mathcal{A}(\sim) \supseteq \mathcal{B} \supseteq \sigma(\llbracket \mathcal{L} \rrbracket_M), \langle S, \mathcal{B}, \tau \rangle \text{ is a Markov process}\}$. \mathbf{F} is a complete lattice under the following lattice operations: for $(\mathcal{B}_i)_{i \in I} \subseteq \mathbf{F}$,

- $\bigwedge_i \mathcal{B}_i = \bigcap_i \mathcal{B}_i$;
- $\bigvee_i \mathcal{B}_i = \bigcap \{\mathcal{B} \in \mathbf{F} : \mathcal{B} \supseteq \mathcal{B}_i \text{ for all } i \in I\}$.

Let \mathcal{F} denote $\bigcup \mathbf{F}$. It follows immediately that $\langle S, \mathcal{F}, \tau \rangle$ is a Markov process and is the biggest stable sub- σ -algebra that is contained in $\mathcal{A}(\sim)$ and defines event bisimilarity \sim .

In the following, we give a general definition of filtration. Essentially, a filtration of a Markov process M' through a sublanguage \mathcal{L}' is its quotient that respects the satisfiability of all formulas in \mathcal{L}' . Let \mathcal{L}' be a subset of language of \mathcal{L} which is *closed under subformulas*. In other words,

- $\top \in \mathcal{L}'$;
- if $L_r \phi \in \mathcal{L}'$, $\phi \in \mathcal{L}'$;
- if $\phi \wedge \psi \in \mathcal{L}'$, $\phi \in \mathcal{L}'$ and $\psi \in \mathcal{L}'$.

\mathcal{L}' defines an equivalence relation $\sim_{\mathcal{L}'}$ on S : $s \sim_{\mathcal{L}'} t$ if they satisfy the same set of formulas in \mathcal{L}' . Any Markov processes $M_{\sim_{\mathcal{L}'}} = \langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}^{\mathcal{A}'} \rangle$ on the set $S/\sim_{\mathcal{L}'}$ of equivalence classes where $\mathcal{A}' \subseteq \mathcal{A}/\sim_{\mathcal{L}'}$ is called a *filtration of M through the sub-language \mathcal{L}'* if it satisfies the following property: for any $s \in S$ and $\phi \in \mathcal{L}'$,

$$M, s \models \phi \text{ if and only if } M_{\sim_{\mathcal{L}'}}([s]_{\mathcal{L}'}) \models \phi$$

When the context is clear, we simply call $M_{\sim_{\mathcal{L}'}}$ a filtration. For the measurable space $\langle S/\sim_{\mathcal{L}'}, \sigma(\llbracket \mathcal{L} \rrbracket_M)/\sim_{\mathcal{L}'} \rangle$, let $\tau_{\sim_{\mathcal{L}'}}^{\sigma(\llbracket \mathcal{L} \rrbracket_M)/\sim_{\mathcal{L}'}}([s]_{\mathcal{L}'}, E) := \tau(s', \bigcup E)$ for some $s' \in [s]_{\mathcal{L}'}$ and $E \in \sigma(\llbracket \mathcal{L} \rrbracket_M)/\sim_{\mathcal{L}'}$. From Proposition 2.2, we know $\sigma(\llbracket \mathcal{L} \rrbracket_M)/\sim_{\mathcal{L}'}$ is stable.

Theorem 5.2 $\langle S/\sim_{\mathcal{L}'}, \sigma(\llbracket \mathcal{L} \rrbracket_M)/\sim_{\mathcal{L}'}, \tau_{\sim_{\mathcal{L}'}}^{\sigma(\llbracket \mathcal{L} \rrbracket_M)/\sim_{\mathcal{L}'}} \rangle$ is a filtration.

It is clear that, for any filtration $M_{\sim_{\mathcal{L}'}} = \langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}} \rangle$, $\mathcal{A}' \supseteq \sigma(\llbracket \mathcal{L}' \rrbracket_{M_{\sim_{\mathcal{L}'}}})$ and $\langle S/\sim_{\mathcal{L}'}, \sigma(\llbracket \mathcal{L}' \rrbracket_{M_{\sim_{\mathcal{L}'}}}), \tau_{\sim_{\mathcal{L}'}}} \rangle$ is a filtration.

In the following, we employ the idea of averaging in [3] to show that, for any σ -algebra \mathcal{A}' such that $\mathcal{A}/\sim_{\mathcal{L}'} \supseteq \mathcal{A}' \supseteq \sigma(\llbracket \mathcal{L}' \rrbracket_{M_{\sim_{\mathcal{L}'}}})$, there is always a filtration $\langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}} \rangle$ with \mathcal{A}' as its σ -algebra of events. The main task is to find a Markov kernel $\tau_{\sim_{\mathcal{L}'}}$ such that $\langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}} \rangle$ is a Markov process.

In order to apply averaging here, we assume that there is a prior probability measure P on the measure space $\langle S, \mathcal{A} \rangle$. Note that any $\sim_{\mathcal{L}'}$ -equivalence class $[s]_{\mathcal{L}'}$ is \mathcal{A} -measurable. We define a mapping $\tau_{\sim_{\mathcal{L}'}}}^{\mathcal{A}'} : S_{\sim_{\mathcal{L}'}} \times \mathcal{A}' \rightarrow [0, 1]$ as follows: for any $[s]_{\mathcal{L}'} \in S_{\sim_{\mathcal{L}'}}$ and $A' \in \mathcal{A}'$,

$$\tau_{\sim_{\mathcal{L}'}}}^{\mathcal{A}'}([s]_{\mathcal{L}'}, A') := \frac{\int_{[s]_{\mathcal{L}'}} \tau(s, \bigcup A') dP(s)}{P([s]_{\mathcal{L}'})}.$$

From [3] and Proposition 2.2, we know that

Theorem 5.3 For such defined $\tau_{\sim_{\mathcal{L}'}}^{A'}$,

- (i) $\langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}^{A'} \rangle$ is a Markov process and hence is a filtration of M through the sub-language \mathcal{L}' .
- (ii) $\langle S, \mathcal{A}', \tau \rangle$ is a Markov process if and only if the natural mapping from M to $\langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}^{A'} \rangle$ is a zigzag morphism, i.e., for any $s \in S$ and $A' \in \mathcal{A}'$, $\tau_{\sim_{\mathcal{L}'}}^{A'}([s]_{\sim_{\mathcal{L}'}}), A') = \tau(s, \bigcup A')$.

For the language \mathcal{L}' , $\langle S/\sim_{\mathcal{L}'}, \sigma(\llbracket \mathcal{L}' \rrbracket_M)/\sim_{\mathcal{L}'}, \tau_{\sim_{\mathcal{L}'}}^{\sigma(\llbracket \mathcal{L}' \rrbracket_M)/\sim_{\mathcal{L}'}}} \rangle$ is called the *smallest filtration* and $\langle S/\sim_{\mathcal{L}'}, \mathcal{A}/\sim_{\mathcal{L}'}, \tau_{\sim_{\mathcal{L}'}}^{A/\sim_{\mathcal{L}'}}} \rangle$ the *greatest filtration* on $S/\sim_{\mathcal{L}'}$. Note that generally the natural mapping from M to $\langle S/\sim_{\mathcal{L}'}, \mathcal{A}', \tau_{\sim_{\mathcal{L}'}}^{A'} \rangle$ is not a zigzag morphism because $\tau_{\sim_{\mathcal{L}'}}^{A'}$ may be different from τ . That is to say, generally we don't have $\tau_{\sim_{\mathcal{L}'}}^{A'}([s]_{\sim_{\mathcal{L}'}}), A') = \tau(s, \bigcup A')$ for $s \in S$ and $A' \in \mathcal{A}'$. But, for $A' \in \sigma(\llbracket \mathcal{L}' \rrbracket_M)$, we always have that, for any $s \in S$, $\tau_{\sim_{\mathcal{L}'}}^{A'}([s]_{\sim_{\mathcal{L}'}}), A') = \tau(s, \bigcup A')$.

In [15], we provides a sequence of filtrations through a sequence of *finite languages* $(\mathcal{L}_i)_{i=1}^\infty$, which are closed under subformulas, to approximate the original Markov process M . This is a kind of approximation based on the so-called *bottom-up* approximating (event) bisimilarity.

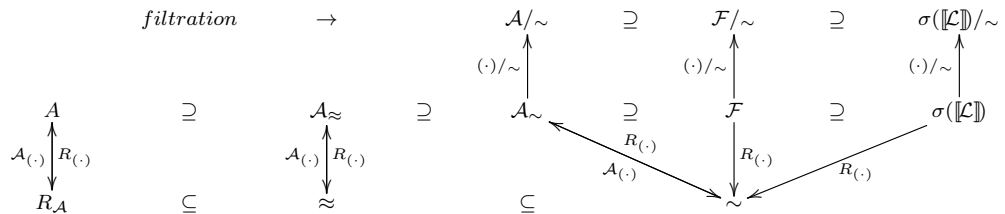
The following theorem is a generalization of Theorem 15 in [10] and provides another characterization of the Hennessy-Milner property through filtration.

Theorem 5.4 For the whole language \mathcal{L} ,

- (i) the natural mapping from M to $\langle S/\sim_{\mathcal{L}}, \mathcal{A}/\sim, \tau_{\sim}^{A(\sim)} \rangle$ is a zigzag morphism if and only if $\mathcal{A}(\approx) = \mathcal{A}(\sim)$ or equivalently $\approx = \sim$, i.e., M satisfies the Hennessy-Milner property.
- (ii) $\mathcal{A}(\sim) = \sigma(\llbracket \mathcal{L} \rrbracket_M)$ iff there is only one filtration through the language \mathcal{L} iff $\approx = \sim$.

Proof. For the second part, we note that there is only one filtration through the language \mathcal{L} iff, for each $A' \in \mathcal{A}(\sim)$, $\tau(\cdot, A')$ is constant on $[s]_{\sim_{\mathcal{L}}}$ for every $s \in S$. \square

The following is about the position of filtration in a general picture of interrelationships among different σ -algebras and equivalence relations.



6 Conclusion

In this paper, we study the difference between event and state bisimilarities from the perspective of fixed point theory. We quantify this difference by counting the

iteration times of the operator \mathbf{O} from event bisimilarity to state bisimilarity. Our work provides insights about the Hennessy-Milner property for general Markov processes. At the end of this paper, we provide another characterization of this property through filtration. Approximate bisimilarity [14] is another important notion to reason about approximate equivalence of processes. It is a subject for future work to study approximating bisimilarity for Markov processes from the perspective of approximate bisimilarity.

References

- [1] P. Billingsley. *Measure and Probability*. John Wiley & Sons, Inc., third edition, 1995. Wiley Series in Probability and Mathematical Statistics.
- [2] V.I. Bogachev. *Measure Theory*. Springer-Verlag Berlin Heidelberg, 2007.
- [3] P. Chaput, V. Danos, P. Panangaden, and G. D. Plotkin. Approximating markov processes by averaging. In *ICALP (2)*, pages 127–138, 2009.
- [4] V. Danos, Josee Desharnais, François Laviolette, and Prakash Panangaden. Bisimulation and cocongruence for probabilistic systems. *Inf. Comput.*, 204(4):503–523, 2006.
- [5] J. Desharnais, A. Edalat, and Panangaden P. Bisimulation for labeled markov processes. *Inf. and Comput.*, 179:163–193, 2002.
- [6] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labeled markov processes. *Inf. and Comp.*, 184:160–200, 2003.
- [7] E. E Doberkat. Stochastic relations: congruence, bisimulations and the hennessy-milner theorem. *SIAM J. Computing*, 35(3):590–626, 2006.
- [8] E. E Doberkat. *Stochastic Coalgebraic Logic*. EATCS Monographs in Theoretical Computer Sciences. Springer, 2010.
- [9] E. E Doberkat. Lattice properties of congruences for stochastic relations. *Ann. Pure Appl. Logic*, 163(8):1016–1029, 2012.
- [10] R. Goldblatt. Saturation and the hennessy-milner property. In M. de Rijke A. Ponse and Y. Venema, editors, *Modal Logic and Process Algebra*, pages 107–129. CSLI Publications, Stanford, CA, 1995.
- [11] D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic*. Foundation of Computing Series. MIT Press, Cambridge, 2000.
- [12] P. Terraf. Unprovability of the logical characterization of bisimulation. *Inf. Comput.*, 209(7):1048–1056, 2011.
- [13] F. van Breugel and Worrell J. Approximating and computing behavioural distances in probabilistic transition systems. *Theor. Comput. Sci.*, 360(1-3):373–385, 2006.
- [14] M. Ying and M. Wirsing. Approximate bisimilarity. In *AMAST*, pages 309–322, 2000.
- [15] C. Zhou and M. Ying. Approximation of markov processes through filtration. *Theoretical Comput. Sci.*, 446:75–97, 2012.