PSC 2020/21 (375AA, 9CFU)

Principles for Software Composition

Roberto Bruni

http://www.di.unipi.it/~bruni/

http://didawiki.di.unipi.it/doku.php/magistraleinformatica/psc/

25a - DTMC
Probability
Probability

Nondeterminism: unpredictable future

Probability: quantitative estimation

- how likely is a series of events?
- how likely is to find the system in a given state?
- what is the expected throughput of the system?
Models

probabilistic models:
when many actions are enabled at the same time
the system uses a probability measure
to choose what to do next

stochastic models:
each event has a duration
a random variable is bound to each action
(it represents the time needed to perform the action)
exponential distribution (memoryless, defined by a rate)
when a race between events is enabled,
the fastest action is taken
Probabilistic programming

Quantum computing
Approximate computing
Randomised algorithms
Bayesian networks
Security protocols
Biological modelling
Reliability analysis
Decision making

functional / imperative programs
randomly drawn values
condition values by observations
Stochastic models

Markovian queueing networks
Stochastic Petri nets
Stochastic activity networks
Stochastic process algebras
Calculi for biological systems
Interactive Markov chains
Performance analysis
probabilistic puzzles
Fair die

Take a fair die and toss it ten times

which sequence is more likely?

1 1 1 1 1 1 1 1 1 1

A - first sequence
B - second sequence
C - equally likely
D - don’t know

1 4 3 2 5 1 6 2 4 5
Fair coins

I take two fair coins, toss one and then the other (without showing the outcomes to you)

You can bet about the fact that the coins give equal results or different ones: your winning chances are greater if you bet on

A - equal results
B - different results
C - equally likely
D - don’t know
Fair coins

I take two fair coins, toss one and then the other (without showing the outcomes to you)

You can bet about the fact that the coins give equal results or different ones: your winning chances are greater if you bet on what if I tell you one is head

A - equal results  H H
B - different results  H T
C - equally likely  T H
D - don’t know
Fair coins

I take two fair coins, toss one and then the other (without showing the outcomes to you)

You can bet about the fact that the coins give equal results or different ones: your winning chances are greater if you bet on what if I tell you the first is head

A - equal results
B - different results  H H
C - equally likely  H T
D - don’t know
Monty Hall problem

my favourite puzzle: Monty Hall problem
(highly controversial)

loosely based on an American TV game show called “Let’s make a deal” (1963)

named after its original host Monty Hall
(serving for nearly 30 years)
Monty Hall problem

first posed and solved in 1975
the puzzle became famous in 1990
after it was posted on a column of an
American Sunday newspaper magazine (Parade)

many readers were disappointed by the solution
and did not believe it (10,000 or more)

people wrote to the magazine claiming the solution was wrong

Paul Erdos, a great mathematician, remained unconvinced
until he was shown a computer simulation
Monty Hall problem

the puzzle comes in many variants, here is the most popular one

you are guest of the show, playing the final game

three closed doors, behind them: a brand new car

two goats

other versions:
three boxes, two empty, one has the key of the car
Monty Hall problem

you have to pick one door

the host opens one of the other doors where he knows there is a goat

you are given the possibility to keep your choice or change it

what is the best strategy to win the car?

A - keep
B - change
C - equally likely
D - don’t know
probabilistic systems
**sigma-field**

\[ \Omega \] elementary events (possible outcomes)

\[ A \subseteq \mathcal{F}(\Omega) \] a set of events we are interested in

a family of subsets of elementary events

such that

\[ \emptyset \in A \] the impossible event is present

\[ A \in A \Rightarrow (\Omega \setminus A) \in A \] closed under complementation

\[ \{A_n\}_{n \in \mathbb{N}} \subseteq A \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in A \] closed under countable union
**sigma-field: properties**

\[ \emptyset \in \mathcal{A} \]

1. the impossible event is present

\[ A \in \mathcal{A} \Rightarrow (\Omega \setminus A) \in \mathcal{A} \]

2. closed under complementation

\[ \{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A} \]

3. closed under countable union

\[ \Omega \in \mathcal{A} \]

by 1 and 2

\[ \{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \Rightarrow \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A} \]

by 2 and 3

\[ \bigcap_{n \in \mathbb{N}} A_n = \Omega \setminus \bigcup_{n \in \mathbb{N}} (\Omega \setminus A_n) \]
**sigma-field: properties**

in simpler terms

if $A$ and $B$ are events

$A \cup B$ is an event (one of the two events happens)

$A \cap B$ is an event (two events happen together)

$\overline{A}$ is an event (one event is not going to happen)

examples:

$\Omega = \{ \text{HH, HT, TH, TT} \}$  \hspace{1cm} $\mathcal{A} = \varnothing(\Omega)$

$\mathcal{A} = \{ \emptyset , \{\text{HH, TT}\} , \{\text{HT, TH}\} , \Omega \}$
Probability space

\[ P : \mathcal{A} \to [0, 1] \]

\[ P(\emptyset) = 0 \]

\[ P \left( \bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} P(A_n) \quad \text{if } \{A_n\}_{n \in \mathbb{N}} \text{ are pairwise disjoint} \]

\[ P(\Omega) = 1 \]

probability space: \( (\Omega, \mathcal{A}, P) \)

a \( \sigma \)-field with a probability measure
Prob space: properties

\[ P(\Omega \setminus A) = 1 - P(A) \]

\[ P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \]

\[ P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) \]

conditional probability:

\[ P(A|B) = \frac{P(A \cap B)}{P(B)} \]

\[ P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \]

\[ P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} \]
Example

two fair coins tosses

\[ \Omega = \{ \text{HH, HT, TH, TT} \} \]

\[ A = \varnothing(\Omega) \]

\[ A = \{ \text{HH} \} \text{ two heads} \]

\[ P(A) = \frac{1}{4} \]

\[ B = \{ \text{HH, HT} \} \text{ first is head} \]

\[ P(B) = \frac{1}{2} \]

\[ P(A \cap B) = P(A) = \frac{1}{4} \]

\[ P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2} \]
Example

two fair coins tosses \( \Omega = \{ \text{HH, HT, TH, TT} \} \)

\( A = \emptyset(\Omega) \)

\( A = \{ \text{HH} \} \) two heads \( P(A) = \frac{1}{4} \)

\( B = \{ \text{HH, HT} \} \) first is head

\( C = \{ \text{HH, HT, TH} \} \) there is one head \( P(C) = \frac{3}{4} \)

\( P(A \cap C) = P(A) = \frac{1}{4} \)

\( P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{1/4}{3/4} = \frac{1}{3} \)
Random variable

$(\Omega, \mathcal{A}, P)$ probability space

$X : \Omega \to \mathbb{R}$ (can just take discrete values)

$\forall x \in \mathbb{R}. \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{A}$

equivalently: $\forall x \in \mathbb{R}. \{\omega \in \Omega \mid X(\omega) > x\} \in \mathcal{A}$

for every $x$ we can assign a probability to the above sets

$P(X \leq x) = P(\{\omega \in \Omega \mid X(\omega) \leq x\})$
Example

$\Omega$ sequences of $n$ fair coin tosses

$X$ counts the number of head in a sequence

for $n = 2$

$X(\text{HH}) = 2$

$X(\text{HT}) = 1$

$X(\text{TH}) = 1$

$X(\text{TT}) = 0$

$P(X \leq 1) = P(\{\text{HT, TH, TT}\}) = \frac{3}{4}$
Stochastic processes and Markov chains
Stochastic process

a family of random variables indexed by $T$

$\{X_t\}_{t \in T}$

discrete $\mathbb{N}$

continuous $\mathbb{R}$

$\forall t \in T. \quad X_t : \Omega \rightarrow \mathbb{R}$

set of states
Stochastic process

\[ \{ X_t \}_{t \in T} \quad \forall t \in T. \ X_t : \Omega \to \mathbb{R} \]

discrete time

we focus on discrete processes
(set of states is finite or countable)

continuous time

we further assume states are positive natural numbers

\[ S = \{ X_t(\omega) | \omega \in \Omega \land t \in T \} = \{1, 2, \cdots, N\} \]

for some \( N \)

\[ X_t = i \]

“the stochastic process \( X \) is in state \( i \) at time \( t \)”
Markov chain

\((\Omega, \mathcal{A}, P)\) probability space \(\{X_t\}_{t \in T}\) stochastic process

\(\forall t_0 < t_1 < \cdots < t_n < t\) possible times
\(\forall x, x_0, x_1, \cdots x_n\) possible states

\[ P(X_t = x|X_{t_n} = x_n, \cdots, X_{t_0} = x_0) = P(X_t = x|X_{t_n} = x_n) \]

Markov property (memoryless)

furthermore, we only consider \textit{homogeneous} Markov chains

\[ P(X_t = x|X_{t_n} = x_n) = P(X_{t-t_n} = x|X_0 = x_n) \]

time independence
Discrete Time MC

$(\Omega, \mathcal{A}, P)$ probability space

$\{X_t\}_{t \in \mathbb{N}}$ homogeneous Markov chain

$$P(X_{n+1} = x | X_n = x_n, \ldots, X_0 = x_0) = P(X_1 = x | X_0 = x_n)$$

$P$ entirely determined by

$$a_{i,j} = P(X_1 = j | X_0 = i) \text{ for } i, j \in \{1, \ldots, N\}$$

called transition probabilities
Continuous Time MC

$(\Omega, \mathcal{A}, P)$ probability space \( \{X_t\}_{t \in \mathbb{R}} \) homogeneous Markov chain

\[
P(X_{t_n + \Delta t} = x | X_{t_n} = x_n, \cdots, X_{t_0} = x_0) = P(X_{\Delta t} = x | X_0 = x_n)
\]

\(P\) entirely determined by the rates \(\lambda_{i,j}\) that govern

\[
P(X_t = j | X_0 = i) = 1 - e^{-\lambda_{i,j} t}
\]

(the exponential distribution is the only memoryless one)
(homogeneous) DTMC
DTMC as matrices

\((\Omega, \mathcal{A}, P)\) probability space \(\{X_t\}_{t \in \mathbb{N}}\) homogeneous Markov chain

\[ a_{i, j} = P(X_1 = j | X_0 = i) \]

\[ P = \begin{bmatrix}
    a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\
    a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{N,1} & a_{N,2} & \cdots & a_{N,N}
\end{bmatrix} \]

\(\forall i, j \in \{1, \ldots, N\}. \ 0 \leq a_{i, j} \leq 1\)

\(\forall i \in \{1, \ldots, N\}. \ \sum_{j=1}^{N} a_{i, j} = 1\)
Example

\[ P = \begin{bmatrix} \frac{4}{5} & \frac{1}{5} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ 1 & 0 & 0 \end{bmatrix} \]

\[
\frac{4}{5} + \frac{1}{5} + 0 = 1 \\
0 + \frac{1}{3} + \frac{2}{3} = 1 \\
1 + 0 + 0 = 1
\]
**DTMC as LTS**

$$P = \begin{bmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\
a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
a_{N,1} & a_{N,2} & \cdots & a_{N,N}
\end{bmatrix}$$

states $S = \{1, \ldots, N\}$, set of labels $[0, 1]$, transitions $i \xrightarrow{a_{i,j}} j$

for each state, the sum of the labels of outgoing arcs is equal to 1

also called *probabilistic transition systems*

$$\begin{array}{ccc}
i & \xleftrightarrow{p} & j \\
\downarrow q & & \downarrow p+q
\end{array}$$

same as

$$i \xrightarrow{p+q} j$$
Example

$$P = \begin{bmatrix}
\frac{4}{5} & \frac{1}{5} & 0 \\
0 & \frac{1}{3} & \frac{2}{3} \\
1 & 0 & 0
\end{bmatrix}$$

Fig. 14.5: A DTMC.
Probabilistic transition systems

transition function:

\[ \alpha_D : S \rightarrow \mathbb{D}(S) \]

set of discrete probabilistic distributions over \( S \)

\[ \mathbb{D}(S) = \left\{ d \mid d : S \rightarrow [0, 1], \sum_{s \in S} d(s) = 1 \right\} \]

more generally, we can allow for deadlock states

\[ \alpha_D : S \rightarrow \mathbb{D}(S) \cup \{ \star \} \]
DTMC: uncertainty

the state of the system at time $t$ is uncertain
we can estimate the likeliness of being in a certain state
the state of the DTMC at time $t$ is a probability distribution

$$\pi(t) = [\pi_1(t), \pi_2(t), \cdots, \pi_N(t)]$$

- probability of being in state $N$ at time $t$
- probability of being in state $2$ at time $t$
- probability of being in state $1$ at time $t$
Example

\[
P = \begin{bmatrix}
  4/5 & 1/5 & 0 \\
  0 & 1/3 & 2/3 \\
  1 & 0 & 0
\end{bmatrix}
\]

\[
\pi^{(0)} = [ 1, 0, 0 ]
\]
the system starts at state 1

\[
\pi^{(0)} = [ 1/2, 0, 1/2 ]
\]
initial states 1 and 3, equally likely

\[
\pi^{(0)} = [ 1/4, 1/2, 1/4 ]
\]
initial state 2 more likely than 1, 3
Computing with uncertainty

if we know the distribution at time $t$

we can estimate the distribution at time $t + 1$

the probability to be in state $k$ at time $t + 1$ is the sum of the probabilities to be in each state $i$ at time $t$ and move to $k$

$$\pi^{(t+1)}_k = \sum_{i=1}^{N} \pi^{(t)}_i \cdot a_{i,k}$$

$$\pi^{(t+1)} = \pi^{(t)} \cdot P$$
Computing with uncertainty

if we know the distribution at time $t$
we can estimate the distribution at time $t+1$

the probability to be in state $k$ at time $t+1$ is the sum
of the probabilities to be in each state $i$ at time $t$ and move to $k$

\[
\pi_k^{(t+1)} = \sum_{i=1}^{N} \pi_i^{(t)} \cdot a_{i,k}
\]

\[
\pi^{(t+1)} = \pi^{(t)} \cdot P
\]
Computing with uncertainty

if we know the distribution at time $t$

we can estimate the distribution at time $t+1$

the probability to be in state $k$ at time $t+1$ is the sum

of the probabilities to be in each state $i$ at time $t$ and move to $k$

$$\pi^{(t+1)} = \pi^{(t)} \cdot P$$

given the initial distribution we can compute the one at time $t$

$$\pi^{(t)} = \pi^{(0)} \cdot P^t$$
Example

\[
\begin{pmatrix}
\frac{4}{5} & \frac{1}{5} & 0 \\
0 & \frac{1}{3} & \frac{2}{3} \\
1 & 0 & 0
\end{pmatrix}
\]

\[
\pi^{(t)} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]
\]

\[
\pi^{(t+1)} = \pi^{(t)} \cdot P = \begin{bmatrix}
\left(\frac{1}{3} \cdot \frac{4}{5} + \frac{1}{3}\right), \\
\left(\frac{1}{3} \cdot \frac{1}{5} + \frac{1}{3} \cdot \frac{1}{3}\right), \\
\left(\frac{1}{3} \cdot \frac{2}{3}\right)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
3 & 8 & 2 \\
\frac{8}{5} & \frac{45}{9}
\end{bmatrix}
\]
Exercise

A printing device has three states: working, faulty, cleaning. When it is working it remains in state working with probability 1/2 and changes state to faulty or cleaning with equal probability. Similarly, when it is cleaning it remains in state cleaning with probability 1/2 and changes state to faulty or working with equal probability. When it is faulty it remains faulty with probability 1/3 or otherwise enters the cleaning state.

1. Represent the system as a DTMC.

\[
P = \begin{bmatrix}
W & F & C \\
W & F & C \\
W & F & C \\
\end{bmatrix}
\]
Exercise

A printing device has three states: working, faulty, cleaning. When it is working it remains in state working with probability 1/2 and changes state to faulty or cleaning with equal probability. Similarly, when it is cleaning it remains in state cleaning with probability 1/2 and changes state to faulty or working with equal probability. When it is faulty it remains faulty with probability 1/3 or otherwise enters the cleaning state.

1. Represent the system as a DTMC.

\[
P = \begin{bmatrix}
W & F & C \\
1/2 & 1/4 & 1/4 \\
0 & 1/3 & 2/3 \\
1/4 & 1/4 & 1/2 \\
\end{bmatrix}
\]

\[
W \quad F \quad C
\]
Finite path probability

let $s_1 s_2 \cdots s_n$ be the states traversed along a finite path on the LTS of a DTMC

i.e. $\forall i \in \{1, \ldots, n - 1\}$ we have $s_i \stackrel{a_{s_i, s_{i+1}}}{\rightarrow} s_{i+1}$

what is the probability of choosing that path?

$$P(s_1 s_2 \cdots s_n) = \prod_{i=1}^{n-1} a_{s_i, s_{i+1}}$$
Now we can calculate the state distribution at time \( t \) for some special classes of DTMCs we can prove the existence of a limit vector that is to say the probability that the system is found in a particular state is stationary in the long run (see Section 14.4.3).

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.

Example 14.5

Let us consider the DTMC in Figure 14.5. We represent the chain algebraically by using the following matrix:

\[
\mathbf{P} = \begin{bmatrix}
\frac{4}{5} & \frac{1}{5} & 0 \\
0 & \frac{1}{3} & \frac{2}{3} \\
1 & 0 & 0
\end{bmatrix}
\]

For some special classes of DTMCs we can prove the existence of a limit vector.

Example 14.5

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.

Example 14.5

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.

Example 14.5

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.

Example 14.5

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.

Example 14.5

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.

Example 14.5

Let \( \mathbf{p} \) be the state probability vector at time \( t \) defined as follows:

\[
\mathbf{p} = \mathbf{p}^0 \times \mathbf{P}^t
\]

where \( \mathbf{p}^0 \) is the state probability vector at time 0.
Example

\[
P = \begin{bmatrix}
1/2 & 1/4 & 1/4 \\
0 & 1/3 & 2/3 \\
1/4 & 1/4 & 1/2
\end{bmatrix}
\]

\[
P(W F C W F) = a_{W,F} \cdot a_{F,C} \cdot a_{C,W} \cdot a_{W,F} = \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{96}
\]
Ergodic DTMC
Steady state distribution

if we let a DTMC work for long enough
can we estimate what is the probability
to find the system in a given state?

\[ \pi_i = \lim_{t \to \infty} \pi_i^{(t)} \]

does it depend on the initial distribution?

\[ \pi^{(t)} = \pi^{(0)} \cdot P^t \]

if the limit exists, it should give a \textit{stationary distribution}

\[ \pi = [\pi_1, \cdots, \pi_n] \]

\[ \pi = \pi \cdot P \]

\[ \sum_{i=1}^{N} \pi_i = 1 \]
Ergodic DTMC

it the DTMC is *ergodic*

the stationary distribution exists

it is unique

it is independent from the initial state distribution

**Ergodic DTMC**

- irreducible: each state is reachable from any other state
  (there is a path between any two nodes in the LTS)

- aperiodic: for any state, the gcd of the lengths of all paths from the state to itself is 1
  (e.g. it is enough to have a self-loop)
Ergodic DTMC: steady state distribution

If the DTMC is ergodic

how to compute the steady state distribution?

take the unique solution of the system of linear equations

\[
\begin{align*}
\pi &= \pi \cdot P \\
\sum_{i=1}^{N} \pi_i &= 1
\end{align*}
\]
Example

\[
P = \begin{bmatrix}
4/5 & 1/5 & 0 \\
0 & 1/3 & 2/3 \\
1 & 0 & 0
\end{bmatrix}
\]

\[
\begin{aligned}
\frac{4}{5} \pi_1 + \pi_3 &= \pi_1 \\
\frac{1}{5} \pi_1 + \frac{1}{3} \pi_2 &= \pi_2 \\
\frac{2}{3} \pi_2 &= \pi_3 \\
\pi_1 + \pi_2 + \pi_3 &= 1
\end{aligned}
\]

\[
\pi = [2/3, 1/5, 2/15]
\]
Example

\[
P = \begin{bmatrix}
W & F & C \\
1/2 & 1/4 & 1/4 \\
0 & 1/3 & 2/3 \\
1/4 & 1/4 & 1/2
\end{bmatrix}
\]

\[
\begin{align*}
\frac{1}{2} \pi W + \frac{1}{4} \pi C &= \pi W \\
\frac{1}{4} \pi W + \frac{1}{3} \pi F + \frac{1}{4} \pi C &= \pi F \\
\frac{1}{4} \pi W + \frac{2}{3} \pi F + \frac{1}{2} \pi C &= \pi C \\
\pi W + \pi F + \pi C &= 1
\end{align*}
\]

\[
\pi = [ \frac{8}{33}, \frac{3}{11}, \frac{16}{33} ]
\]