Modeling and Verification of Probabilistic Systems

Joost-Pieter Katoen

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

http://moves.rwth-aachen.de/teaching/ss-14/movep14/

April 15, 2014

Overview

Theme of the course

The theory of modelling and verification
of probabilistic systems

Probabilities help

- When analysing system performance and dependability
  - to quantify arrivals, waiting times, time between failure, QoS, ...
- When modelling unreliable and unpredictable system behavior
  - to quantify message loss, processor failure
  - to quantify unpredictable delays, express soft deadlines, ...
- When building protocols for networked embedded systems
  - randomized algorithms
- When problems are undecidable deterministically
  - repeated reachability of lossy channel systems, . . .

Illustrative example: Security

Security: Crowds protocol **[Reiter & Rubin, 1998]**

- A protocol for anonymous web browsing (variants: mCrowds, BT-Crowds)
- Hide user's communication by random routing within a crowd
  - sender selects a crowd member randomly using a uniform distribution
  - selected router flips a biased coin:
    - with probability $1 - p$: direct delivery to final destination
    - otherwise: select a next router randomly (uniformly)
  - once a routing path has been established, use it until crowd changes
- Rebuild routing paths on crowd changes
- Property: Crowds protocol ensures "probable innocence":
  - probability real sender is discovered $< \frac{1}{2}$ if $N \geqslant \frac{p}{p-\frac{1}{2}} \cdot (c+1)$
  - where N is crowd's size and c is number of corrupt crowd members

Illustrative example: Leader election

Distributed system: Leader election **[Itai & Rodeh, 1990]**

- A round-based protocol in a synchronous ring of $N > 2$ nodes
  - the nodes proceed in a lock-step fashion
  - each slot = 1 message is read + 1 state change + 1 message is sent
  - $\Rightarrow$ this synchronous computation yields a discrete-time Markov chain
- Each round starts by each node choosing a uniform id $\in \{1, \ldots, K\}$
- Nodes pass their selected id around the ring
- If there is a unique id, the node with the maximum unique id is leader
- If not, start another round and try again ...

Properties of leader election

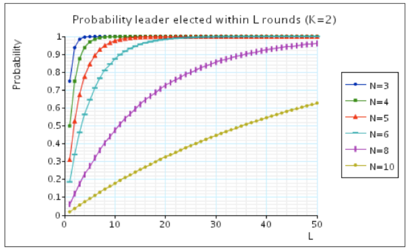
Almost surely eventually a leader will be elected

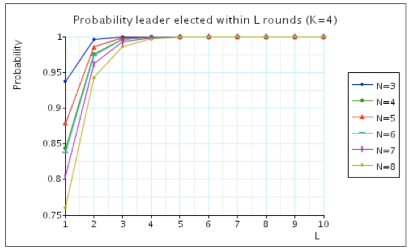
$$\mathbb{P}_{=1}\left(\Diamond \textit{leader elected}\right)$$

With probability at least 0.8, a leader is elected within k steps

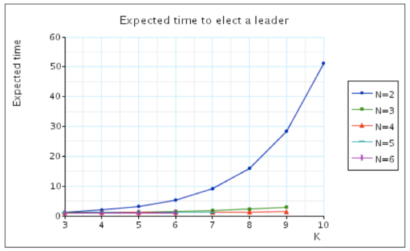
$$\mathbb{P}_{\geqslant 0.8}\left(\Diamond^{\leqslant k} \textit{leader elected}\right)$$

Probability to elect a leader within L rounds

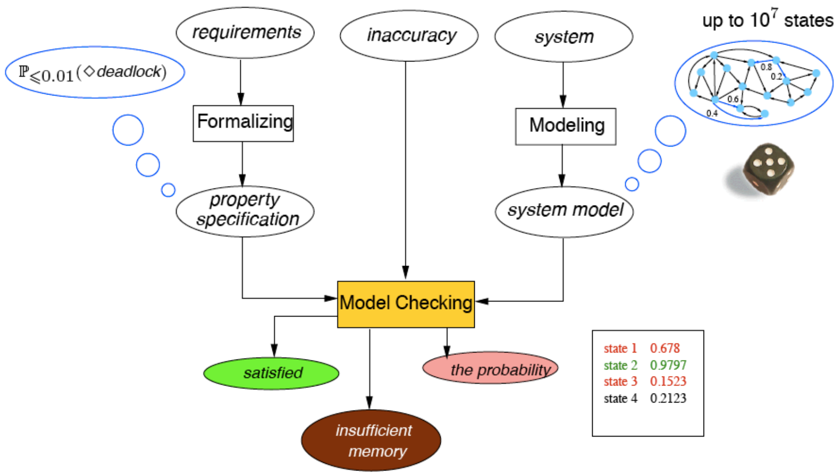




$$\mathbb{P}_{\leqslant q}\left(\Diamond^{\leqslant (N+1)\cdot L} \textit{ leader elected}\right)$$



What is probabilistic model checking?



Probabilistic models

	Nondeterminism no	Nondeterminism yes
Discrete time	discrete-time Markov chain (DTMC)	Markov decision process (MDP)
Continuous time	CTMC	interactive MC

Probabilistic models

	Nondeterminism no	Nondeterminism yes
Discrete time	discrete-time Markov chain (DTMC)	Markov decision process (MDP)
Continuous time	CTMC	CTMDP

Some other models: probabilistic variants of (priced) timed automata

Properties

	Logic	Monitors
Discrete time	probabilistic CTL	deterministic automata (safety and LTL)
Continuous time	probabilistic timed CTL	deterministic timed automata

Core problem: computing (timed) reachability probabilities

Overview

1. Introduction
2. Course details
3. Probability refresher
   - Random variables
   - Probability spaces
   - Random variables
   - Stochastic processes

Course topics

A probability theory refrehser

- measurable spaces, σ-algebra, measurable functions
- geometric, exponential and binomial distributions
- Markov and memoryless property
- limiting and stationary distributions

What are probabilistic models?

- discrete-time Markov chains
- continuous-time Markov chains
- extensions of these models with rewards
- Markov decision processes (or: probabilistic automata)
- interactive Markov chains

Course topics

What are properties?

- reachability probabilities, i.e., $\Diamond G$
- long-run properties
- linear temporal logic
- probabilistic computation tree logic

How to check temporal logic properties?

- graph analysis, solving systems of linear equations
- deterministic Rabin automata, product construction
- linear programming, integral equations
- uniformisation, Volterra integral equations

Course topics

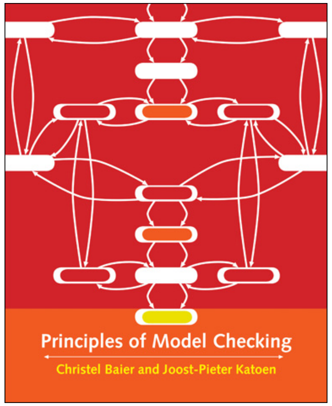
How to make probabilistic models smaller?

- Equivalences and pre-orders
- Which properties are preserved?

How to model probabilistic models?

- parallel composition and hiding
- compositional modeling and minimisation

Course material



Ch. 10, Principles of Model Checking

CHRISTEL BAIER

TU Dresden, Germany

JOOST-PIETER KATOEN

RWTH Aachen University, Germany, and University of Twente, the Netherlands

Other literature

- H.C. Tijms: A First Course in Stochastic Models. Wiley, 2003.
- H. Hermanns: Interactive Markov Chains: The Quest for Quantified Quality. LNCS 2428, Springer-Verlag, 2002.
- J.-P. Katoen. Model Checking Meets Probability: A Gentle Introduction. IOS Press, 2013. (see course web-page for download)
- M. Stoelinga. An Introduction to Probabilistic Automata. Bull. of the ETACS, 2002.
- M. Kwiatkowska *et al.*. Stochastic Model Checking. LNCS 4486, Springer-Verlag, 2007.

Lectures

Lecture

- Tue 13:00 - 14:30 (9U10), Thu 13:00-14:30 (9U10)
- April 15, 17, 22, 24, 29
- May 8, 13, 15, 20, 22, 27
- June 3*, 5, 17, 24, 26
- July 1, 3, 8, 10, 15
- Check regularly course webpage for possible "no shows"

Material

- Lecture slides (with gaps) are made available on webpage
- Copies of the books are available in the CS library

Website

http://moves.rwth-aachen.de/teaching/ss-14/movep14/

Exercises and exam

Exercise classes

- Thu 15:00 - 16:30 in 9U10 (start: April 24)
- Instructors: Nils Jansen and Benjamin Kaminski

Weekly exercise series

- Intended for groups of 2 students
- New series: every Thu on course webpage (start: April 17)
- Solutions: Thu (before 15:00) one week later

Exam:

- unknown date (written or oral exam)
- participation if $\geqslant$ 40% of all exercise points are gathered

Course embedding

Aim of the course

It's about the foundations of verifying and modeling probabilistic systems

Prerequisites

- Automata and language theory
- Algorithms and data structures
- Probability theory
- Introduction to model checking

Some related courses

- Advanced Model Checking (Katoen)
- Modeling and Verification of Hybrid Systems (Abráhám)
- Applied Automata Theory (Thomas)

Questions?

Overview

1. Introduction
2. Course details
3. Probability refresher
   - Random variables
   - Probability spaces
   - Random variables
   - Stochastic processes

Probability theory is simple, isn't it?

In no other branch of mathematics
is it so easy to make mistakes
as in probability theory

Henk Tijms, "Understanding Probability" (2004)

Measurable space

Sample space

A *sample space* Ω of a chance experiment is a set of elements that have a 1-to-1 relationship to the possible outcomes of the experiment.

σ-algebra

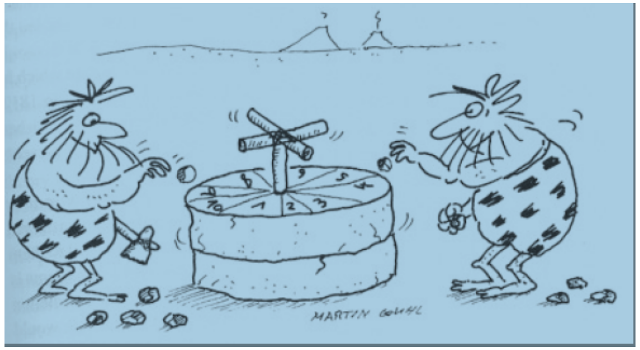
A *σ-algebra* is a pair $(\Omega, \mathcal{F})$ with $\Omega \neq \varnothing$ and $\mathcal{F} \subseteq 2^{\Omega}$ a collection of subsets of sample space Ω such that:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \;\Rightarrow\; \Omega - A \in \mathcal{F}$ — complement
3. $(\forall i \geqslant 0.\ A_i \in \mathcal{F}) \;\Rightarrow\; \bigcup_{i \geqslant 0} A_i \in \mathcal{F}$ — countable union

The elements in $\mathcal{F}$ of a σ-algebra $(\Omega, \mathcal{F})$ are called *events*.
The pair $(\Omega, \mathcal{F})$ is called a *measurable space*.

Let Ω be a set. $\mathcal{F} = \{ \varnothing, \Omega \}$ yields the smallest σ-algebra; $\mathcal{F} = 2^{\Omega}$ yields the largest one.

Probabilities



Probability space

Probability space

A *probability space* $\mathcal{P}$ is a structure $(\Omega, \mathcal{F}, Pr)$ with:

- $(\Omega, \mathcal{F})$ is a σ-algebra, and
- $Pr : \mathcal{F} \to [0, 1]$ is a *probability measure*, i.e.:
  1. $Pr(\Omega) = 1$, i.e., Ω is the certain event
  2. $Pr\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} Pr(A_i)$ for any $A_i \in \mathcal{F}$ with $A_i \cap A_j = \varnothing$ for $i \neq j$, where $\{ A_i \}_{i \in I}$ is finite or countably infinite.

The elements in $\mathcal{F}$ of a probability space $(\Omega, \mathcal{F}, Pr)$ are called *measurable* events.

Some lemmas

Properties of probabilities

For measurable events A, B and A_i and probability measure Pr:

- $Pr(A) = 1 - Pr(\Omega - A)$
- $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$
- $Pr(A \cap B) = Pr(A \mid B) \cdot Pr(B)$
- $A \subseteq B$ implies $Pr(A) \leqslant Pr(B)$
- $Pr(\bigcup_{n \geqslant 1} A_n) \;=\; \sum_{n \geqslant 1} Pr(A_n)$ provided A_n are pairwise disjoint

Discrete probability space

Discrete probability space

Pr is a *discrete* probability measure on $(\Omega, \mathcal{F})$ if

- there is a countable set $A \subseteq \Omega$ such that for $a \in A$:

$$\{ a \} \in \mathcal{F} \quad \text{and} \quad \sum_{a \in A} Pr(\{ a \}) = 1$$

- e.g., a probability measure on $(\Omega, 2^{\Omega})$

$(\Omega, \mathcal{F}, Pr)$ is then called a *discrete* probability space; otherwise, it is a *continuous probability* space.

Example

Example discrete probability space: throwing a die, number of customers in a shop, ...

Example

Example continuous probability space: throwing a dart on a circular board (see black board), water tank level, ...

Random variable

Measurable function

Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be measurable spaces. Function $f : \Omega \to \Omega'$ is a *measurable function* if

$$f^{-1}(A) = \{ a \mid f(a) \in A \} \in \mathcal{F} \quad \text{for all } A \in \mathcal{F}'$$

Random variable

Measurable function $X : \Omega \to \mathbb{R}$ is a *random variable*.

The *probability distribution* of X is $Pr_X = Pr \circ X^{-1}$ where Pr is a probability measure on $(\Omega, \mathcal{F})$.

Example: rolling a pair of fair dice

Distribution function

Distribution function

The *distribution function* F_X of random variable X is defined by:

$$F_X(d) = Pr_X((-\infty, d]) = Pr(\underbrace{\{ a \in \Omega \mid X(a) \leqslant d \}}_{\{ X \leqslant d \}}) \quad \text{for real } d$$

Properties

- F_X is monotonic and right-continuous
- $0 \leqslant F_X(d) \leqslant 1$
- $\lim_{d \to -\infty} F_X(d) = 0$ and
- $\lim_{d \to \infty} F_X(d) = 1$.

Discrete / continuous random variables

Distribution function

The *distribution function* F_X of random variable X is defined for $d \in \mathbb{R}$ by:

$$F_X(d) \;=\; Pr_X(X \in (-\infty, d]) = Pr(\{\, a \in \Omega \mid X(a) \leqslant d \,\})$$

In the continuous case, F_X is called the *cumulative density function*.

Distribution function

- For discrete random variable X, F_X can be written as:

$$F_X(d) = \sum_{d_i \leqslant d} Pr_X(X{=}d_i)$$

- For continuous random variable X, F_X can be written as:

$$F_X(d) = \int_{-\infty}^{d} f_X(u)\, du \quad \text{with } f \text{ the density function}$$

Expectation and variance

Expectation

The *expectation* of discrete r.v. X with range I is defined by

$$E[X] \;=\; \sum_{x_i \in I} x_i \cdot Pr_X(X{=}x_i)$$

provided that this series converges absolutely, i.e., the sum must remain finite on replacing all x_i's with their absolute values.

The expectation is the weighted average of all possible values that X can take on.

Variance

The *variance* of discrete r.v. X is given by $Var[X] = E[X^2] - (E[X])^2$.

Stochastic process

Stochastic process

A *stochastic process* is a collection of random variables $\{\, X_t \mid t \in T \,\}$.

- casual notation $X(t)$ instead of X_t
- with all X_t defined on probability space $\mathcal{P}$
- parameter t (mostly interpreted as "time") takes values in the set T

X_t is a random variable whose values are called *states*. The set of all possible values of X_t is the *state space* of the stochastic process.

	Parameter space T	
State space	Discrete	Continuous
Discrete	# jobs at k-th job departure	# jobs at time t
Continuous	waiting time of k-th job	total service time at time t

Example stochastic processes

- Waiting times of customers in a shop
- Interarrival times of jobs at a production lines
- Service times of a sequence of jobs
- Files sizes that are downloaded via the Internet
- Number of occupied channels in a wireless network
- ......

Bernouilli process

Bernouilli random variable

Random variable X on state space $\{0, 1\}$ defined by:

$$Pr(X = 1) = p \quad \text{and} \quad Pr(X = 0) = 1 - p$$

is a *Bernouilli* random variable.

The mass function is given by $f(k; p) = p^k \cdot (1-p)^{1-k}$ for $k \in \{0, 1\}$.

Expectation $E[X] = p$; variance $Var[X] = E[X^2] - (E[X])^2 = p \cdot (1-p)$.

Bernouilli process

A *Bernouilli process* is a sequence of independent and identically distributed Bernouilli random variables $X_1, X_2, \ldots$.

Binomial process

Binomial process

Let $X_1, X_2, \ldots$ be a Bernouilli process. The *binomial* process S_n is defined by $S_0 = 0$ and $S_n = \sum_{i=1}^{n} X_i$. The probability distribution of "counting process" S_n is given by:

$$Pr\{ S_n = k \} = \binom{n}{k} p^k \cdot (1-p)^{n-k} \quad \text{for } 0 \leqslant k \leqslant n$$

Moments: $E[S_n] = n \cdot p$ and $Var[S_n] = n \cdot p \cdot (1-p)$.

Geometric distribution

Let r.v. T_i be the number of steps between increments of counting process S_n. Then:

$$Pr\{ T_i = k \} = (1-p)^{k-1} \cdot p \quad \text{for } k \geqslant 1$$

This is a *geometric distribution*. We have $E[T_i] = \frac{1}{p}$ and $Var[T_i] = \frac{1-p}{p^2}$.

Intuition: Geometric distribution $=$ number of Bernoulli trials needed for one success.

Geometric distribution

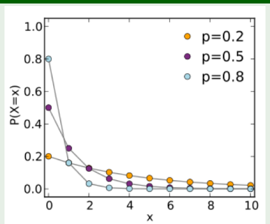
Geometric distribution

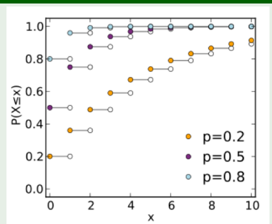
Let X be a discrete random variable, natural $k > 0$ and $0 < p \leqslant 1$. The mass function of a *geometric distribution* is given by:

$$Pr\{ X = k \} = (1-p)^{k-1} \cdot p$$

We have $E[X] = \frac{1}{p}$ and $Var[X] = \frac{1-p}{p^2}$ and cdf $Pr\{ X \leqslant k \} = 1 - (1-p)^k$.

Geometric distributions and their cdf's





Memoryless property

Theorem

1. For any random variable X with a geometric distribution:

   $$Pr\{X = k + m \mid X > m\} \;=\; Pr\{X = k\} \quad \text{for any} \quad m \in T, k \geqslant 1$$

   This is called the memoryless property, and X is a memoryless r.v..
2. Any discrete random variable which is memoryless is geometrically distributed.

Proof:

On the black board.

Joint distribution function

Joint distribution function

The *joint* distribution function of stochastic process $X = \{ X_t \mid t \in T \}$ is given for n, $t_1, \ldots, t_n \in T$ and $d_1, \ldots, d_n$ by:

$$F_X(d_1, \ldots, d_n; t_1, \ldots, t_n) = Pr\{ X(t_1) \leqslant d_1, \ldots, X(t_n) \leqslant d_n \}$$

The shape of F_X depends on the stochastic dependency between $X(t_i)$.

Stochastic independence

Random variables X_i on probability space $\mathcal{P}$ are *independent* if:

$$F_X(d_1, \ldots, d_n; t_1, \ldots, t_n) \ = \ \prod_{i=1}^{n} F_X(d_i; t_i) \ = \ \prod_{i=1}^{n} Pr\{ X(t_i) \leqslant d_i \}.$$

A renewal process is a discrete-time stochastic process where $X(t_1), X(t_2), \ldots$ are independent, identically distributed, non-negative random variables.

The next state of the stochastic process only depends on the current state, and not on states assumed previously. This is the Markov property.