Model Checking

Discrete-Time Markov Chains

Prof. Dr. Nils Jansen
Radboud University, Nijmegen, 2024/2025

Based on Slides by Dave Parker and Ralf Wimmer

Introduction

What are the coming weeks about? Probabilities!

(x_R, y_R, x_O, y_O)

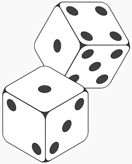
$$(0,0,1,1) \xrightarrow{\text{right}} \begin{cases} \xrightarrow{0.1} (1,0,1,0) \quad \text{bad} \\ \xrightarrow{0.5} (1,0,0,1) \end{cases}$$

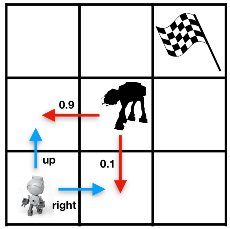
$$(0,0,1,1) \xrightarrow{\text{up}} \begin{cases} \xrightarrow{0.5} (0,1,0,1) \quad \text{bad} \\ \xrightarrow{0.1} (0,1,1,0) \end{cases}$$

bad: $x_R = x_O \wedge y_O = y_R$

$Pr(F \text{ "bad"}) = 0.1$

$\Diamond \qquad \leq$





Why Probabilities in Model Checking?

- Analyzing system performance and dependability
  - to quantify arrivals, waiting times, time between failures, QoS, ...
- Modeling unreliable and unpredictable system behavior
  - to capture machine learning models
  - to quantify message loss, processor failure
  - to quantify unpredictable delays, express soft deadlines, ...
- Building protocols for networked embedded systems
  - randomized algorithms often much simpler than deterministic ones

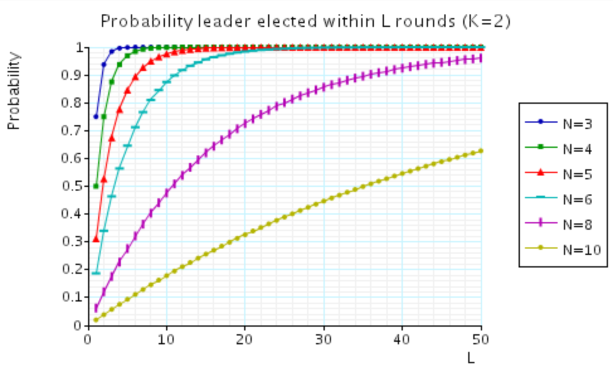
Observation
Answer "correct" / "erroneous" often not sufficient!
We need quantitative information about the system.

Example: Leader election

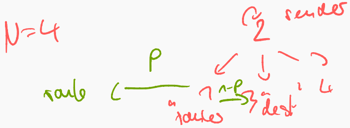
Distributed system: Leader election

- **System**:
  - Synchronous ring of $N > 2$ identical nodes
  - Task: select a leader node
- **Protocol**:
  - Each round starts by each node randomly choosing a number from $\{1, \ldots, K\}$ (uniformly distributed).
  - Nodes pass their selected number around the ring.
  - If there is a unique number, the node with the *maximal unique number* is leader.
  - Otherwise start a new round.
- **Desirable properties**:
  - Almost surely eventually a leader will be elected: $\mathrm{P}_{=1}(\mathrm{F}\,\text{leader elected})$
  - With probability at least 0.8, a leader is elected within k steps: $\mathrm{P}_{\geq 0.8}(\mathrm{F}^{\leq k}\,\text{leader elected})$.
  - The probability that node i becomes leader is $\frac{1}{N}$ for all $1 \leq i \leq N$.

Example: Leader election



Example: Crowds protocol



Security: Crowds protocol

- A protocol for anonymous web browsing [Reiter & Rubin, 1998]
- 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
    - with probability p: select next router randomly (uniformly)
  - Once a routing path has been established, use it until crowd changes
- Rebuild routing paths on crowd changes
- c of N crowd members are corrupt and try to identify sender
- Property: Crowds protocol ensures "probable innocence":
  - probability that real sender is discovered is $< \frac{1}{2}$ if $N \geq \frac{p}{p-0.5} \cdot (c+1)$.

Further examples

Examples: Real-world protocols featuring randomization

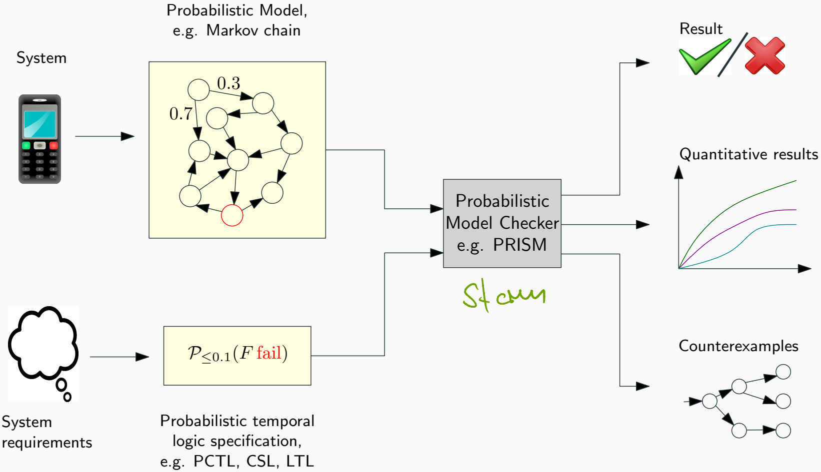
- Randomized back-off schemes:
  - IEEE 802.3 CSMA/CD, IEEE 802.11 Wireless LAN
- Random choice of waiting time
  - IEEE 1394 Firewire (root contention), Bluetooth (device discovery)
- Random choice over a set of possible addresses
  - IPv4 Zeroconf dynamic configuration (link-local addressing)

`https://qcomp.org/benchmarks/`

`https://www.prismmodelchecker.org/benchmarks/`

Probabilistic Model Checking: An Overview

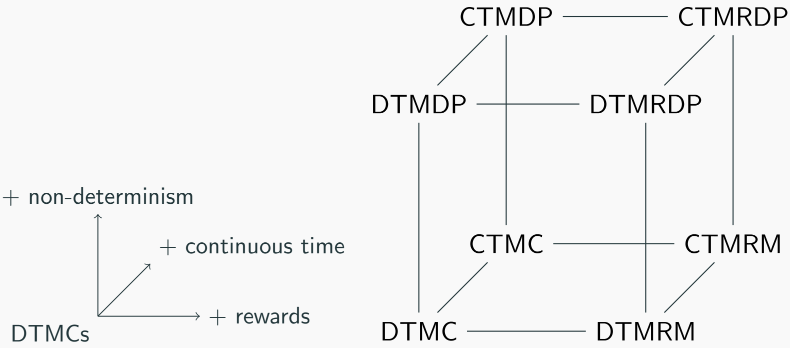
The Model Checking Flow



Probabilistic model checking inputs

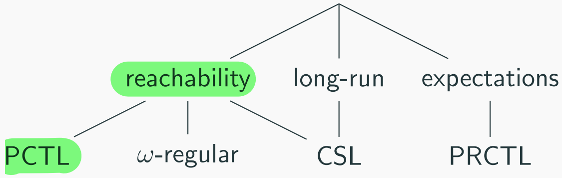
- Models: Variants of Markov chains
  - discrete-time Markov chains (DTMCs)
  - continuous-time Markov chains (CTMCs)
  - Markov decision processes (MDPs)
    = DTMCs + non-determinism
  - Markov reward models (MRMs)
    = DTMCs + rewards/costs
  - Partially observable MDPs
    = MDPs where a state is not fully observable
- Specifications
  - Informally:
    - "probability of delivery within time deadline is ..."
    - "expected time until delivery is ..."
    - "expected power consumption is ..."
  - Formally:
    - probabilistic temporal logics (PCTL, CSL, LTL, PCTL*)
    - e.g. $P_{<0.05}(\mathrm{F}\,\mathrm{critical})$, $P_{=?}(\neg\mathrm{warning}\ \mathrm{U}\ \mathrm{msg\_received})$

The probabilistic model space



DTMC	=	Discrete-time Markov chain
DTMRM	=	Discrete-time Markov reward model
DTMDP	=	Discrete-time Markov decision process
DTMRDP	=	Discrete-time Markow reward decision process
CTMC	=	Continuous-time Markov chain
CTMRM	=	Continuous-time Markov reward model
CTMDP	=	Continuous-time Markov decision process
CTMRDP	=	Continuous-time Markow reward decision process

Property space



Probabilistic model checking involves . . .

- Construction of models
  from a description in a high-level language

- Probabilistic model checking algorithms
  - graph-theoretical algorithms
    - for reachability, identifying strongly connected components, . . .
  - numerical computation
    - linear equation systems, linear optimization problems
    - iterative methods, direct methods
    - uniformization, shortest path problems
  - automata for regular languages
  - sampling-based methods for approximate analysis

- Efficient implementation techniques
  - essential for scalability to real-life applications
  - symbolic data structures based on BDDs
  - algorithms for model minimization, abstraction, . . .

Goals

Lecture:

- Introduce main types of probabilistic models and specification notations
- Algorithms for probabilistic model checking

Exercises:

- Deepening the understanding of the theoretical part
- Working with software tools (PRISM, Storm)
- Prototypic implementation
- Theoretical problems

Measurable space

Dice: $\Omega = \{1, 2, \ldots, 6\}$

σ-Algebra: $\{ \emptyset, \Omega, \{1\}, \{2\}, \ldots, \{6\}, \{2, \ldots, 6\}, \{1, 3, \ldots, 6\}, \{1, 2\}, \{1, 2, 3\}$ $\ldots = 2^\Omega$

Sample space

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

σ-Algebra

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

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \quad \Rightarrow \quad \Omega \setminus A \in \mathcal{F}$ complement
3. $(\forall i \geq 0 : A_i \in \mathcal{F}) \quad \Rightarrow \quad \bigcup_{i=0}^{\infty} A_i \in \mathcal{F}$. 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*.

If Ω is a set, $\mathcal{F} = \{\emptyset, \Omega\}$ yields the smallest, $\mathcal{F} = 2^\Omega$ the largest σ-algebra.

Probability space

$Pr(\{1\}) = \frac{1}{6}$

$Pr(\{1,3\}) = Pr(\{1\}) + Pr(\{3\}) = \frac{2}{6} = \frac{1}{3}$

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 $A_i \cap A_j = \emptyset$ 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

$Pr(\{2, \ldots, 6\}) = 1 - Pr(\{1\}) = 5/6$

$Pr(\{1,2\} \cup \{1\}) = \frac{2}{6} + \frac{1}{6} - \frac{1}{6} = 5/6$

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

- $\Pr(A) = 1 - \Pr(\Omega \setminus A)$
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
- $A \subseteq B$ implies $\Pr(A) \leq \Pr(B)$
- $\Pr(\bigcup_{n \geq 1} A_n) = \sum_{n \geq 1} \Pr(A_n)$ provided that 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 all $a \in A$:

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

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

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

Examples

Discrete

- throwing a dice
- number of customers in a shop
- drawn numbers in the Lotto game

Continuous

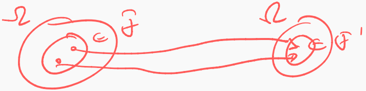
- Weight of a baby at birth
- Time until a system fails
- Throwing a dart on a circular board

Probability spaces: Example 1

- Sample space:
  - $\Omega = \{1, 2, 3\}$
- Event set Σ:
  - e. g., power set of Ω
  - $\Sigma = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\big\}$
  - closed under complement/(countable) union, contains $\emptyset$
- Probability measure $\Pr$:
  - e. g., $\Pr(1) = \Pr(2) = \Pr(3) = \frac{1}{3}$
  - $\Pr(\{1, 2\}) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$, etc.

Probability spaces: Example 2

- Sample space:
  - $\Omega = \{0, 1, 2, 3, \ldots\} = \mathbb{N}$
- Event set Σ:
  - e. g., $\Sigma = \{\emptyset, \text{“odd”}, \text{“even”}, \mathbb{N}\}$
  - closed under complement/(countable) union, contains $\emptyset$
- Probability measure Pr:
  - e. g., $\Pr(\text{“odd”}) = \Pr(\text{“even”}) = \frac{1}{2}$



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

$$\forall A \in \mathcal{F}' : f^{-1}(A) = \{a \in \Omega \mid f(a) \in A\} \in \mathcal{F}$$

Random variable
A 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 $(\Omega, \mathcal{F})$.

We omit the subscript X in $\Pr_X$ when clear from context. We consider only discrete random variables.

$x_1 \to x_2 \to x_4$

$\downarrow x_3$

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

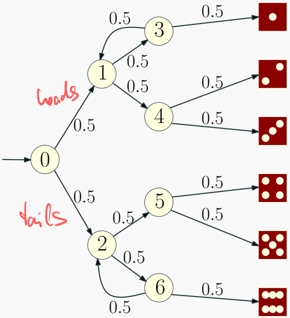
Bad Religion

Examples of stochastic processes

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

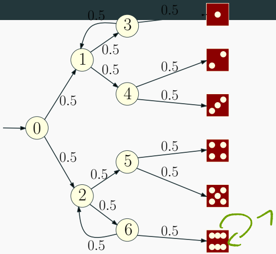
Probability example (1)

- Modelling a 6-sided dice using a fair coin
  - Algorithm due to Knuth/Yao:
  - Start at 0, toss a coin
  - upper branch when "*H*"
  - lower branch when "*T*"
  - repeat until value chosen



- Is this algorithm correct?
  - e. g. probability of obtaining "4"
  - Obtained as disjoint union of events
  - *THH*, *TTTHH*, *TTTTTHH*, ...
  - $\Pr(\text{"eventually 4"}) = (1/2)^3 + (1/2)^5 + (1/2)^7 + \cdots = 1/6$

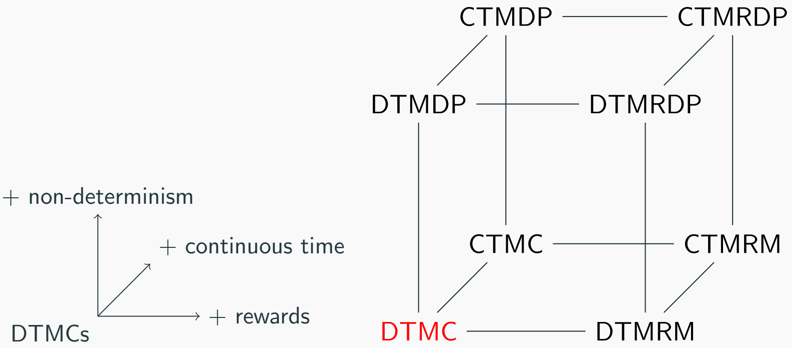
Probability example (2)



- Other properties?
  - “what is the probability of termination?”
  - efficiency:
    - What is the probability that more than four tosses are needed?
    - On average, how many tosses are needed?

- Probabilistic model checking provides a framework for these kinds of properties . . .
  - modeling languages
  - property specification languages
  - model checking algorithms, techniques, and tools

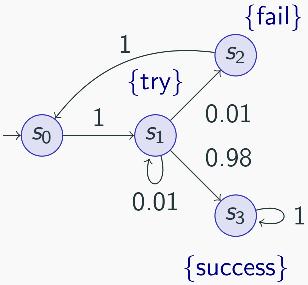
The probabilistic model space



DTMC	=	Discrete-time Markov chain
DTMRM	=	Discrete-time Markov reward model
DTMDP	=	Discrete-time Markov decision process
DTMRDP	=	Discrete-time Markow reward decision process
CTMC	=	Continuous-time Markov chain
CTMRM	=	Continuous-time Markov reward model
CTMDP	=	Continuous-time Markov decision process
CTMRDP	=	Continuous-time Markov reward decision process

Discrete-time Markov chains

- States
  - set of states representing possible configurations of the system being modelled
- Transitions
  - transitions between states model evolution of the system's state; occur in discrete time steps.
- Probabilities
  - probabilities of making transitions between states are given by discrete probability distributions.



Markov property

If the current state is known, the future states are independent of the past states.

- The current state contains all information that can influence the future evolution of the system.
- We do not need to store the history, i. e. the way how the current state was reached.
- This property is also known as "memorylessness".

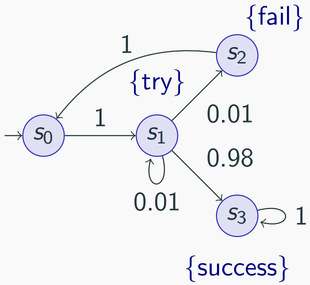
$$\Pr(X_{n+1} = s_{n+1} \mid X_n = s_n, X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) \\ = \Pr(X_{n+1} = s_{n+1} \mid X_n = s_n)$$

Bad Religion

Simple DTMC example

Modelling a very simple communication protocol

- After one step, process starts trying to send a message
- With probability 0.01, the channel is not ready. So wait a step.
- With probability 0.01, message sending fails. Restart.
- With probability 0.98, message is sent successfully. Stop.



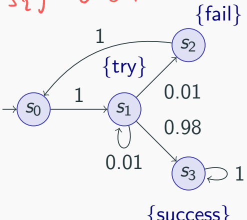
Formal definition of DTMCs

A discrete-time Markov chain (DTMC) is a tuple $M = (S, s_{\text{init}}, P, L)$ where

- S is a finite or countably infinite set of states ("state space").
- $s_{\text{init}} \in S$ is the initial state.
- $P : S \times S \to [0, 1]$ is the *transition probability matrix* such that $\sum_{s' \in S} P(s, s') = 1$ for all $s \in S$.
- $L : S \to 2^{\text{AP}}$ is a function labeling states with atomic propositions (taken from a set AP).

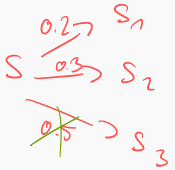
$$\begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0.01 & 0.01 & 0.98 \\ 2 & 1 & \text{—} & \text{—} & \text{—} \\ 3 & \text{—} & \text{—} & \text{—} & 1 \end{array}$$

$P(s_1, s_2) = 0.01$



Some more terminology

- Matrix $A : S \times S \to \mathbb{R}$ is stochastic if
  - $A(s, s') \in [0, 1]$ for all $s, s' \in S$ and
  - $\sum_{s' \in S} A(s, s') = 1$ for all $s \in S$.



- Matrix $A : S \times S \to \mathbb{R}$ is sub-stochastic if
  - $A(s, s') \in [0, 1]$ for all $s, s' \in S$ and
  - $\sum_{s' \in S} A(s, s') \leq 1$ for all $s \in S$.

- State $s \in S$ is absorbing if
  - $P(s, s) = 1$ and $P(s, s') = 0$ for all $s' \in S \setminus \{s\}$.

  The transition for s to itself is called a self-loop.

Some more terminology

- Matrix $A : S \times S \to \mathbb{R}$ is stochastic if
  - $A(s, s') \in [0, 1]$ for all $s, s' \in S$ and
  - $\sum_{s' \in S} A(s, s') = 1$ for all $s \in S$.

- Matrix $A : S \times S \to \mathbb{R}$ is sub-stochastic if
  - $A(s, s') \in [0, 1]$ for all $s, s' \in S$ and
  - $\sum_{s' \in S} A(s, s') \leq 1$ for all $s \in S$.

- State $s \in S$ is absorbing if
  - $P(s, s) = 1$ and $P(s, s') = 0$ for all $s' \in S \setminus \{s\}$.

  The transition for s to itself is called a self-loop.

We assume that P is stochastic, i. e.

- every state has at least one outgoing transition
- there are no deadlocks.

Other assumptions made here

$\rightarrow s_0$

$s_{init}(s_0) = 0.5$

$s_{init}(s_0') = 0.2$

$s_{init}(s_0'') = 0.3$

- Finite state space
  In general: arbitrary countable set

s_0

s_0'

s_0''

- Single initial state
  In general: initial probability distribution:

$$s_{\text{init}} : S \to [0, 1]$$

$s_I \xrightarrow{0.5} s_0$

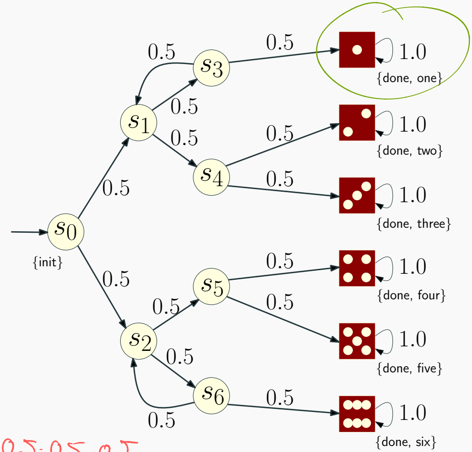
$s_I \xrightarrow{0.2} s_0'$

$s_I \xrightarrow{0.3} s_0''$

- Rational transition probabilities for algorithmic purposes (finite representation . . .).
  In general: real numbers.

Example: Coins and dice

- Recall Knuth/Yao's dice algorithm from earlier:



$0.5 \cdot 0.5 \cdot 0.5$

$0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 \sim 0.5 \cdot (0.5)^i \ldots$

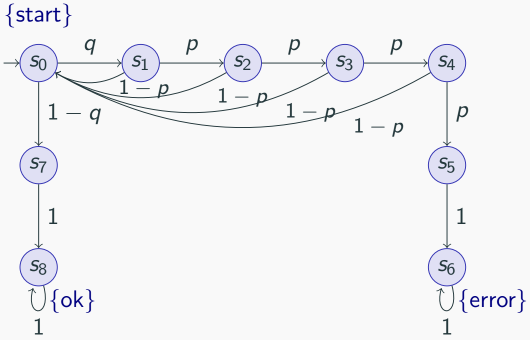
- $S = \{s_0, s_1, \ldots, s_6, 1, 2, \ldots, 6\}$
- $s_{\text{init}} = s_0$
- $P(s_0, s_1) = 0.5$
  $P(s_0, s_2) = 0.5$
  etc.
- $L(s_0) = \{\text{init}\}$
  $L(s_1) = \emptyset$
  etc.

Example: Zeroconf protocol

- Zeroconf = "Zero configuration networking"
  - self-configuration for local ad-hoc networks
  - automatic configuration of unique IP address for new devices
  - simple, no DHCP, DNS, ...
- Basic idea:
  - 65 024 available IP addresses (IANA-specified range)
  - new node picks address U at random
  - broadcasts "probe" messages: "Who is using U?"
  - a node already using U replies to the probe
  - in this case, protocol is restarted
  - messages may not get sent (transmission fails, host busy, ...)
  - so: node sends multiple (n) probes, waiting after each one.

DTMC for Zeroconf

- $n = 4$ probes, m existing nodes in the network
- probability of message loss: p
- probability that new address is in use: $q = m/65\,024$



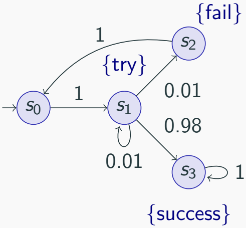
Properties of DTMCs

- Path-based properties
  - What is the probability of observing a particular behavior (or class of behaviors)?
  - e.g.: What is the probability of running into a safety-critical state without issuing a warning before?
- Transient properties
  - What is the probability of being in state s after k steps?
- Steady-state properties
  - What is the probability to be in an failure state on the long run?
- Expectations
  - What is the average number of coin tosses required?

DTMCs and paths

A path in a DTMC represents an execution (i. e., one possible behavior) of the system being modelled.

- Formally:
  - finite sequence of states $s_0 s_1 s_2 \ldots s_n$ such that $P(s_i, s_{i+1}) > 0$ for all $0 \leq i < n$, or
  - infinite sequence of states $s_0 s_1 s_2 \ldots$ such that $P(s_i, s_{i+1}) > 0$ for all $i \geq 0$.
- Examples:
  - never succeeds: $(s_0 s_1 s_2)^\omega$
  - tries, waits, fails, retries, succeeds: $s_0 s_1 s_1 s_2 s_0 s_1 (s_3)^\omega$
- Notation:
  - finite paths starting in state s: $\text{Paths}_{\text{fin}}(s)$
  - infinite paths starting in state s: $\text{Paths}_{\text{inf}}(s)$

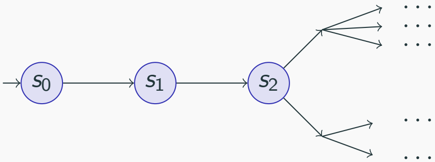


Paths and probabilities

- To reason (quantitatively) about this system, we need to define a probability space over paths.

- Intuitively:
  - sample space: $\text{Paths}_{\text{inf}}(s)$ = set of all infinite paths starting in s.
  - events: sets of infinite paths from s
  - basic events: cylinder sets (or "cones")
  - cylinder set $\text{Cyl}(\omega)$ for a finite path $\omega \in \text{Paths}_{\text{fin}}(s)$ = set of the infinite paths with common prefix ω

- Example: Cylinder set $\text{Cyl}(s_0 s_1 s_2)$:



Probability spaces (refresher 1)

- Ω: arbitrary non-empty set
- A σ-algebra on Ω is a family Σ of subsets of Ω which is closed under complementation and countable union, i. e.,
  - if $A \in \Sigma$, then $\Omega \setminus A \in \Sigma$,
  - if $A_i \in \Sigma$ for $i \in \mathbb{N}$, then $\bigcup_{i \in \mathbb{N}} A_i \in \Sigma$, and
  - $\emptyset \in \Sigma$.
- Elements of Σ: measurable sets or events.

- Theorem:
  For any family F of subsets of Ω there exists a unique smallest σ-algebra on Ω containg F.

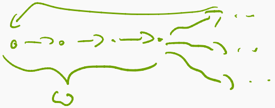
Probability spaces (refresher 2)

Probability space $(\Omega, \Sigma, \Pr)$ such that

- Ω is the sample space
- Σ is a σ-algebra, the set of events
- $\Pr : \Sigma \to [0, 1]$ is the probability measure:
    - $\Pr(\emptyset) = 0$
    - $\Pr(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \Pr(A_i)$ for pairwise disjoint A_i.

Probability space over paths

- Sample space Ω_s:
  - $\Omega_s = \mathsf{Paths}_{\mathsf{inf}}(s)$ = set of infinite paths starting in s
- Event set: Σ_s
  - Cylinder set $\mathsf{Cyl}(\omega) = \{\omega' \in \mathsf{Paths}_{\mathsf{inf}}(s) \mid \omega \text{ is prefix of } \omega'\}$ for $\omega \in \mathsf{Paths}_{\mathsf{fin}}(s)$
  - Σ_s is the least σ-algebra on $\mathsf{Paths}_{\mathsf{inf}}(s)$ containing $\mathsf{Cyl}(\omega)$ for $\omega \in \mathsf{Paths}_{\mathsf{fin}}(s)$.



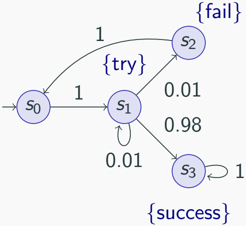
- Probability measure $\Pr_s$:
  - define $P_s(\omega)$ for finite path $\omega = s_0 s_1 \ldots s_n$ with $s_0 = s$ by

$$P_s(\omega) = \prod_{i=0}^{n-1} P(s_i, s_{i+1}).$$

  - define $\Pr_s(\mathsf{Cyl}(\omega)) = P_s(\omega)$.
  - $\Pr_s$ extends uniquely to a probability measure $\Pr_s : \Sigma_s \to [0, 1]$

Paths and probabilities: Example

- Paths where sending fails the first time:
  - $\omega = s_0 s_1 s_2$
  - $\mathrm{Cyl}(\omega)$ = all paths starting with $s_0 s_1 s_2 \ldots$
  - $P_{s_0}(\omega) = P(s_0, s_1) \cdot P(s_1, s_2)$
    $= 1 \cdot 0.01 = 0.01$
  - $\mathrm{Pr}_{s_0}(\mathrm{Cyl}(\omega)) = 0.01$



- Paths which are eventually successful and with no failure
  - $\mathrm{Cyl}(s_0 s_1 s_3) \cup \mathrm{Cyl}(s_0 s_1 s_1 s_3) \cup \mathrm{Cyl}(s_0 s_1 s_1 s_1 s_3) \cup \ldots$
  - $\mathrm{Pr}_{s_0}(\mathrm{Cyl}(s_0 s_1 s_3) \cup \mathrm{Cyl}(s_0 s_1 s_1 s_3) \cup \cdots)$
    $= P_{s_0}(s_0 s_1 s_3) + P_{s_0}(s_0 s_1 s_1 s_3) + \cdots$
    $= 1 \cdot 0.98 + 1 \cdot \underbrace{0.1}_{0} \cdot 0.98 + 1 \cdot \underbrace{0.1^2}_{0} \cdot 0.98 + \cdot 1 \cdot \underbrace{0.1^3}_{0} \cdot 0.98 + \cdots = \frac{98}{99}.$

Reachability

$s_0 \to s_1 \to s_2 \rightsquigarrow s_i \in T \; \cancel{\rightsquigarrow}$

- Key property: reachability
  - probability of a path reaching a state in some target set $T \subseteq S$
  - e. g., "probability of the algorithm terminating successfully?"
  - e. g., "probability that an error occurs during execution?"

- Dual of reachability: invariance
  - probability of remaining within some class of states
  - $\Pr($"remain in set of states T"$) = 1 - \Pr($"reach set $S \setminus T$"$)$
  - e. g., "probability that an error never occurs"

- Variants of reachability
  - time-bounded, constrained ("until"), ...

Rechability probabilities

$s \rightsquigarrow T$

- Formally:
  - $\mathrm{PrReach}(s, T) = \mathrm{Pr}_s(\mathrm{Reach}(s, T))$
  - $\mathrm{Reach}(s, T) = \{s_0 s_1 s_2 \ldots \in \mathrm{Paths}_{\mathrm{inf}}(s) \mid s_i \in T \text{ for some } i \in \mathbb{N}\}$

- Is $\mathrm{Reach}(s, T)$ measurable for arbitrary $T \subseteq S$? Yes ...
  - $\mathrm{Reach}(s, T)$ is the union of all basic cylinders $\mathrm{Cyl}(s_0 s_1 \ldots s_n)$ where $s_0 s_1 \ldots s_n \in \mathrm{Reach}_{\mathrm{fin}}(s, T)$
  - $\mathrm{Reach}_{\mathrm{fin}}(s, T)$ contains all finite paths $s_0 s_1 \ldots s_n$ such that $s_0 = s$, $s_0, \ldots, s_{n-1} \notin T$, $s_n \in T$.
  - The set of such paths $s_0 s_1 \ldots s_n$ is countable.

$s_0 \to s_i \to T$

- Probability
  - The above is a disjoint union
  - so probability is obtained by simply summing ...

s_0, s_1, T

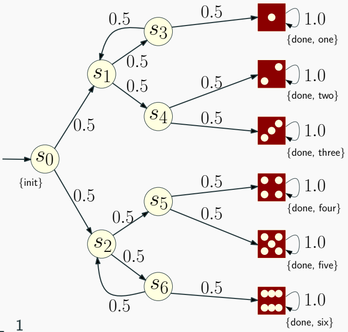
s_0, s_1, s_1, T

Computing reachability probabilities

- Compute as infinite sum ...
- $$\sum_{\omega \in \mathrm{Reach}_{\mathrm{fin}}(s,T)} \mathrm{Pr}_s(\mathrm{Cyl}(\omega)) = \sum_{\omega \in \mathrm{Reach}_{\mathrm{fin}}(s,T)} P_s(\omega)$$
- Example:
  - $\mathrm{PrReach}(s_0, \{4\}) = \sum_{i=0}^{\infty} 0.5 \cdot (0.5 \cdot 0.5)^i \cdot 0.5 \cdot 0.5 =$

Computing reachability probabilities

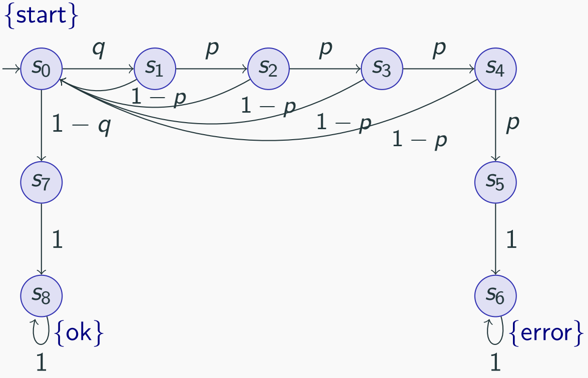
- Compute as infinite sum ...
- $$\sum_{\omega \in \mathrm{Reach}_{\mathrm{fin}}(s,T)} \mathrm{Pr}_s(\mathrm{Cyl}(\omega)) = \sum_{\omega \in \mathrm{Reach}_{\mathrm{fin}}(s,T)} P_s(\omega)$$
- Example:
  - $\mathrm{PrReach}(s_0, \{4\}) = \sum_{i=0}^{\infty} 0.5 \cdot (0.5 \cdot 0.5)^i \cdot 0.5 \cdot 0.5 = \frac{1}{8} \cdot \frac{1}{1-\frac{1}{4}} = \frac{1}{6}$



$$\sum_{i=0}^{\infty} p^i = \frac{1}{1-p}$$

Computing reachability probabilities

- PrReach($s_0, \{s_6\}$): compute as infinite sum?
  - Doesn't scale!!



Computing reachability probabilities

- Alternative: derive a linear equation system
  - solve for all states simultaneously
  - i. e., compute vector PrReach(T)

$$\begin{pmatrix} s_0 \\ \vdots \\ s_n \end{pmatrix} \rightsquigarrow T$$

- Let x_s denote PrReach(s, T)

x_s: $s \to s'$ ($x_{s'}$) $\to s''$; $s \to s''''$ ($x_{s''''}$) $\xrightarrow{1} s_t \in T$, $x_{s_t} = 1$; $s \to s'''$ ($x_{s'''}$)

- Solve:

$$x_s = \begin{cases} 1 & \text{if } s \in T, \\ 0 & \text{if } T \text{ is not reachable from } s, \\ \sum_{s' \in S} P(s,s') \cdot x_{s'} & \text{otherwise.} \end{cases}$$

Example

- Compute PrReach($s_0, \{4\}$)

  $x_1 = x_2 = x_3 = x_5 = x_6 = 0$

  $x_{s_3} = x_{s_4} = x_{s_1} = 0$

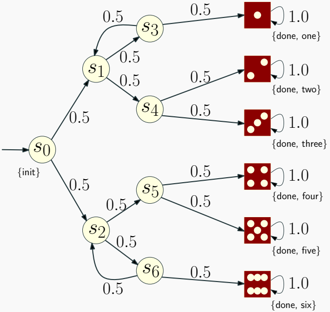
  $x_4 = 1$

  $x_{s_0} = 0.5\overset{=0}{x_{s_1}} + 0.5x_{s_2}$

  $x_{s_2} = 0.5x_{s_5} + 0.5x_{s_6}$

  $x_{s_5} = 0.5\overset{=1}{x_4} + 0.5x_5$

  $x_{s_6} = 0.5x_6 + 0.5x_{s_2}$



- Simplification:

  $x_4 = 1$

  $x_{s_5} = 0.5$

  $x_{s_2} = 0.25 + 0.5x_{s_6}$

  $x_{s_6} = 0.5x_{s_2}$

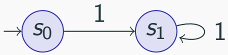
  $x_{s_0} = 0.5x_{s_2}$

- Solution $x_{s_0} = \frac{1}{6}$.

Unique solutions

- Why do we need to identify states that can reach T?

- Consider this simple DTMC:



- Compute probability of reaching s_0 from s_1.
  - Linear equation system:

    $x_{s_0} = 1$

    $x_{s_1} = x_{s_1}$
  - Solutions: $(x_{s_0}, x_{s_1}) = (1, p)$ for any $p \in [0, 1]$.

Bounded reachability

$$\in T$$

$$s_0 \to s_1 \to \dots s_i$$

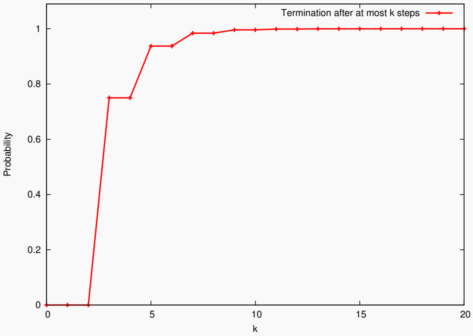
$$i \leq k$$

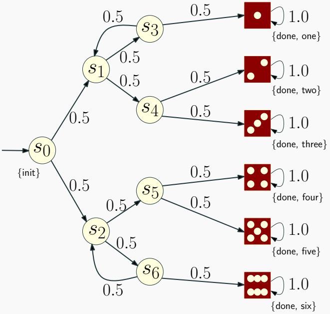
- Probability of reaching T from s within k steps
- Formally:
    - $\mathrm{PrReach}^{\leq k}(s, T) = \mathrm{Pr}_s(\mathrm{Reach}^{\leq k}(s, T))$ where
    - $\mathrm{Reach}^{\leq k} = \{s_0 s_1 s_2 \ldots \in \mathrm{Paths}_{\mathrm{inf}}(s) \mid s_i \in T \text{ for some } i \leq k\}$.

$$\mathrm{PrReach}^{\leq k}(s, T) = \begin{cases} 1 & \text{if } s \in T, \\ 0 & \text{if } k = 0 \text{ and } s \notin T, \\ \sum_{s' \in S} P(s, s') \cdot \mathrm{PrReach}^{\leq k-1}(s', T) & \text{if } k > 0 \text{ and } s \notin T. \end{cases}$$

(Bounded) reachability: Example

- $\mathrm{PrReach}(s_0, \{1,2,3,4,5,6\}) = 1$
- $\mathrm{PrReach}^{\leq k}(s_0, \{1,2,3,4,5,6\}) = \cdots$





Summing up so far ...

- Discrete-time Markov chains (DTMCs)
  - state-transition systems augmented with probabilities
- Formalizing path-based properties of DTMCs
  - probability space over infinite paths
- Probabilistic reachability
  - infinite sum
  - linear equation system
  - least fixed point characterization
  - bounded reachability

If Time Permits: NO

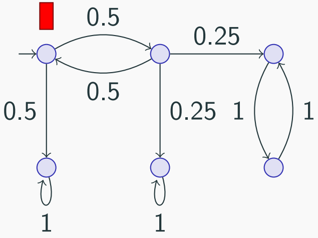
- Transient properties
  - What is the probability of being in state s after k steps?
- Steady-state properties
  - What is the probability to be in a failure state on the long run?
- Expectations
  - What is the average number of coin tosses required?

Transient state probabilities

- What is the probability, having started in state s, of being in state s' at time k?
  - i. e., after exactly k steps/transition have occured
  - transient state probability: $\pi_{s,k}(s')$

- This is a discrete probability distribution
  - we have $\pi_{s,k} : S \to [0,1]$
  - rather than $\Pr_s : \Sigma_s \to [0,1]$ where $\Sigma_s \subseteq 2^{\mathrm{Paths}_{\mathrm{inf}}(s)}$

$k = 0$:

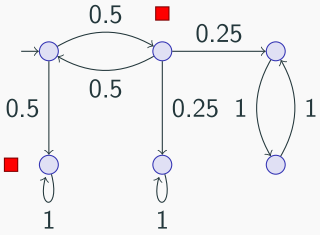


Transient state probabilities

- What is the probability, having started in state s, of being in state s' at time k?
  - i. e., after exactly k steps/transition have occured
  - transient state probability: $\pi_{s,k}(s')$

- This is a discrete probability distribution
  - we have $\pi_{s,k} : S \to [0,1]$
  - rather than $\mathrm{Pr}_s : \Sigma_s \to [0,1]$ where $\Sigma_s \subseteq 2^{\mathrm{Paths}_{\mathrm{inf}}(s)}$

$k = 1$:

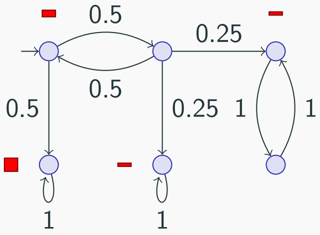


Transient state probabilities

- What is the probability, having started in state s, of being in state s' at time k?
    - i. e., after exactly k steps/transition have occured
    - transient state probability: $\pi_{s,k}(s')$

- This is a discrete probability distribution
    - we have $\pi_{s,k} : S \to [0,1]$
    - rather than $\Pr_s : \Sigma_s \to [0,1]$ where $\Sigma_s \subseteq 2^{\mathrm{Paths}_{\mathrm{inf}}(s)}$

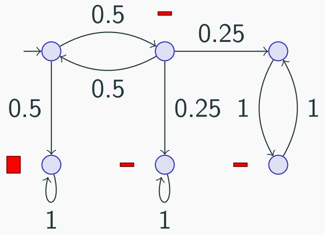
$k = 2$:



Transient state probabilities

- What is the probability, having started in state s, of being in state s' at time k?
  - i. e., after exactly k steps/transition have occured
  - transient state probability: $\pi_{s,k}(s')$
- This is a discrete probability distribution
  - we have $\pi_{s,k} : S \to [0,1]$
  - rather than $\Pr_s : \Sigma_s \to [0,1]$ where $\Sigma_s \subseteq 2^{\mathrm{Paths}_{\mathrm{inf}}(s)}$

$k = 3$:



Computing transient probabilities

- Transient state probabilities:

$$\pi_{s,k}(s') = \sum_{s'' \in S} P(s'', s') \cdot \pi_{s,k-1}(s'')$$

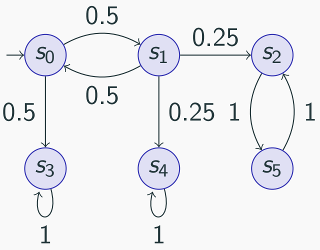
- Computation of transient state distribution:
  - $\pi_{s,0}$ is the initial probability distribution
  - e. g., in our case $\pi_{s,0}(s') = 1$ if $s' = s$ and $\pi_{s,0}(s') = 0$ otherwise.
  - $\pi_{s,k} = \pi_{s,k-1} \cdot P$
- $\Rightarrow$ successive vector-matrix multiplications

Computing transient probabilities

$$\pi_{s,k} = \pi_{s,k-1} \cdot P = \pi_{s,0} \cdot P^k$$

- k-th matrix power P^k
  - P gives one-step transition probabilities
  - P^k gives k-step transition probabilities
  - i. e., $P^k(s, s') = \pi_{s,k}(s')$
- A possible optimization: iterative squaring
  - e. g., $P^8 = ((P^2)^2)^2$
  - only requires $\log k$ multiplications
  - but potentially inefficient, e. g,. if P is large and sparse
  - in practice, successive vector-matrix multiplications preferred.

Example



$$\pi_{s_0,0} = \left(1, 0, 0, 0, 0, 0\right)$$

$$\pi_{s_0,1} = \left(0, \frac{1}{2}, 0, \frac{1}{2}, 0, 0\right)$$

$$\pi_{s_0,2} = \left(\frac{1}{4}, 0, \frac{1}{8}, \frac{1}{2}, \frac{1}{8}, 0\right)$$

$$\pi_{s_0,3} = \left(0, \frac{1}{8}, 0, \frac{5}{8}, \frac{1}{8}, \frac{1}{8}\right)$$

$$\cdots$$

$$P = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.25 & 0 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Notion of time in DTMCs

Two possible views on the timing aspects of a system modelled as a DTMC:

1. Discrete time-steps model time accurately
   - e. g., clock ticks in a model of an embedded device
   - or like dice example: interested in the number of steps (tosses)

2. Time-abstract
   - no information assumed about the time transitions take
   - e. g., Zeroconf protocol model

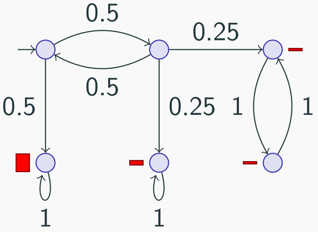
In the latter case, transient probabilities are not very useful!

$\Rightarrow$ Study long-run behavior

Long-run behavior

- Consider the limit $\pi_s := \lim_{k \to \infty} \pi_{s,k}$
  - $\pi_{s,k}$ is the transient state distribution at time k having started in state s
  - this limit, where it exists, is called the limiting distribution.
- Intuitive idea:
  - The percentage of time, in the long run, spent in each state.
  - e. g., availability: "In the long run, what percentage of time is the system in an operational state?"

Limiting distribution: Example



$$\pi_{s_0,0} = \Big(1, 0, 0, 0, 0, 0\Big)$$

$$\pi_{s_0,1} = \left(0, \frac{1}{2}, 0, \frac{1}{2}, 0, 0\right)$$

$$\pi_{s_0,2} = \left(\frac{1}{4}, 0, \frac{1}{8}, \frac{1}{2}, \frac{1}{8}, 0\right)$$

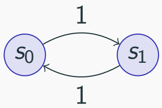
$$\pi_{s_0,3} = \left(0, \frac{1}{8}, 0, \frac{5}{8}, \frac{1}{8}, \frac{1}{8}\right)$$

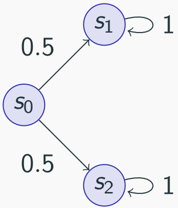
$$\dots$$

$$\pi_{s_0} = \left(0, 0, \frac{1}{12}, \frac{2}{3}, \frac{1}{6}, \frac{1}{12}\right)$$

Long-run behavior

- Questions:
    - When does the limiting distribution exist?
    - Does it depend on the initial state?
    - How to efficiently compute it?



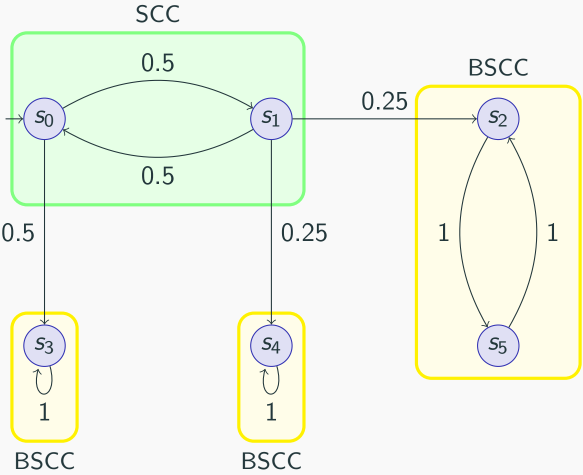


- We need to consider the underlying graph
    - (V, E) where V are vertices and $E \subseteq V \times V$ are edges
    - $V = S$ and $E = \{(s, s') \in S \times S \mid P(s, s') > 0\}$

Graph terminology

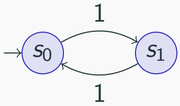
- A state s' is reachable from s if there is a finite path starting in s and ending in s'.
- A subset T of S is strongly connected if, for each pair of states s and s' in T, s' is reachable from s passing only through states in T.
- A strongly connected component (SCC) is a maximal strongly connected set of states (i. e., no proper superset of it is also strongly connected)
- A bottom strongly connected component (BSCC) is an SCC T from which no state outside T is reachable from T.

BSCCs: Example



Graph terminology

- A DTMC is irreducible if all its states belong to a single BSCC; otherwise reducible.



- A state s is periodic with period d, if
    - the greatest common divisor of the set $\{n \mid f_s^{(n)} > 0\}$ equals d,
    - where $f_s^{(n)}$ is the probability of, when starting in state s, returning to state s in exactly n steps.
    - A DTMC is aperiodic if its period is 1.

Steady-state probabilities

- For a finite, irreducible, aperiodic DTMC ...
  - the limiting distribution always exists
  - and is independent of the initial state/distribution.

- These are known as steady-state probabilities
- They can be computed as the unique solution of the linear equation system

$$\pi \cdot P = \pi$$
$$\sum_{s \in S} \pi(s) = 1.$$

Qualitative properties

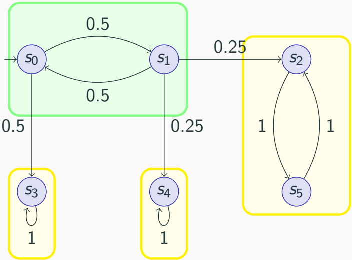
- Quantitative properties:
  - "What is the probability of event A?"

- Qualitative properties:
  - "Is the probability of event $A = 1$?" ("almost surely A")
  - "Is the probability of event $A > 0$?" ("possibly A")

For finite DTMCs, qualitative properties do not depend on the transition probabilities – only need the underlying graph.

Fundamental property of BSCCs

With probability 1, a BSCC will be reached and all of its states visited infinitely often.

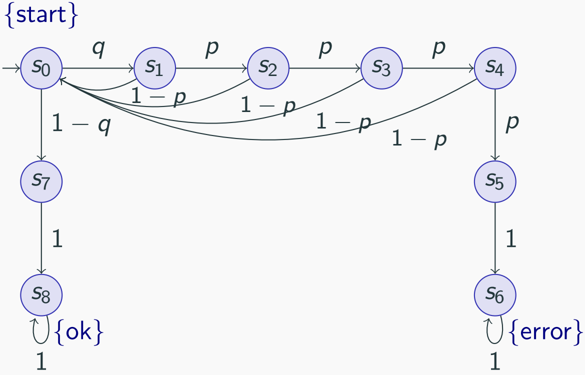


Formally:

$$\begin{aligned}\Pr_{s_0}(s_0 s_1 s_2 \ldots \mid\ & \exists i \geq 0, \exists \text{ BSCC } T \text{ such that} \\ & \forall j \geq i : s_j \in T \text{ and} \\ & \forall s \in T : s_k = s \text{ for infinitely many } k) = 1.\end{aligned}$$

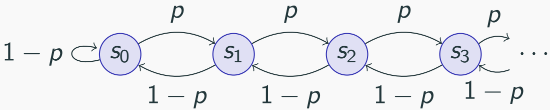
Zeroconf example

- 2 BSCCs: $\{s_6\}$ and $\{s_8\}$
- Probability of trying to acquire a new address infinitely often is 0



Remark: Infinite Markov chains

- Infinite-state random walk:



- Value of probability p does affect qualitative proerties:
  - RepeatedReachability$(s, \{s_0\}) = 1$ if $p \leq 0.5$
  - RepeatedReachability$(s, \{s_0\}) = 0$ if $p > 0.5$

Repeated reachability

- Repeated reachability:
    - "always eventually", "infinitely often"
- $\Pr_{s_0}(s_0 s_1 s_2 \ldots \mid \forall i \geq 0 \; \exists j \geq i : s_j \in B\}$
    - where $B \subseteq S$ is a set of states.
- e. g., "What is the probability that the protocol successfully sends a message infinitely often?"
- Is this measurable? Yes ...
    - set of satisfying paths is $\bigcap_{n \geq 0} \bigcup_{n \geq m} C_m$
    - where C_m is the union of all cylinder sets $\mathrm{Cyl}(s_0 s_1 \ldots s_m)$ for finite paths $s_1 s_1 \ldots s_m$ such that $s_m \in B$.

Qualitative repeated reachability

$$\Pr_{s_0}(s_0 s_1 s_2 \ldots \mid \forall i \geq 0 \exists j \geq i : s_j \in B) = 1$$

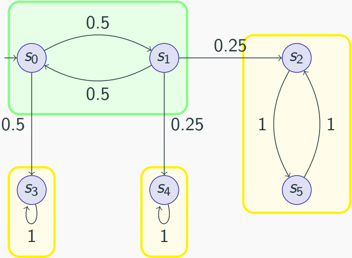
if and only if

$$\Pr_{s_0}(\text{“always eventually } B\text{”}) = 1$$

if and only if

$T \cap B \neq \emptyset$ for each BSCC T that is reachable from s_0.

Example:



$B = \{s_3, s_4, s_5\}$

Persistence

- Persistence properties:
  - "eventually forever"
- $\mathrm{Pr}_{s_0}(s_0 s_1 s_2 \ldots \mid \exists i \geq 0\ \forall j \geq i : s_j \in B)$
  - where $B \subseteq S$ is a set of states.
- Examples
  - "What is the probability of the leader election algorithm reaching, and staying in, a stable state?"
  - "What is the probability that an irrecoverable error occurs?"
- Is this measurable? Yes . . .

Qualitative persistence

$$\Pr_{s_0}(s_0 s_1 s_2 \ldots \mid \exists i \geq 0\ \forall j \geq i : s_j \in B) = 1$$

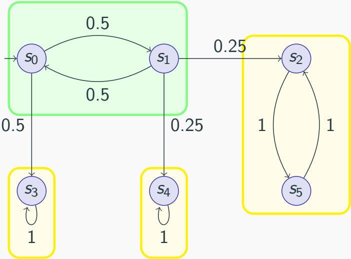
if and only if

$$\Pr_{s_0}(\text{“eventually forever } B\text{”}) = 1$$

if and only if

$T \subseteq B$ for each BSCC T that is reachable from s_0.

Example:



$$B = \{s_2, s_3, s_4, s_5\}$$

Summary DTMCs

- Basic Probability Theory
- Discrete-Time Markov Chains
- First Model Checking Approaches
- Now: Markov Decision Processes