Discrete-time Markov chains

Lecture #13 of Modeling Concurrent and Probabilistic Systems

Joost-Pieter Katoen and Thomas Noll

Lehrstuhl 2: Software Modeling and Verification

E-mail: `katoen@cs.rwth-aachen.de`

June 19, 2009

Overview Lecture #13

$\Rightarrow$ *Discrete-time Markov chains*

- What is a discrete-time Markov chain?
- Computing n-step transition probabilities
- Transient distribution
- Limiting and stationary distribution

Markov property

Stochastic process $\{ X(t) \mid t \in T \}$ is a *Markov process* if
for any $t_0 < t_1 < \ldots < t_n < t_{n+1}$:

$$\Pr\{ X(t_{n+1}) \leqslant d_{n+1} \mid X(t_0) = d_0, X(t_1) = d_1, \ldots, X(t_n) = d_n \}$$
$$=$$
$$\Pr\{ X(t_{n+1}) \leqslant d_{n+1} \mid X(t_n) = d_n \}$$

The distribution of $X(t_{n+1})$, given the values $X(t_0)$ through $X(t_n)$, *only depends on the current state* $X(t_n)$

$\Rightarrow$ the history has no influence on the future behaviour

Invariance to time-shifts

Markov process $\{ X(t) \mid t \in T \}$ is *time-homogeneous* iff for any $t' < t$:

$$\Pr\{ X(t) \leqslant d \mid X(t') = d' \} = \Pr\{ X(t - t') \leqslant d \mid X(0) = d' \}$$

the next state only depends on the current state, and
not on how long we have already been in that state

Discrete-time Markov chain

- A *time-homogeneous discrete-time Markov chain* (DTMC) is
  - a Markov process
  - with discrete parameter T and discrete state space $X(t)$
  - which is time-homogeneous

- $p_s(n) = \Pr\{\, X(n) = s \,\}$ probability to be in state s at step n

- Probability of being in state s' at step n when in s at step $m < n$:

$$\begin{aligned} p_{s,s'}(m,n) &= \Pr\{\, X(n) = s' \mid X(m) = s \,\} \\ &= \Pr\{\, X(n-m) = s' \mid X(0) = s \,\} \end{aligned}$$

  - these are the $(n-m)$-step *transition probabilities*

Stochastic matrix

- 1-step transition probabilities can be gathered in matrix $\mathbf{P}$
  - $\mathbf{P}(s, s') = p_{s,s'}(n+1, n) = p_{s,s'}(0, 1) = \Pr\{\, X(1) = s' \mid X(0) = s \,\}$
  - i.e., the probability to move from s to s' in a single step

- $\mathbf{P}$ is a *stochastic* matrix:
  - quadratic cardinality
  - $0 \leqslant \mathbf{P}(s, s') \leqslant 1$ for all states s, s'
  - $\sum_{s'} \mathbf{P}(s, s') = 1$ for any state s

- For stochastic matrix $\mathbf{P}$ it holds:
  - $\mathbf{P}^n$ is a stochastic matrix, for all n
  - $\mathbf{P}$ has Eigenvalue 1 and all Eigenvalues are at most 1

Another perspective

A *discrete-time Markov chain* (DTMC) is a tuple $(S, \mathbf{P})$ where:

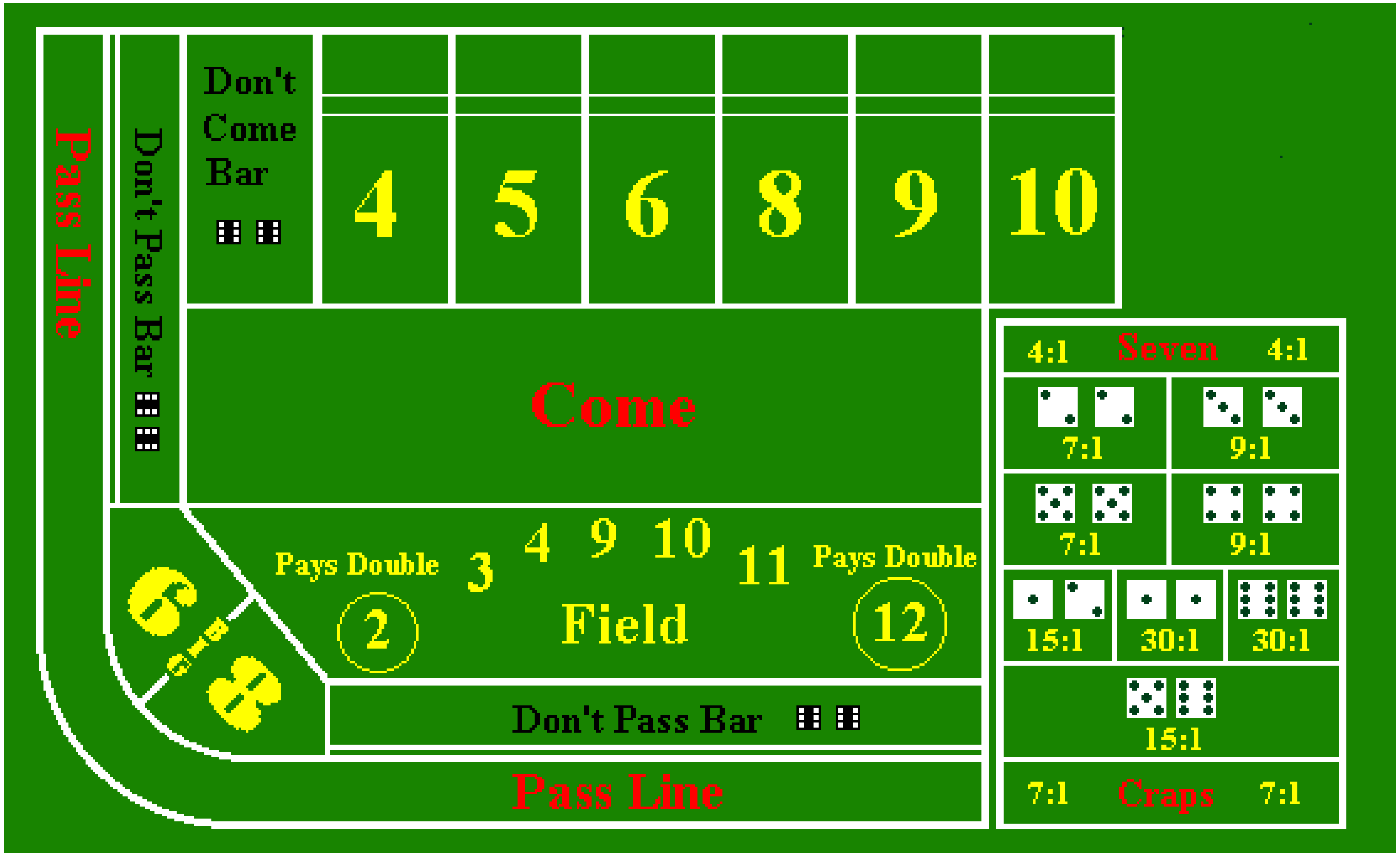
- S is a countable set of states
- $\mathbf{P} : S \times S \to [0, 1]$ is a probability matrix satisfying

$$\sum_{s' \in S} \mathbf{P}(s, s') = 1 \qquad \text{for all } s \in S$$

- (state s is *absorbing* whenever $\mathbf{P}(s, s) = 1$)

a DTMC is a transition system (unlabeled transitions)
where transitions are equipped with discrete probabilities

Craps


Pass Line
Don't Pass Bar
Don't Come Bar
4
5
6
8
9
10
Come
4:1 Seven 4:1
7:1
9:1
7:1
9:1
15:1
30:1
30:1
15:1
7:1 Craps 7:1
Pays Double
2
3 4 9 10 11
Field
Pays Double
12
6
8
Don't Pass Bar
Pass Line

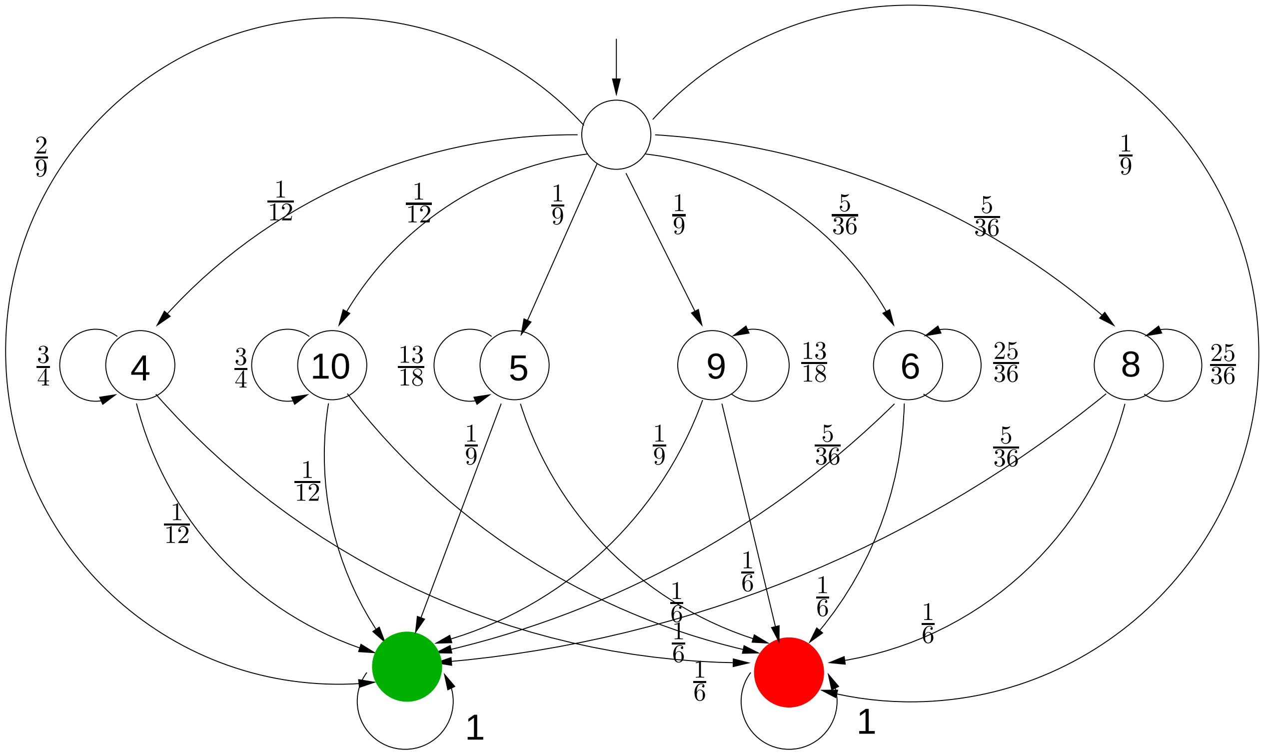
Craps

- Roll two dice and bet on outcome
- Come-out roll ("pass line" wager):
  - outcome 7 or 11: win
  - outcome 2, 3, or 12: loss ("craps")
  - any other outcome: roll again (outcome is "point")
- Repeat until 7 or the "point" is thrown:
  - outcome 7: loss ("seven-out")
  - outcome the point: win
  - any other outcome: roll again

A DTMC model of Craps

- Come-out roll:
  - 7 or 11: win
  - 2, 3, or 12: loss
  - else: roll again

- Next roll(s):
  - 7: loss
  - point: win
  - else: roll again



State residence time distribution

Let T_s be the remaining number of epochs to stay in state s:

$$
\begin{array}{rcl}
\Pr\{T_s = 1\} & = & 1 - \mathbf{P}(s,s) \\
\Pr\{T_s = 2\} & = & \mathbf{P}(s,s) \cdot (1 - \mathbf{P}(s,s)) \\
\Pr\{T_s = 3\} & = & \mathbf{P}(s,s)^2 \cdot (1 - \mathbf{P}(s,s)) \\
\cdots\cdots & \cdots & \cdots\cdots \\
\Pr\{T_s = n\} & = & \mathbf{P}(s,s)^{n-1} \cdot (1 - \mathbf{P}(s,s))
\end{array}
$$

$\Rightarrow$ the state residence times in a DTMC obey a *geometric* distribution

with expectation $E[T_s] = \frac{1}{1-\mathbf{P}(s,s)}$ and variance *Var*$[T_s] = \frac{\mathbf{P}(s,s)}{(1-\mathbf{P}(s,s))^2}$

This is not a surprise: the geometric distribution is the *only*

discrete probability distribution that exhibits the memoryless property

Initial distribution

- What is the state-distribution after n steps?
  - $\Pr\{ X(n) = s \}$ for each state s
- This can only be answered when the *initial distribution* is known!
- Starting distribution $\underline{p}(0) : S \rightarrow [0,1]$ such that $\underline{p}_s(0) = \Pr\{ X(0) = s \}$
  - comparable to initial states of transition systems
  - $\sum_{s \in S} \underline{p}_s(0) \; = \; 1$
- Example starting distributions: ......

Computing n-step transition probabilities

- The probability to move from s to s' in $n \geqslant 0$ steps:

$$p_{s,s'}(n) = \sum_{s''} p_{s,s''}(l) \cdot p_{s'',s'}(n-l) \quad \text{for all } 0 \leqslant l \leqslant n$$

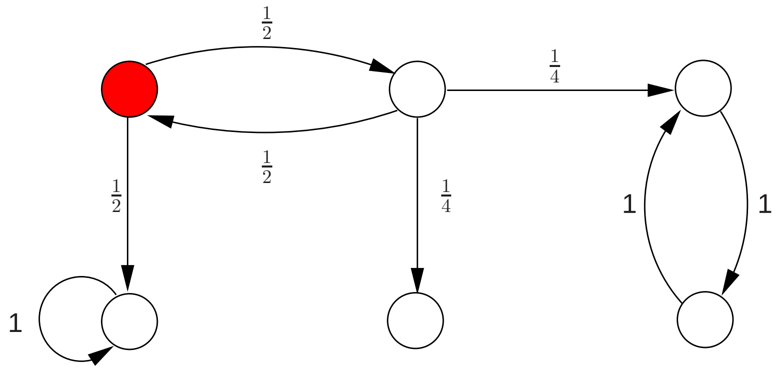
  - this is known as the *Chapman-Kolmogorov* equation

- For $l = 1$ and $n > 0$ we obtain: $p_{s,s'}(n) = \sum_{s''} p_{s,s''}(1) \cdot p_{s'',s'}(n-1)$
  - in matrix-form: $\mathbf{P}^{(n)} = \mathbf{P}^{(1)} \cdot \mathbf{P}^{(n-1)} = \mathbf{P} \cdot \mathbf{P}^{(n-1)}$
  - where $\mathbf{P}^{(n)}$ is the n-step transition probability matrix

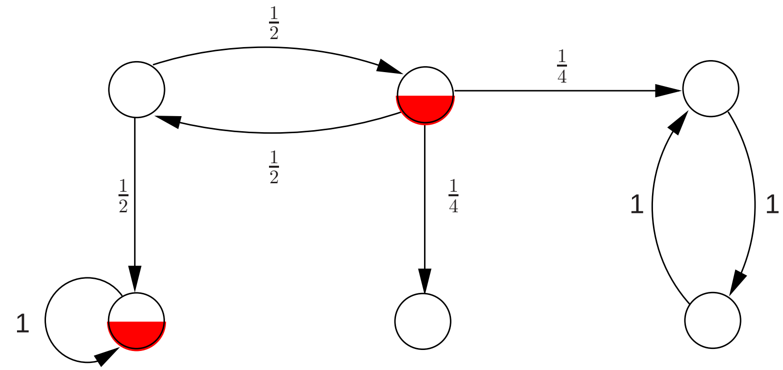
- Repeating this scheme: $\mathbf{P}^{(n)} = \mathbf{P} \cdot \mathbf{P}^{(n-1)} = \ldots = \mathbf{P}^{n-1} \cdot \mathbf{P}^{(1)} = \mathbf{P}^{n}$

Note: the difference between $\mathbf{P}^{n}$ and $\mathbf{P}^{(n)}$

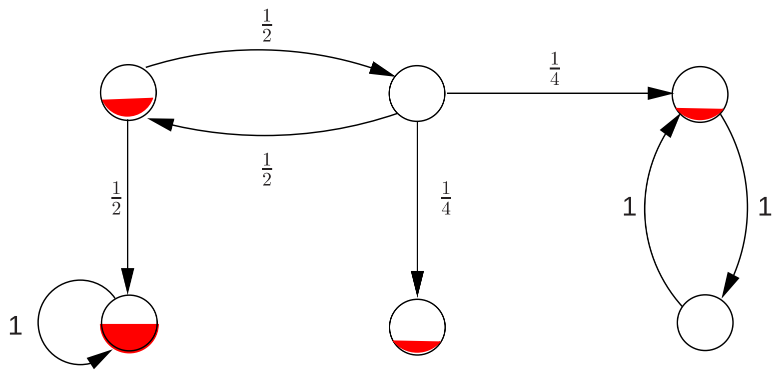
Evolution of an example DTMC



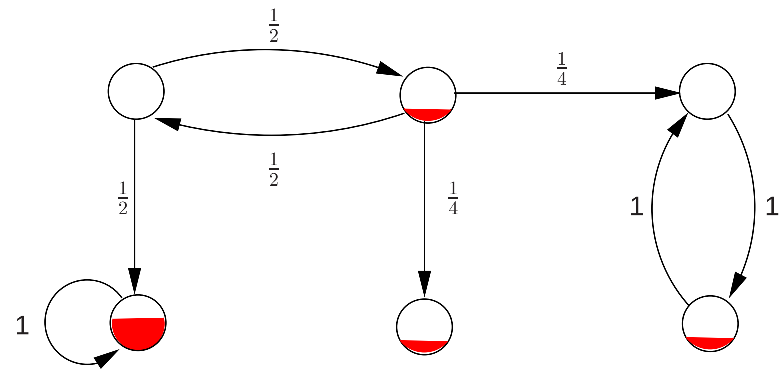
zero-th epoch



first epoch

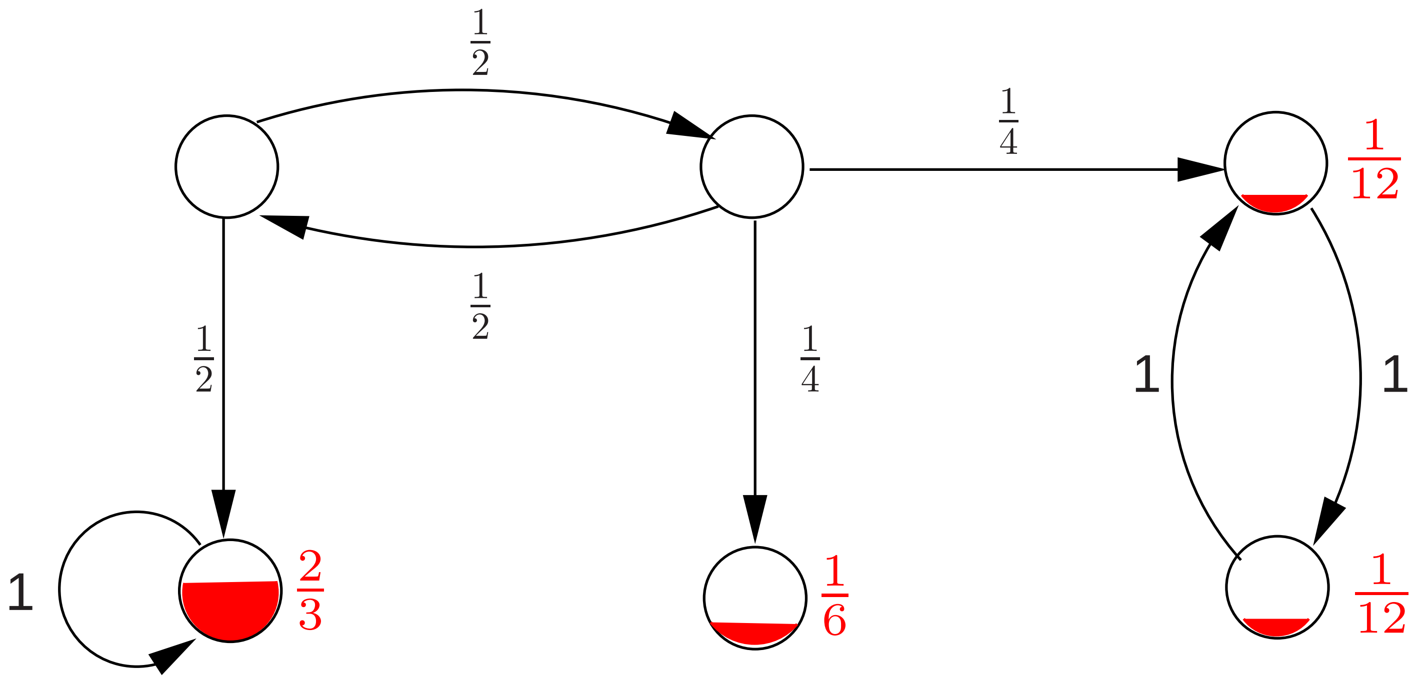


second epoch



third epoch

On the long run



Transient distribution of a DTMC

Probability to be in state s at step n:

$$
\begin{aligned}
p_s(n) &= \Pr\{\, X(n) = s \,\} \\
&= \sum_{s' \in S} \underbrace{\Pr\{\, X(0) = s' \,\}}_{p_{s'}(0)} \cdot \underbrace{\Pr\{\, X(n) = s \mid X(0) = s' \,\}}_{p_{s',s}(n)}
\end{aligned}
$$

Using $\underline{p}(n) = (p_{s_0}(n), p_{s_1}(n), \ldots, p_{s_k}(n))$ we obtain in matrix form:

$$\underline{p}(n) = \underline{p}(0) \cdot \mathbf{P}^n \quad \text{given} \quad \underline{p}(0)$$

where $\mathbf{P}^n$ is the n-step transition probability matrix

$\underline{p}(n)$ is called the *n-step transient-state* probability vector

Example

Limiting distribution

- Stochastic matrix $\mathbf{P}$ is called *ergodic* if:

$$\mathbf{P}^{\infty} = \lim_{n\to\infty} \mathbf{P}^n \quad \text{exists and has identical rows}$$

- *Theorem*: if the transition probability matrix $\mathbf{P}$ of a DTMC is ergodic:
  - $\underline{p}(n)$ converges to a limiting distribution $\underline{v}$ independent from $\underline{p}(0)$
  - and each row of $\mathbf{P}^{\infty}$ equals the limiting distribution

- This can be seen as follows:

$$\lim_{n\to\infty} \underline{p}(n) = \lim_{n\to\infty} \underline{p}(0) \cdot \mathbf{P}^n = \underline{p}(0) \cdot \underbrace{\lim_{n\to\infty} \mathbf{P}^n}_{\mathbf{P}^{\infty}} = \underline{p}(0) \cdot \begin{pmatrix} v_{s_0} & \cdots & v_{s_n} \\ \cdots & \cdots & \cdots \\ v_{s_0} & \cdots & v_{s_n} \end{pmatrix} = \underline{v}$$

Limiting distribution

- We also have:

$$\underline{v} = \lim_{n\to\infty} \underline{p}(n+1) = \lim_{n\to\infty} \underline{p}(0)\cdot\mathbf{P}^{n+1} = \left(\lim_{n\to\infty} \underline{p}(0)\cdot\mathbf{P}^{n}\right)\cdot\mathbf{P} = \underline{v}\cdot\mathbf{P}$$

- Thus, limiting probabilities can be obtained by solving the (homogeneous) system of linear equations:

$$\boxed{\underline{v} = \underline{v}\cdot\mathbf{P} \quad \text{or} \quad \underline{v}\cdot(\mathbf{I}-\mathbf{P}) = \underline{0} \quad \text{under } \textstyle\sum_i \underline{v}(i) = 1}$$

  - vector $\underline{v}$ is the left Eigenvector of $\mathbf{P}$ with Eigenvalue 1

- $\underline{v}$ is called the *limiting* state-probability vector

Examples

Limiting distribution

Two interpretations of $\underline{v}(s)$:

- the long-run proportion of time that the DTMC “spends” in state s
- the probability the DTMC is in s when making a snapshot after a very long time

Transient and recurrent states

- The probability of *first return* to state s after exactly n epochs is:

$$f_s(n) = \Pr\{ X(n) = s, X(n{-}1) \neq s, \ldots, X(1) \neq s \mid X(0) = s \}$$

  - not to be confused with $p_{s,s}(n) = \mathbf{P}^n(s, s)$
  - relationship: $p_{s,s}(n) = \sum_{k=1}^{n} f_s(k) \cdot p_{s,s}(n{-}k)$ for $n \geqslant 1$

- Probability to *eventually return* to s: $f_s = \sum_{n=1}^{\infty} f_s(n)$

  - state s is called *transient* if $f_s < 1$

  $\Rightarrow$ there is a non-zero probability that the DTMC will *not* return to s

  - state s is called *recurrent* if $f_s = 1$

  $\Rightarrow$ it is impossible to never come back to a recurrent state

A graph-theoretical characterization

Irreducibility and periodicity

- A DTMC is *irreducible* if its underlying digraph is strongly connected
  - otherwise it is called reducible
- The period $d(s)$ of recurrent state s is
  - greatest common divisor of $\{ n > 0 \mid p_{s,s}(n) > 0 \}$
  - if $d(s) = 1$, recurrent state s is called *aperiodic*
- Some facts:
  - all states in the same strongly component have the same period
  - an irreducible DTMC is aperiodic if all its states are aperiodic

Mean recurrence time

- For recurrent state s, the mean number of epochs between two successive visits to s:

$$m_s = \sum_{n=1}^{\infty} n \cdot f_s(n)$$

  - state s is called *positive recurrent* if $m_s < \infty$
  - state s is called *null recurrent* if $m_s = \infty$

- State s is *ergodic* if s is aperiodic *and* positive recurrent
  - a DTMC is called ergodic when all its states are ergodic

- Fact: in a *finite*, aperiodic and irreducible DTMC all states are ergodic
  - there are also infinite DTMCs that are ergodic

Connected states have the same "type"

Let s and s' be mutually reachable from each other. Then:

s is transient	iff	s' is transient
s is null-recurrent	iff	s' is null-recurrent
s is positive recurrent	iff	s' is positive recurrent
s has period d	iff	s' has period d

Stationary distribution

- $\underline{\pi}$ is the *stationary* distribution of DTMC with matrix $\mathbf{P}$ if: $\underline{\pi} = \underline{\pi} \cdot \mathbf{P}$

- in elementwise notation:

$$\underline{\pi}(s) = \sum_{s'} \underline{\pi}(s') \cdot \mathbf{P}(s', s)$$

- or equivalently:

$$\underbrace{\underline{\pi}(s) \cdot (1 - \mathbf{P}(s,s))}_{\text{the "outflux" of } s} = \underbrace{\sum_{s' \neq s} \underline{\pi}(s') \cdot \mathbf{P}(s', s)}_{\text{the "influx" of } s}$$

Stationary distribution

- An irreducible, positive recurrent DTMC has a *unique* stationary distribution $\underline{v}$:

$$\underline{v}(s) = \frac{1}{m_s}$$

- . . . but the limiting distribution does not need to exist
  - since the DTMC could be periodic!

Stationary versus limiting distribution

For ergodic DTMCs: the limiting and stationary distribution coincide