Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in Theoretical Computer Science

Electronic Notes in Theoretical Computer Science 162 (2006) 73–78

www.elsevier.com/locate/entcs

Bisimulation and Simulation Relations for Markov Chains

Christel Baier

Universität Bonn, D-53117 Bonn, Germany

Holger Hermanns

Universität des Saarlandes, D-66123 Saarbrücken, Germany

Joost-Pieter Katoen

RWTH Aachen, D-52074 Aachen, Germany

Verena Wolf

University of Mannheim, D-68131 Mannheim, Germany

Abstract

Formal notions of bisimulation and simulation relation play a central role for any kind of process algebra. This short paper sketches the main concepts for bisimulation and simulation relations for probabilistic systems, modelled by discrete- or continuous-time Markov chains.

Keywords: Markov chain, (weak) simulation, (weak) bisimulation, temporal logic

1 Introduction

To compare the stepwise behaviour of states in labeled transition systems, simulation and bisimulation relations have been widely considered. They play a crucial role for the compositional design and reasoning within a process algebra framework, and for abstraction purposes. Bisimulation relations are equivalences requiring two bisimilar states to exhibit identical stepwise behaviour. Simulation relations are uni-directed requiring that whenever s' simulates s then state s' can mimic all stepwise behaviour of s; but possibly not vice versa. Typically, (bi)simulation relations enjoy many nice properties such as congruence properties for parallel composition and other operators of process algebras, preservation properties for linear

1571-0661
doi:10.1016/j.entcs.2005.12.078

and branching-time logics; they have sound and complete axiomatizations, efficient decision algorithms and allow for coinductive reasoning.

In this short paper, we consider probabilistic systems modelled by action-labelled Markov chains and summarize the main concepts of (bi)simulation relations for them. Markov chains are an important class of stochastic processes that are widely used in practice to determine system performance and dependability characteristics, see e.g. [28,20]. A variety of process algebras with an operational Markov chain semantics has been defined, see e.g. [27,17,9,23,22,30] for an overview. Based on the seminal works of Jonsson and Larsen [26] and Larsen and Skou [29], various notions of simulation and bisimulation relations have been studied for both discrete and continuous-time Markov chains. This paper surveys the results on comparative semantics of branching-time relations for time-abstract fully probabilistic systems (discrete-time Markov chains) and continuous-time Markov chains. We skip many details, which can be found in the above mentioned literature, and focus on the ideas of stochastic notions of bisimulation and simulation relations.

In the sequel, let Act be a fixed, finite set of actions. We assume that $\tau \in Act$ is a special action symbol for non-observable activities, i.e., computations that are internal to some process. All actions in $Act \setminus \{\tau\}$ are called visible. The symbol $\hat{a}$ equals a if a is a visible action, while $\hat{\tau} = \epsilon$ is the empty word in Act^*.

2 Markov chains

An action-labelled discrete-time Markov chain (DTMC for short) is a labelled transition system where each state is associated with a probability distribution that specifies the probabilities for the actions and successor states. That is, in any state s there is a probabilistic choice between the enabled transitions $s \xrightarrow{a} s'$. Formally, a DTMC is a tuple $\mathcal{D} = (S, \mathbf{P})$ where S is a countable set of states, $\mathbf{P} : S \times Act \times S \to [0,1]$ is a probability matrix satisfying $\sum_{s' \in S, a \in Act} \mathbf{P}(s,a,s') = 1$ for all $s \in S$.

We consider DTMCs as time-abstract models. The name DTMC has historical reasons. A (discrete-)timed interpretation is appropriate in settings where all state changes occur at equidistant time points. In contrast, CTMCs are considered as time-aware, as they have an explicit reference to time, in the form of transition rates which determine the stochastic evolution of the system in time. Formally, a CTMC is a tuple $\mathcal{C} = (S, \mathbf{R})$ with S as before, and *rate matrix* $\mathbf{R}$ a function that assigns to any triple (s,a,s') a non-negative real number such that $\sum_{s' \in S, a \in Act} \mathbf{R}(s,a,s')$ converges. If $\mathbf{R}(s,a,s') = 0$ then there is no a-labelled transition from s to s', otherwise the a-transition from s to s' has rate $\lambda = \mathbf{R}(s,a,s')$ which roughly means that $1/\lambda$ is the average delay of the transition $s \xrightarrow{a} s'$. The mean time spend in s without performing any action is $1/E(s)$ where $E(s) = \sum_{s' \in S, a \in Act} \mathbf{R}(s,a,s')$ is the so-called exit rate of state s. For simplicity, we assume here that all states have at least one outgoing transition, i.e., $E(s) > 0$ for all states s. The time-abstract probability for moving from s to s' via action a is $\mathbf{P}(s,a,s') = \mathbf{R}(s,a,s')/E(s)$. Then, $(S, \mathbf{P})$ is a DTMC, called the embedded DTMC of $\mathcal{C}$.

3 Strong bisimulation [29,27,17,11,23]

While in the non-probabilistic setting, bisimulation equivalence of two states requires that any transition of one state has at least one matching transition of the other state, probabilistic bisimulation takes the "quantity" (probabilities or rates) of transitions into account. For DTMCs, bisimulation equivalence denotes the coarsest equivalence $\sim_d$ on the state space such that for all $s_1 \sim_d s_2$, all actions a and all bisimulation equivalence classes C we have $\mathbf{P}(s_1, a, C) = \mathbf{P}(s_2, a, C)$ where $\mathbf{P}(s, a, C) = \sum_{s' \in C} \mathbf{P}(s, a, s')$ denotes the probability for s to move via an a-transition to a state in C. Similarly, for CTMCs, bisimulation equivalence denotes the coarsest equivalence $\sim_c$ on the state space such that for all $s_1 \sim s_2$, all actions a and all bisimulation equivalence classes C we have $\mathbf{R}(s_1, a, C) = \mathbf{R}(s_2, a, C)$ where $\mathbf{R}(s, a, C) = \sum_{s' \in C} \mathbf{R}(s, a, s')$ denotes the total rate to move from s via action a to a C-state.

$\sim_c$ refines $\sim_d$ in the sense that $\sim_c$ for a CTMC $\mathcal{C}$ is finer than $\sim_d$ for its embedded DTMC which again is finer than standard bisimulation equivalence in the labelled transition system obtained by ignoring the probabilities. Moreover, $\sim_d$ and $\sim_c$ have analogous properties as standard bisimulation equivalence in labelled transition systems. They fulfill several congruence properties for composition operators of probabilistic process calculi [17,23], have complete axiomatizations [27], logical characterizations by means of CTL-like branching time logics [2,4,15], coalgebraic characterizations [16,8] and polynomial-time decision algorithms [24,3,12].

4 Weak bisimulation [5,7]

While in *strong* (bi)simulations, all visible or non-visible steps are considered *weak* (bi)simulations abstract away from internal, non-observable steps. In the non-probabilistic setting several notions of weak bisimulation exist that differ in the underlying "weak transition relation" which combines the effect of consecutive τ-transitions. Corresponding notions for Markov chains can be provided by considering the cumulative probabilistic effect of τ-transitions. For instance, the analogue to Milner's observational equivalence can be defined for DTMCs as the coarsest equivalence $\approx_d$ such that for all $s_1 \approx_d s_2$, actions $a \in Act$ and equivalence classes $C \in S/\approx_d$ we have

$$\Pr(s_1, \tau^* \hat{a} \tau^*, C) = \Pr(s_2, \tau^* \hat{a} \tau^*, C)$$

where $\Pr(s, \tau^* \hat{a} \tau^*, C)$ denotes the the probability to move from s to a C-state via action sequences in $\tau^* \hat{a} \tau^*$. In contrast to the non-probabilistic setting, this notion of observational equivalence for DTMCs coincides with branching bisimulation equivalence á la van Glabbeek and Weijland [18]. Roughly speaking, branching bisimulation is defined as observational bisimulation equivalence except that the intermediate states in the $\tau^* \hat{a} \tau^*$-paths have to be equivalent to the starting state in the τ^*-prefix and to the target state in the τ^*-suffix. For DTMCs, branching bisimulation and observation bisimulation equivalence agree and they can be characterized by (1) a local probability condition and (2) a global reachability condition.

The local probability condition requires that for any equivalence class $B \in S/\approx_d$ the conditional probabilities

$$\frac{\mathbf{P}(s,a,C)}{1-\mathbf{P}(s,\tau,B)}$$

to move from s via action a to some equivalence class C, provided that either a visible action is executed or a non-visible action leading to some other equivalence class (i.e., $(a,C) \neq (\tau,B)$), agree for all states $s \in B$ where $\mathbf{P}(s,\tau,B) < 1$. The reachability condition is needed to distinguish divergent states from non-divergent ones. Formally, it requires that if there is some state $s \in B$ that can perform a visible action or can reach another equivalence class B' then the same holds for all states in B.

This latter characterization of observational equivalence can easily be adapted to CTMCs where we may deal with (1) the local probability condition in the embedded DTMC and (2) a rate condition that refines the reachability condition by the requirement $\mathbf{R}(s_1,a,C) = \mathbf{R}(s_2,a,C)$ for all $s_1 \approx_c s_2$ and $(a,C) \in Act \times (S/\approx_c)$ where $a \neq \tau$ or $s_i \notin C$, $i = 1,2$. This notion of weak bisimulation equivalence on CTMCs has a simple characterization: $\approx_c$ is the coarsest equivalence on the state space such that $\mathbf{R}(s_1,a,C) = \mathbf{R}(s_2,a,C)$ for all $s_1 \approx_c s_2$, $a \in Act$ and all equivalence classes C with $a \neq \tau$ or $s_1, s_2 \notin C$.

Although $\approx_d$ and $\approx_c$ are rather strong equivalences, they are the coarsest relations that preserve all branching-time properties of a temporal CTL-like logic [14,7]. A coarser notion of weak bisimulation for DTMCs has been suggested in [1] which relies on a nondeterministic transition relation for the τ-transitions and probabilistic choices for the visible actions. The local characterizations of weak bisimulation equivalence for DTMCs or CTMCs allow for decision algorithms that use similar ideas as the strong bisimulation algorithms [24] and run in polynomial-time.

5 Simulation relations [26,13,7]

The formal definition of simulation relations is more complicated for probabilistic systems than for labeled transition systems. The reason is that probability distributions rather than single states have to be compared. We skip the details and just mention that the formal strong simulation relies on (1) a local condition for the probabilities and (2) an additional rate condition for CTMCs. The local probability condition (1) can be formalized by means of so-called weight functions [25,26] that combine fragments of states, or alternatively by a quantitative criteria for the upward-closed sets:

if s_1 is simulated by s_2 then $\mathbf{P}(s_1,a,C\uparrow) \leq \mathbf{P}(s_2,a,C\uparrow)$ for all actions a and $C \subseteq S$.

Here, $C\uparrow$ denotes the upward-closure of C, i.e. the set of all states u that simulate a state $t \in C$. The formal definition of weak simulation is more complex as it relies on the identification of appropriate fragments of observable transitions for which the local probability condition and rate condition as for strong simulation are re-

quired. As for weak bisimulation, an additional reachability condition is needed to treat divergence in an appropriate way. Whereas (strong or weak) simulation equivalence in labeled transition systems is coarser than (strong or weak) bisimulation equivalence they agree for Markov chains.

Although these definition are rather complex, polynomial-time decision algorithms for finite-state Markov chains exists that rely on network-flow algorithms [3] or linear programs [6].

6 Conclusion

This note provided a brief introduction to simulation and bisimulation relations on Markov chain models. A comparative discussion of their features and properties including preservation results for fragments of the branching-time logics PCTL[19] and CSL [4] is provided in [7].

References

[1] S. Andova and J. Baeten. Abstraction in probabilistic process algebra. *Tools and Algorithms for the Construction and Analysis of Systems*, LNCS 2031, pp. 204–219, 2001.

[2] A. Aziz, V. Singhal, F. Balarin, R. Brayton and A. Sangiovanni-Vincentelli. It usually works: the temporal logic of stochastic systems. *Computer-Aided Verification*, LNCS 939, pp. 155–165, 1995.

[3] C. Baier, B. Engelen, and M. Majster-Cederbaum. Deciding bisimilarity and similarity for probabilistic processes. *J. of Comp. and System Sc.*, **60**(1):187–231, 2000.

[4] C. Baier, B.R. Haverkort, H. Hermanns and J.-P. Katoen. Model-checking algorithms for continuous-time Markov chains. *IEEE Trans. on Software Eng.*, **29**(6):524–541, 2003.

[5] C. Baier and H. Hermanns. Weak bisimulation for fully probabilistic processes. *Computer-Aided Verification*, LNCS 1254, pp. 119-130, 1997.

[6] C. Baier, H. Hermanns and J.-P. Katoen. Probabilistic weak simulation is decidable in polynomial time. *Inf. Proc. Lett.*, **89**(3):123–130, 2004.

[7] C. Baier, J.-P. Katoen, H. Hermanns and V. Wolf. Comparative branching-time semantics for Markov chains. to appear in *Information and Computation*.

[8] C. Baier and M.Z. Kwiatkowska. Domain Equations for Probabilistic Processes, *Math. Structures in Computer Science*, 10(6): 665-717, 2000.

[9] M. Bernardo and R. Gorrieri. Extended Markovian process algebra. *Concurrency Theory*, LNCS 1119, pp. 315-330, 1996.

[10] M. Bernardo and R. Cleaveland. A theory of testing for Markovian processes. *Concurrency Theory*, LNCS 1877, pp. 305–319, 2000.

[11] P. Buchholz. Exact and ordinary lumpability in finite Markov chains. *J. of Appl. Prob.*, **31**: 59–75, 1994.

[12] S. Derisavi, H. Hermanns and W.H. Sanders. Optimal state-space lumping in Markov chains. *Inf. Proc. Lett.*, **87**(6): 309–315, 2003.

[13] J. Desharnais. *Labelled Markov Processes.* PhD Thesis, McGill University, 1999.

[14] J. Desharnais, V. Gupta, R. Jagadeesan, P. Panangaden. Weak bisimulation is sound and complete for PCTL*. *Concurrency Theory*, LNCS 2421, pp. 355-370, 2002.

[15] J. Desharnais, P. Panangaden. Continuous stochastic logic characterizes bisimulation of continuous-time Markov processes. *J. of Logic and Alg. Progr.*, **56**: 99–115, 2003.

[16] E.P. De Vink, J.J.M.M. Rutten. Bisimulation for Probabilistic Transition Systems: a Coalgebraic Approach, *ICALP*, LNCS 1256, pp 460-470, 1997.

[17] R.J. van Glabbeek, S.A. Smolka, B. Steffen. Reactive, generative, and stratified models of probabilistic processes. *Inf. & Comput.*, **121**: 59–80, 1995.

[18] R.J. van Glabbeek, W.P. Weijland. Branching time and abstraction in bisimulation semantics. *J. ACM*, **43**(3): 555-600, 1996.

[19] H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing* **6**: 512–535, 1994.

[20] B. Haverkort. *Performance of Computer Communication Systems. A Model-Based Approach.* John Wiley & Sons. 1998.

[21] H. Hermanns. *Interactive Markov Chains.* LNCS 2428, 2002.

[22] H. Hermanns, U. Herzog and J.-P. Katoen. Process algebra for performance evaluation. *Theoretical Computer Science*, **274**(1-2), 2002.

[23] J. Hillston. *A Compositional Approach to Performance Modelling.* Cambr. Univ. Press, 1996.

[24] T. Hyunh, L. Tian. On some equivalence relations for probabilistic processes. *Fund. Inf.*, **17**: 211–234, 1992.

[25] C. Jones. *Probabilistic Non-Determinism.* Ph.D.Thesis, University of Edinburgh. 1990.

[26] B. Jonsson and K.G. Larsen. Specification and refinement of probabilistic processes. *IEEE Symp. on Logic in Comp. Sc.*, pp. 266-277, 1991.

[27] C.-C. Jou and S.A. Smolka. Equivalences, congruences, and complete axiomatizations for probabilistic processes. *Concurrency Theory*, LNCS 458, pp. 367–383, 1990.

[28] V.G. Kulkarni. *Modeling and Analysis of Stochastic Systems.* Chapman & Hall, 1995.

[29] K.G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Inf. and Comput.*, **94**(1): 1–28, 1991.

[30] N. Lopez, M. Nunez. An overview of probabilistic process algebras and their equivalences. Tutorial Proceedings *Validation of Stochastic Systems. A Guide to Current Research.* Lecture Notes in Computer Science 2925, pp 89–123, 2004.