Modeling and Verification of Probabilistic Systems

Lecture 19: Interactive Markov Chains

Joost-Pieter Katoen

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

http://www-i2.informatik.rwth-aachen.de/i2/mvps11/

July 12, 2011

Overview

Overview

Model-based performance evaluation

- Analyse performance metrics based on an abstract system model
  - formalisms: stochastic Petri nets, queueing networks, SANs, . . .
- The prevailing paradigm is continuous-time randomness
  - exponential distributions, i.e., continuous-time Markov processes
- Complexity of systems requires compositional approach
  - reflecting system architecture
- Enormous model sizes require compositional abstraction mechanisms
  - like bisimulation minimization
- Nondeterminism is at heart of compositionality

We need: Compositional Continuous-Time Markov Chains

Interactive Markov chains

Interactive Markov chain

An interactive Markov chain is a tuple $\mathcal{I} = (S, Act, \longrightarrow, \Longrightarrow, s_0)$ where

- S is a nonempty set of states with initial state $s_0 \in S$
- Act is a finite set of actions; $\tau \in Act$ is internal action
- $\longrightarrow \ \subseteq S \times Act \times S$ is a set of interactive transitions, and
- $\Longrightarrow \ \subseteq S \times \mathbb{R}_{>0} \times S$ is a set of Markovian transitions.

Thus:

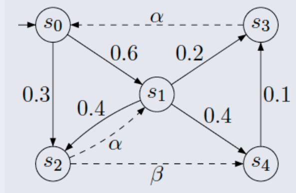
IMCs are labeled transition systems with action-labeled transitions, as well as Markovian transitions that are labeled with rates of exponential distributions. Any CTMC is an IMC; any LTS is an IMC.

1. $IT(s)$ the set of interactive transitions that leave s.
2. $MT(s)$ the set of Markovian transitions that leave s.

Interactive Markov chains

Classification of states

- s is Markovian if $MT(s) \neq \varnothing$ and $IT(s) = \varnothing$
- s is interactive if $MT(s) = \varnothing$ and $IT(s) \neq \varnothing$
- s is hybrid if $MT(s) \neq \varnothing$ and $IT(s) \neq \varnothing$
- s is timelock if $MT(s) = IT(s) = \varnothing$



For Markovian state s, let:

- $\mathbf{R}(s, s') = \sum \left\{ \lambda \mid s \overset{\lambda}{\Longrightarrow} s' \right\}$ be the rate to move from s to s',
- $r(s) = \sum_{s' \in S} \mathbf{R}(s, s')$ be the exit rate of s
- $\mathbf{P}(s, s') = \frac{\mathbf{R}(s,s')}{E(s)}$ is the probability to move from s to s'

Example

$\mathbf{R}(s_0, s_2) = 0.3$, $r(s_0) = 0.3 + 0.6 = 0.9$ and $\mathbf{P}(s_0, s_2) = \frac{1}{3}$.

Maximal progress assumption

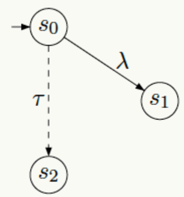
Maximal progress

1. Internal (action) transitions are labeled with the action τ.
2. These transitions will not be subject to interaction.
3. They cannot be delayed by other components.
4. Thus, internal interactive transitions can trigger immediately.
5. But, the probability to execute Markovian transitions immediately is zero.

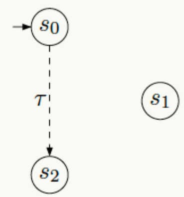
Maximal progress assumption

Internal transitions take precedence over Markovian ones.

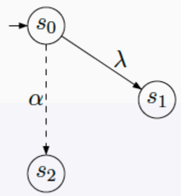
Maximal progress



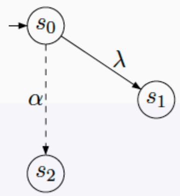
reduces to



But as visible actions may be subject to delaying by other components:



remains



Overview

Parallel composition

Let $\mathcal{I}_1 = (S_1, Act_1, \longrightarrow_1, \Longrightarrow_1, s_{0,1})$ and $\mathcal{I}_2 = (S_2, Act_2, \longrightarrow_2, \Longrightarrow_2, s_{0,2})$ be IMCs. The *parallel composition* of $\mathcal{I}_1$ and $\mathcal{I}_2$ wrt. $A \subseteq Act \setminus \{ \tau \}$ is:

$$\mathcal{I}_1 \,||_A\, \mathcal{I}_2 = (S_1 \times S_2, Act_1 \cup Act_2, \longrightarrow, \Longrightarrow, (s_{0,1}, s_{0,2}))$$

where $\longrightarrow$ and $\Longrightarrow$ are defined as the smallest relations satisfying:

(SYNC) $$\frac{s_1 \xrightarrow{\alpha}_1 s_1' \text{ and } s_2 \xrightarrow{\alpha}_2 s_2' \text{ and } \alpha \in A}{(s_1, s_2) \xrightarrow{\alpha} (s_1', s_2')}$$

(ASYNC) $$\frac{s_1 \xrightarrow{\alpha}_1 s_1' \text{ and } \alpha \notin A}{(s_1, s_2) \xrightarrow{\alpha} (s_1', s_2)} \qquad \frac{s_2 \xrightarrow{\alpha}_2 s_2' \text{ and } \alpha \notin A}{(s_1, s_2) \xrightarrow{\alpha} (s_1, s_2')}$$

(DELAY) $$\frac{s_1 \overset{\lambda}{\Longrightarrow}_1 s_1'}{(s_1, s_2) \overset{\lambda}{\Longrightarrow} (s_1', s_2)} \qquad \frac{s_2 \overset{\lambda}{\Longrightarrow}_2 s_2'}{(s_1, s_2) \overset{\lambda}{\Longrightarrow} (s_1, s_2')}$$

Processes delay independently as in interleaving.
Due to the memoryless property, this is correct.

Parallel composition: example

Hiding

Hiding

The *hiding* of IMC $\mathcal{I} = (S, Act, \longrightarrow, \Longrightarrow, s_0)$ wrt. the set $A \subseteq Act \setminus \{ \tau \}$ of actions is the IMC $\mathcal{I} \setminus A = (S, Act \setminus A, \longrightarrow', \Longrightarrow, s_0)$ where $\longrightarrow'$ is the smallest relation defined by:

1. $s \xrightarrow{\alpha} s'$ and $\alpha \notin A$ implies $s \xrightarrow{\alpha}{}' s'$, and
2. $s \xrightarrow{\alpha} s'$ and $\alpha \in A$ implies $s \xrightarrow{\tau}{}' s'$.

- Hiding transforms α-transitions with $\alpha \in A$ into τ-transitions.
- Turning an α-transition emanating from state s into a τ-transition may change the semantics of the IMC, as now —due to maximal progress— never a Markovian transition in s will be taken.

Hiding: example

Overview

Strong bisimulation

For $C \subseteq S$, $\alpha \in Act$, let $\mathbf{T}(s, \alpha, C) = 1$ iff $\{s' \in C \mid s \xrightarrow{\alpha} s'\} \neq \varnothing$.

Intuition: $\mathbf{T}(s, \alpha, C) = 1$ iff s can move to C via an α-transition.

Strong bisimulation

Let $\mathcal{I} = (S, Act, \longrightarrow, \Longrightarrow, s_0)$ be an IMC. Equivalence $R \subseteq S \times S$ is a *strong bisimulation* on $\mathcal{I}$ if for any $(s, t) \in R$ and $C \in S/R$:

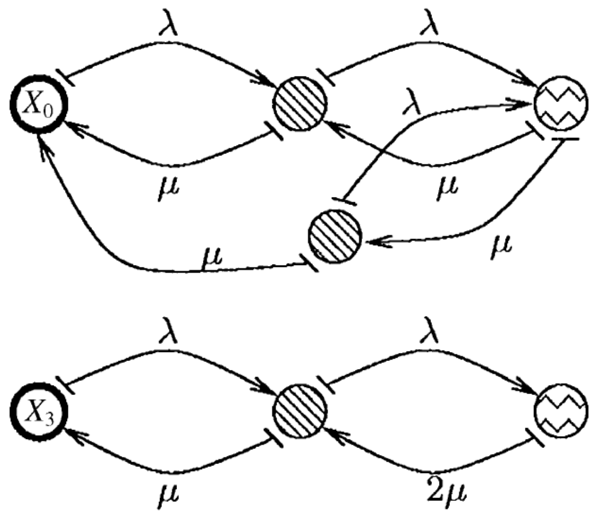
1. for any $\alpha \in Act$, $\mathbf{T}(s, \alpha, C) = \mathbf{T}(t, \alpha, C)$, and
2. $s \stackrel{\tau}{\not\longrightarrow}$ implies $\mathbf{R}(s, C) = \mathbf{R}(t, C)$, where $\mathbf{R}(s, C) = \sum_{s' \in C} \mathbf{R}(s, s')$.

Let $\sim$ be the largest strong bisimulation.

Congruence

$\sim$ is a congruence wrt. parallel composition and hiding.

Strong bisimulation – Example



Weak bisimulation

- Main idea: IMCs are equivalent when they exhibit similar observable system behaviour.
- Sequences of $\xrightarrow{\tau}$-transitions can be collapsed as for branching bisimulation.
- What about sequences of $\Longrightarrow$-transitions? Can they be mimicked?
- Yes, but a sequence of exponential distributions is a phase-type distribution, and not an exponential distribution.
- Thus, sequences of $\Longrightarrow$-transitions cannot be mimicked exactly by exponential distributions.
- $\Longrightarrow$-transitions will therefore be treated as in strong bisimulation.
- As before, rate equality is only required for stable states.

Weak bisimulation

For $C \subseteq S$, let $\mathbf{W}(s, \alpha, C) = 1$ iff $\{s' \in C \mid s \xrightarrow{\tau^*} \xrightarrow{\alpha \neq \tau} \xrightarrow{\tau^*} s'\} \neq \varnothing$.

$\mathbf{W}(s, \alpha, C) = 1$ iff s can weakly move to C via an α-transition.

Weak bisimulation

Let $\mathcal{I} = (S, Act, \longrightarrow, \Longrightarrow, s_0)$ be an IMC. Equivalence $R \subseteq S \times S$ is a *weak bisimulation* on $\mathcal{I}$ if for any $(s, t) \in R$ and $C \in S/R$:

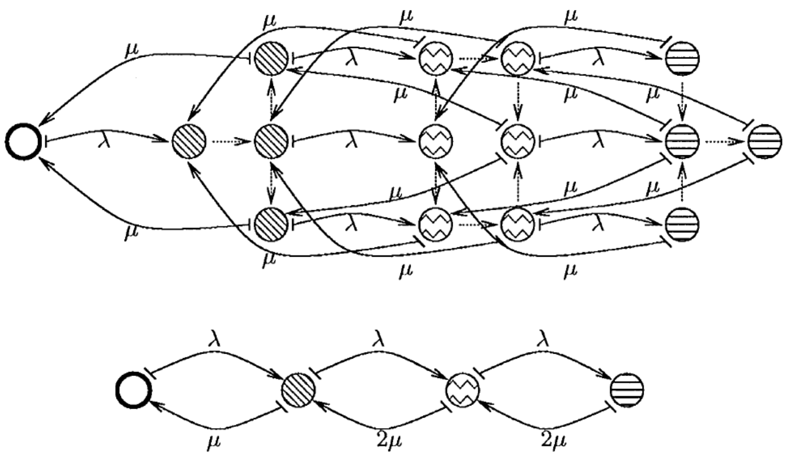
1. for any $\alpha \in Act$, $\mathbf{W}(s, \alpha, C) = \mathbf{W}(t, \alpha, C)$, and
2. $s \xrightarrow{\tau^*} s'$ and $s' \stackrel{\tau}{\not\longrightarrow}$ implies
   $t \xrightarrow{\tau^*} t'$ and $t' \stackrel{\tau}{\not\longrightarrow}$ and $\mathbf{R}(s', C) = \mathbf{R}(t', C)$, for some $t' \in S$.

Let $\approx$ be the largest weak bisimulation.

Congruence

$\approx$ is a congruence wrt. parallel composition and hiding.

Weak bisimulation – Example



Constraint-oriented performance modeling

- Let α and β be two successive actions in LTS P.
- Let D_p be of the form $\alpha; D; \beta$ where D is a phase-type distribution.
- A phase-type distribution is given by the (random) time until absorption in an absorbing CTMC. D is thus an IMC.
- Adding a random time constraint on top of P yields: $P \,||_{\{\alpha,\beta\}}\, D_p$.
- Applying this to two processes yields: $(P \,||_A\, Q) \,||_{A_p \cup A_q}\, (D_p \,||_\varnothing\, D_q)$
- If D_p (D_q) only delays local actions from P (Q), then this is weak bisimilar to:

$$\underbrace{\left(P \,||_{A_p}\, D_p\right)}_{\text{local constraints of } P} \;||_A\; \underbrace{\left(Q \,||_{A_q}\, D_q\right)}_{\text{local constraints of } Q}$$

- Functional and performance aspects are separated constraints.

Closed IMC models

Closed IMC model

The typical specification that is subject to analysis is of the form:

$$\left(\mathcal{I}_1 \,||_{A_1}\, \mathcal{I}_2 \,||_{A_2}\, \ldots \,||_{A_{N-1}}\, \mathcal{I}_N\right) \setminus A$$

where A is the union of all actions, i.e., $A = \cup_{i=1}^{N-1} Act_i \setminus \{\, \tau \,\}$.

It is closed, as no action is subject to further interaction.

Compositional minimisation

- The preservation results suggest to compute the quotient IMC prior to their analysis.
- This leads to significant state-space reductions and efficiency gains in computation times.
- Even better: bisimulation being a congruence wrt. $||$, enables compositional minimisation:

$$\forall 0 < j \leqslant N.\ \mathcal{I}_j \sim \widehat{\mathcal{I}_j} \quad \text{implies}$$

$$\mathcal{I}_1 \,||_{A_1}\, \ldots \,||_{A_{N-1}}\, \mathcal{I}_N \ \sim\ \widehat{\mathcal{I}_1} \,||_{A_1}\, \ldots \,||_{A_{N-1}}\, \widehat{\mathcal{I}_N}.$$

- This technique has allowed the performance analysis of some systems that could not be handled before.

An exemplary case

MONOLITHIC CONSTRUCTION FOR ETCS WITH 2 TRAINS

Phases	Monolithic Construction			
	States	Transitions	G Time (sec.)	M Time (sec.)
1	33600	518464	12	3
5	302400	4142016	22	402
10	1016400	13521376	46	5154

EXPLICIT STEPS: COMPOSITION AND MINIMIZATION STATISTICS

Trains	Phases	Compositional Construction			Final Quotient IMC	
		States	Transitions	G + M Time (sec.)	States	Transitions
2	1	600	2505	42	355	1590
	5	10000	53625	61	5875	39500
	10	37500	207500	511	20000	154750
3	1	3240	16064	58	1375	5225
	5	64440	354100	813	36070	159119
	10	249480	1382900	10666	113650	533500
4	1	2870	11260	53	1435	5475
	5	57950	260350	420	30575	141000
	10	224900	1022700	7391	119650	558500

Reduction to a CTMC

Reduction strategy:

1. Apply maximal progress: remove all Markovian transitions in unstable states.
2. Apply weak bisimulation minimisation. This requires computing transitive closure.
3. In absence of nondeterminism, the resulting IMC is a CTMC.
4. Analyse the CTMC (transient/stationary) using standard numerical techniques.
5. Or, apply model checking, e.g., branching time (CSL), or linear time (DTA).

Nondeterminism

What if nondeterminism is still present? Can we still analyse IMCs? How? Which measures? At which cost? This is a current topic of research.

Overview

Summary

Main points

- IMCs are combine LTS and CTMCs in an orthogonal manner.
- Maximal progress = immediate transitions take precedence over Markovian ones.
- IMCs are closed under parallel composition and hiding.
- Strong bisimulation requires rate equality for stable states only.
- Weak bisimulation allows for collapsing sequences of internal transitions.
- Strong and weak bisimulation are substitutive wrt. parallel composition and hiding.
- Weak bisimulation may allow for removing nondeterminism yielding CTMCs.

Practical usage

Take-home message

IMCs are used as semantical model for practical modeling languages such as: dynamic fault trees, AADL, generalized stochastic Petri nets, and so forth.