Lecture 3
Discrete-time Markov Chains

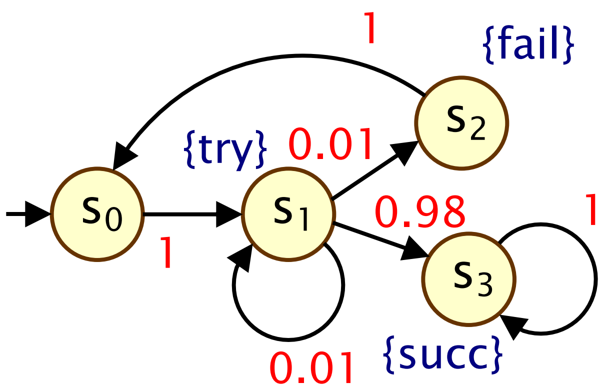
Alessandro Abate

Department of Computer Science
University of Oxford

Discrete-time Markov chains

- State-transition systems augmented with probabilities

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



Overview

- Previous lecture: path-based properties,
  probabilistic reachability
- Transient state probabilities
- Long-run / steady-state probabilities
- Qualitative properties
  - repeated reachability
  - persistence

Transient state probabilities

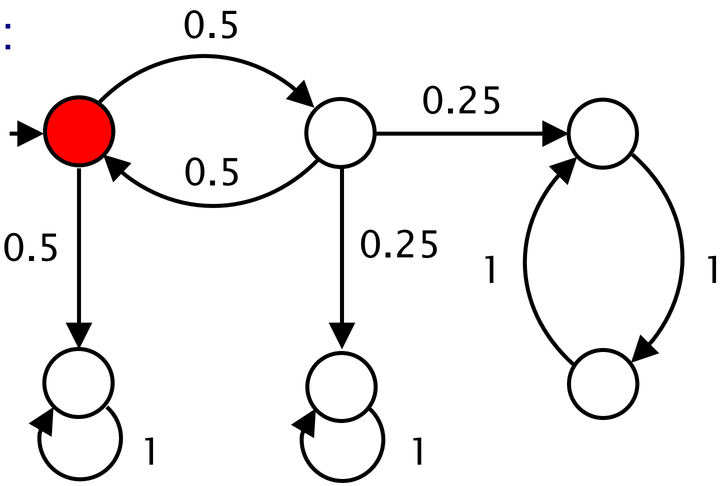
- What is the probability, having started in state s, of being in state s' at time k?
  - i.e. after exactly k steps/transitions have occurred
  - this is the transient state probability: $\pi_{s,k}(s')$

- Transient state distribution: $\underline{\pi}_{s,k}$
  - (row) vector $\underline{\pi}_{s,k}$ i.e. $\pi_{s,k}(s')$ for all states s'

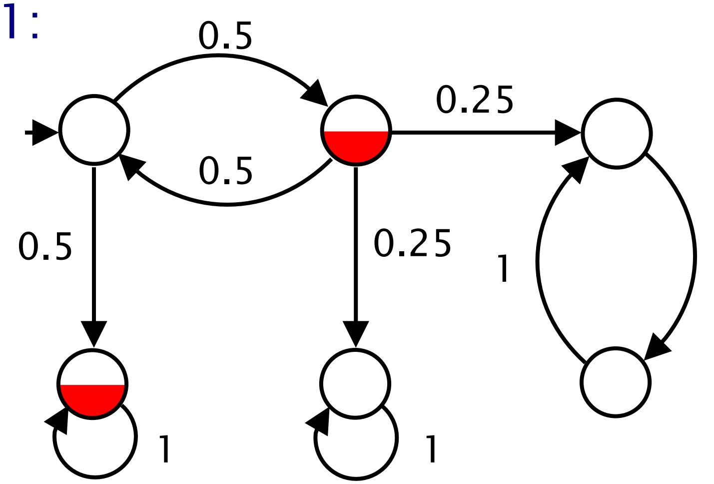
- Note: this is a discrete probability distribution
  - so we have $\underline{\pi}_{s,k} : S \rightarrow [0,1]$
  - recall instead $Pr_s : \Sigma_{Path(s)} \rightarrow [0,1]$, where $\Sigma_{Path(s)} \subseteq 2^{Path(s)}$

Transient distributions

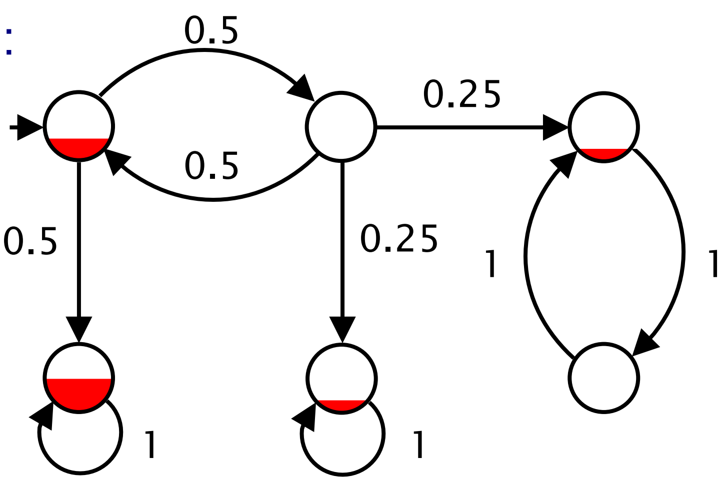
k=0:



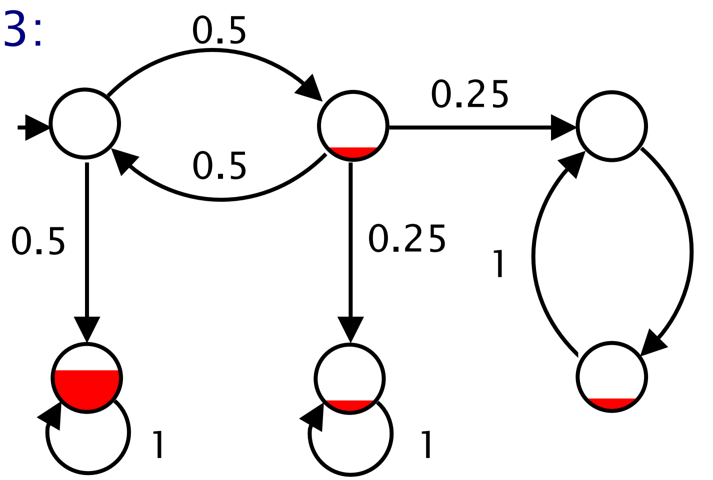
k=1:



k=2:



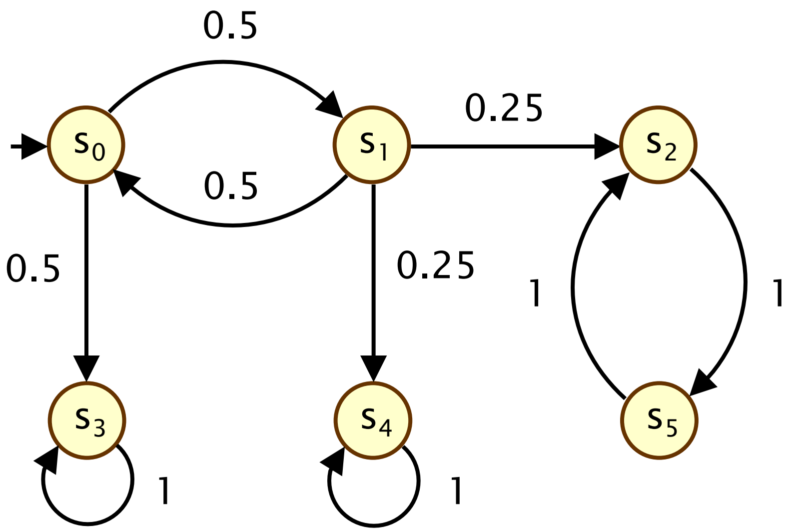
k=3:



Computing transient probabilities

- Transient state probabilities:
  - $\pi_{s,k}(s') = \Sigma_{s''\in S}\ \pi_{s,k-1}(s'') \cdot P(s'',s')$
  - (i.e. look at incoming transitions, into s')
- Computation of transient state distribution:
  - $\underline{\pi}_{s,0}$ is the initial probability distribution
  - e.g. in our case $\underline{\pi}_{s,0}(s') = 1$ if $s'=s$ and $\underline{\pi}_{s,0}(s') = 0$ otherwise
  - $\underline{\pi}_{s,k} = \underline{\pi}_{s,k-1} \cdot P$
- i.e. successive vector-matrix multiplications

Computing transient probabilities



$$\underline{\pi}_{s0,0} = \left[1,0,0,0,0,0\right]$$

$$\underline{\pi}_{s0,1} = \left[0,\tfrac{1}{2},0,\tfrac{1}{2},0,0\right]$$

$$\underline{\pi}_{s0,2} = \left[\tfrac{1}{4},0,\tfrac{1}{8},\tfrac{1}{2},\tfrac{1}{8},0\right]$$

$$\underline{\pi}_{s0,3} = \left[0,\tfrac{1}{8},0,\tfrac{5}{8},\tfrac{1}{8},\tfrac{1}{8}\right]$$

$\ldots$

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

Computing transient probabilities

- $\underline{\pi}_{s,k} = \underline{\pi}_{s,k-1} \cdot \mathbf{P} = \underline{\pi}_{s,0} \cdot \mathbf{P}^k$

- k^{th} matrix power: $\mathbf{P}^k$
  - $\mathbf{P}$ gives one-step transition probabilities
  - $\mathbf{P}^k$ gives probabilities of k-step transition probabilities
  - i.e. $\mathbf{P}^k(s,s') = \pi_{s,k}(s')$

- A possible optimisation: iterative squaring
  - e.g. $\mathbf{P}^8 = ((\mathbf{P}^2)^2)^2$
  - only requires log k multiplications
  - but potentially inefficient, e.g. if **P** is large and sparse

  - in practice, successive vector-matrix multiplications preferred

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 number of steps (tosses)

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

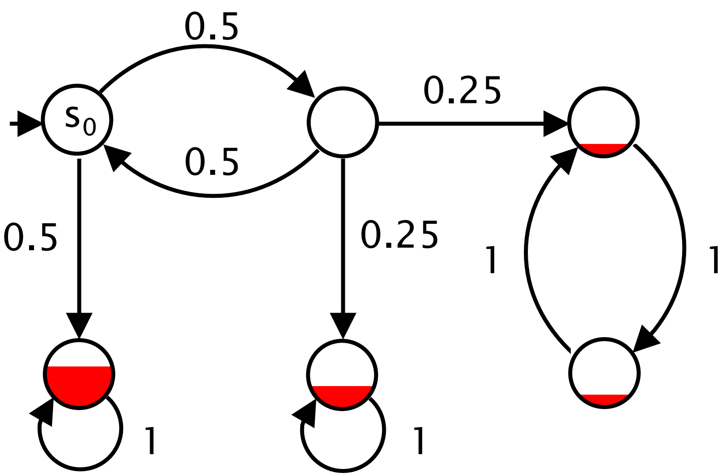
- In both cases, often beneficial to study long-run behaviour

Long-run behaviour

- Consider the limit: $\underline{\pi}_s = \lim_{k \to \infty} \underline{\pi}_{s,k}$
  - where $\underline{\pi}_{s,k}$ is the transient state distribution at time k, having started in state s
  - this limit, where it exists, is called the limiting distribution
  - steady-state of the model
- Intuitive idea
  - the percentage of time, in the long run, spent in each state
  - e.g. reliability: "in the long-run, what portion of time is the system in an operational state"

Limiting distribution

- Example:



$$\underline{\pi}_{s0,0} = \left[1,0,0,0,0,0\right]$$

$$\underline{\pi}_{s0,1} = \left[0,\frac{1}{2},0,\frac{1}{2},0,0\right]$$

$$\underline{\pi}_{s0,2} = \left[\frac{1}{4},0,\frac{1}{8},\frac{1}{2},\frac{1}{8},0\right]$$

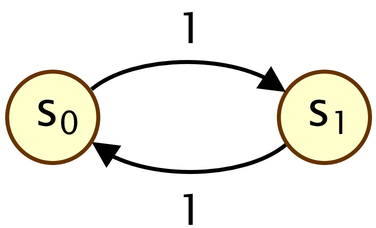
$$\underline{\pi}_{s0,3} = \left[0,\frac{1}{8},0,\frac{5}{8},\frac{1}{8},\frac{1}{8}\right]$$

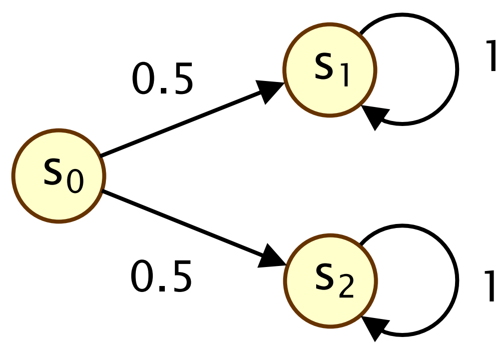
…

$$\underline{\pi}_{s0} = \left[0,0,\frac{1}{12},\frac{2}{3},\frac{1}{6},\frac{1}{12}\right]$$

Long-run behaviour

- Questions:
  - when does this limit exist?
  - does it depend on the initial state/distribution?





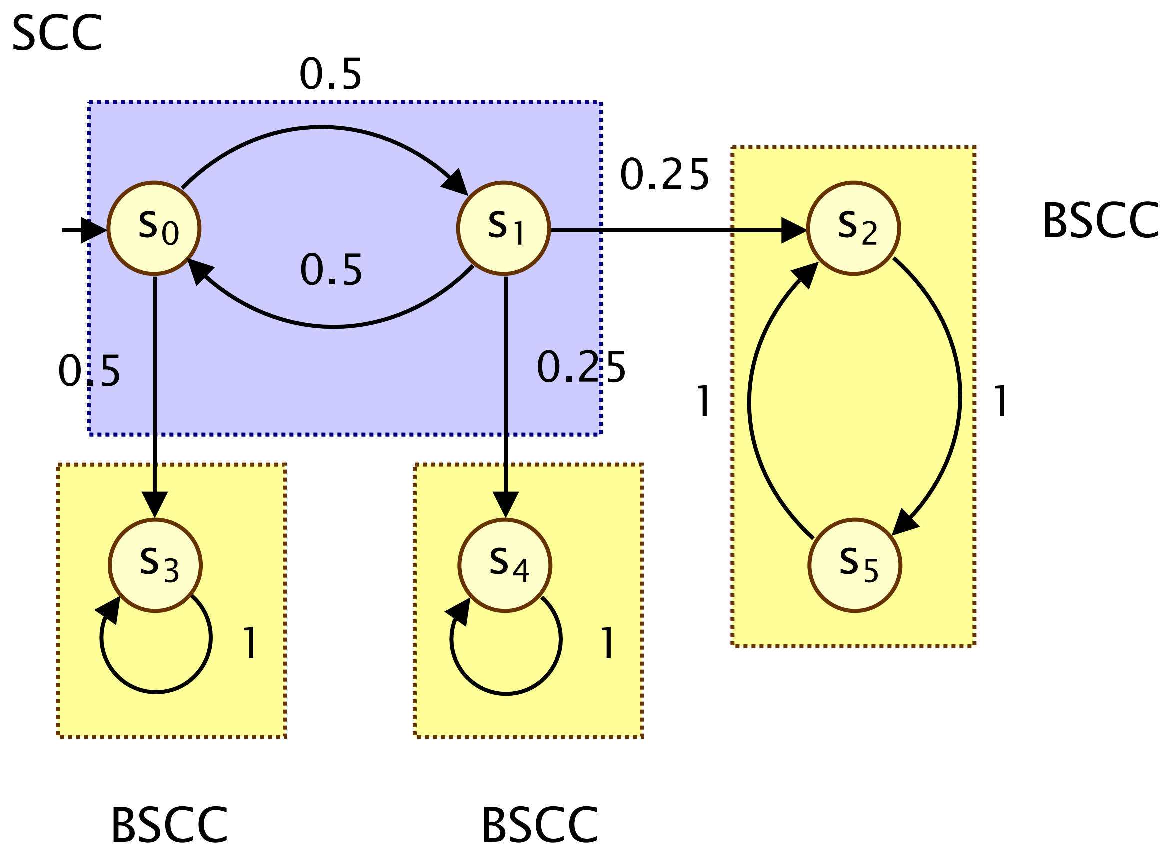
- Need to consider underlying graph
  - (V,E) where V are vertices and E ⊆ VxV are edges
  - V = S and E = { (s,s') s.t. **P**(s,s') > 0 }

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 *maximally* strongly connected set of states (i.e. no 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

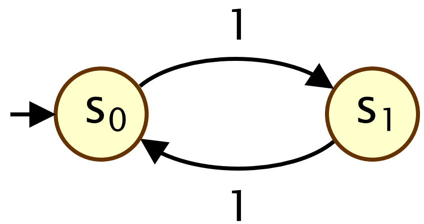
- Alternative terminology: "s communicates with s' ", "communicating class", "recurrent class"

Example – (B)SCCs



Graph terminology

- Markov chain 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 Markov chain is aperiodic if all its states have period 1

Steady-state probabilities

- For a finite, irreducible, aperiodic DTMC (a.k.a., ergodic)
  - limiting distribution always exists
  - and is independent of initial state/distribution
- These are known as steady-state probabilities
  - (or equilibrium probabilities)
  - effect of initial distribution has disappeared, denoted $\underline{\pi}$
- These probabilities can be computed as the unique solution of the linear equation system:

$$\underline{\pi} \cdot \mathbf{P} = \underline{\pi} \quad \text{and} \quad \sum_{s \in S} \underline{\pi}(s) = 1$$

Steady-state – Balance equations

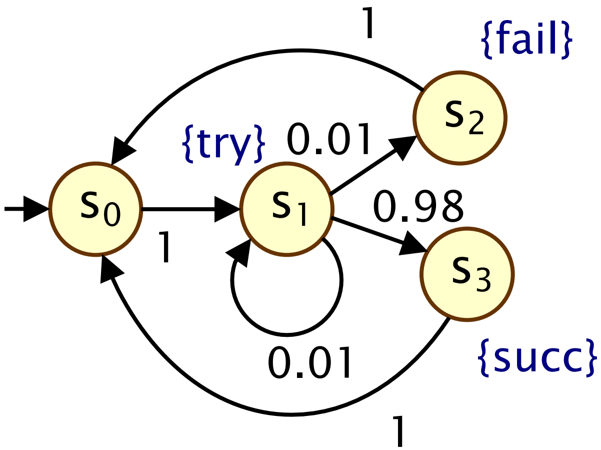
- Known as balance equations

$$\underline{\pi} \cdot \mathbf{P} = \underline{\pi} \quad \text{and} \quad \sum_{s \in S} \underline{\pi}(s) = 1$$

- That is:
  - $\underline{\pi}(s') = \Sigma_{s \in S}\ \underline{\pi}(s) \cdot \mathbf{P}(s,s')$ ← balance the probability of leaving and entering a state s'
  - $\Sigma_{s \in S}\ \underline{\pi}(s) = 1$ ← normalisation

Steady-state – Example

- Let $\underline{x} = \underline{\pi}$
- Solve: $\underline{x} \cdot \mathbf{P} = \underline{x}$, $\Sigma_s \underline{x}(s) = 1$



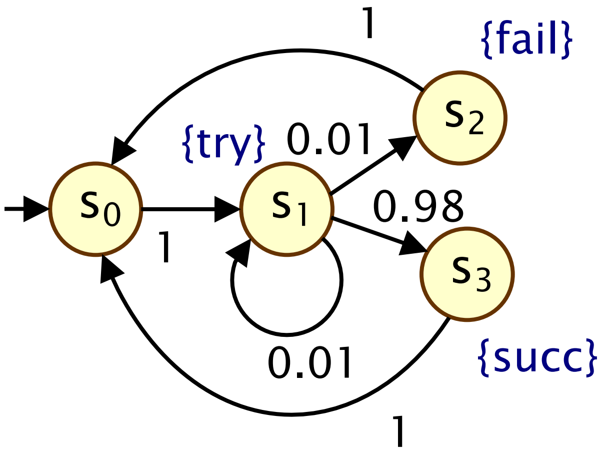
$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.01 & 0.01 & 0.98 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$\underline{x} \approx$ [0.332215, 0.335570, 0.003356, 0.328859]

Steady-state - Example

- Let $\underline{x} = \underline{\pi}$
- Solve: $\underline{x} \cdot \mathbf{P} = \underline{x}$, $\Sigma_s \underline{x}(s) = 1$

$\underline{x} \approx$ [0.332215, 0.335570, 0.003356, 0.328859]



Long-run percentage of time spent in the state "try" $\approx$ 33.6%

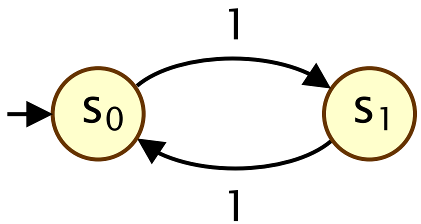
$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.01 & 0.01 & 0.98 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Long-run percentage of time spent in "fail"/"succ" $\approx 0.003356 + 0.328859$ $\approx$ 33.2%

Periodic DTMCs

- For (finite, irreducible) periodic DTMCs, this limit:

$$\underline{\pi}_s(s') = \lim_{k \to \infty} \underline{\pi}_{s,k}(s')$$



- in general does not exist, but this limit does:

$$\lim_{n \to \infty} \frac{1}{n} \cdot \sum_{k=1}^{n} \underline{\pi}_{s,k}(s')$$

(and where both limits exist, e.g. for aperiodic DTMCs, these 2 limits coincide)

- Steady-state probabilities for periodic DTMCs can still be computed, again by solving the same set of linear equations:

$$\underline{\pi} \cdot \mathbf{P} = \underline{\pi} \quad \text{and} \quad \sum_{s \in S} \underline{\pi}(s) = 1$$

Steady-state – General case

- General case: reducible DTMC
- there are multiple solutions of steady-state equation

$$\underline{\pi} \cdot \mathbf{P} = \underline{\pi} \quad \text{and} \quad \sum_{s \in S} \underline{\pi}(s) = 1$$

- number of (lin. Independent) solutions = number of BSCCs

- limiting distribution obtained by iterations exists
- limiting distribution depends on initial one

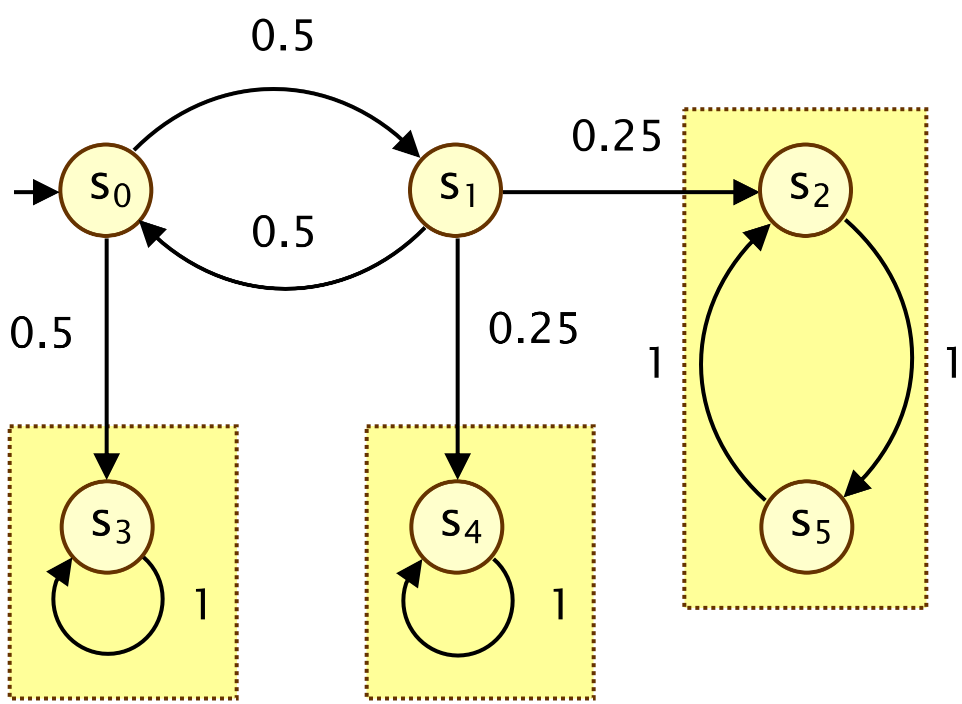
Steady-state – General case

- General case: reducible DTMC
  - compute vector $\underline{\pi}_s$

- Compute BSCCs for DTMC; then two cases to consider:
- (1) s is in a BSCC T
  - compute steady-state probabilities $\underline{x}$ in sub-DTMC for T
  - $\underline{\pi}_s(s') = \underline{x}(s')$ if s' in T
  - $\underline{\pi}_s(s') = 0$ if s' not in T
- (2) s is not in any BSCC
  - compute steady-state probabilities $\underline{x}_T$ for sub-DTMC of each BSCC T and combine with reachability probabilities to BSCCs
  - $\underline{\pi}_s(s') = \text{ProbReach}(s, T) \cdot \underline{x}_T(s')$ if s' is in BSCC T
  - $\underline{\pi}_s(s') = 0$ if s' is not in a BSCC

Steady-state – Example 2

- $\underline{\pi}_s$ depends on initial state s



$\underline{\pi}_{s3}$ = [0 0 0 1 0 0]

$\underline{\pi}_{s4}$ = [0 0 0 0 1 0]

$\underline{\pi}_{s2} = \underline{\pi}_{s5} = \left[0,0,\frac{1}{2},0,0,\frac{1}{2}\right]$

$\underline{\pi}_{s0} = \left[0,0,\frac{1}{12},\frac{2}{3},\frac{1}{6},\frac{1}{12}\right]$

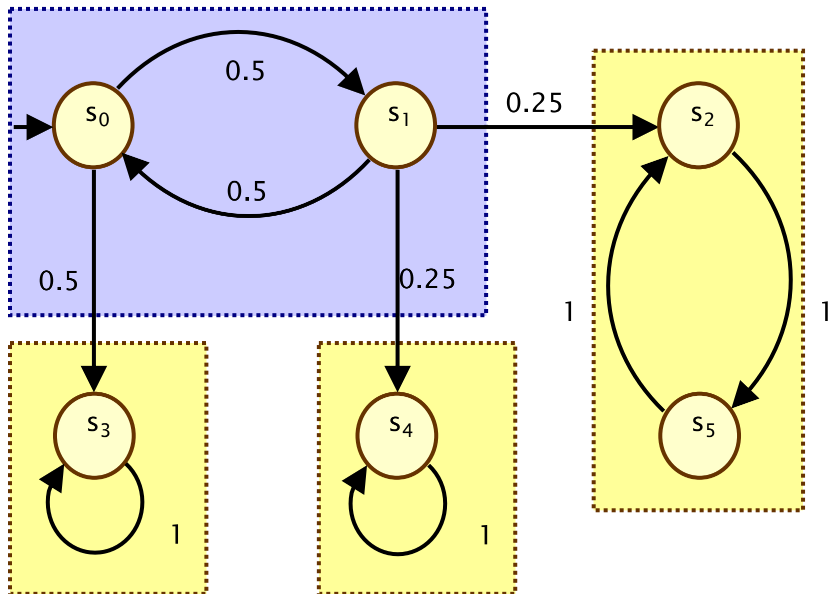
$\underline{\pi}_{s1}$ = …

Qualitative properties

- Quantitative properties:
  - "what is the probability of event A?"
- Qualititative properties:
  - "the probability of event A is 1" ("almost surely A")
  - or: "the probability of event A is > 0" ("possibly A")
- For finite DTMCs, qualititative properties do not depend on the transition probabilities - only need underlying graph
  - e.g. to determine "is target set T reached with probability 1?" (more in the DTMC model checking lecture later)
  - computing BSCCs of a DTMCs yields information about long-run qualitative properties…

Fundamental property

- Fundamental property of finite DTMCs…

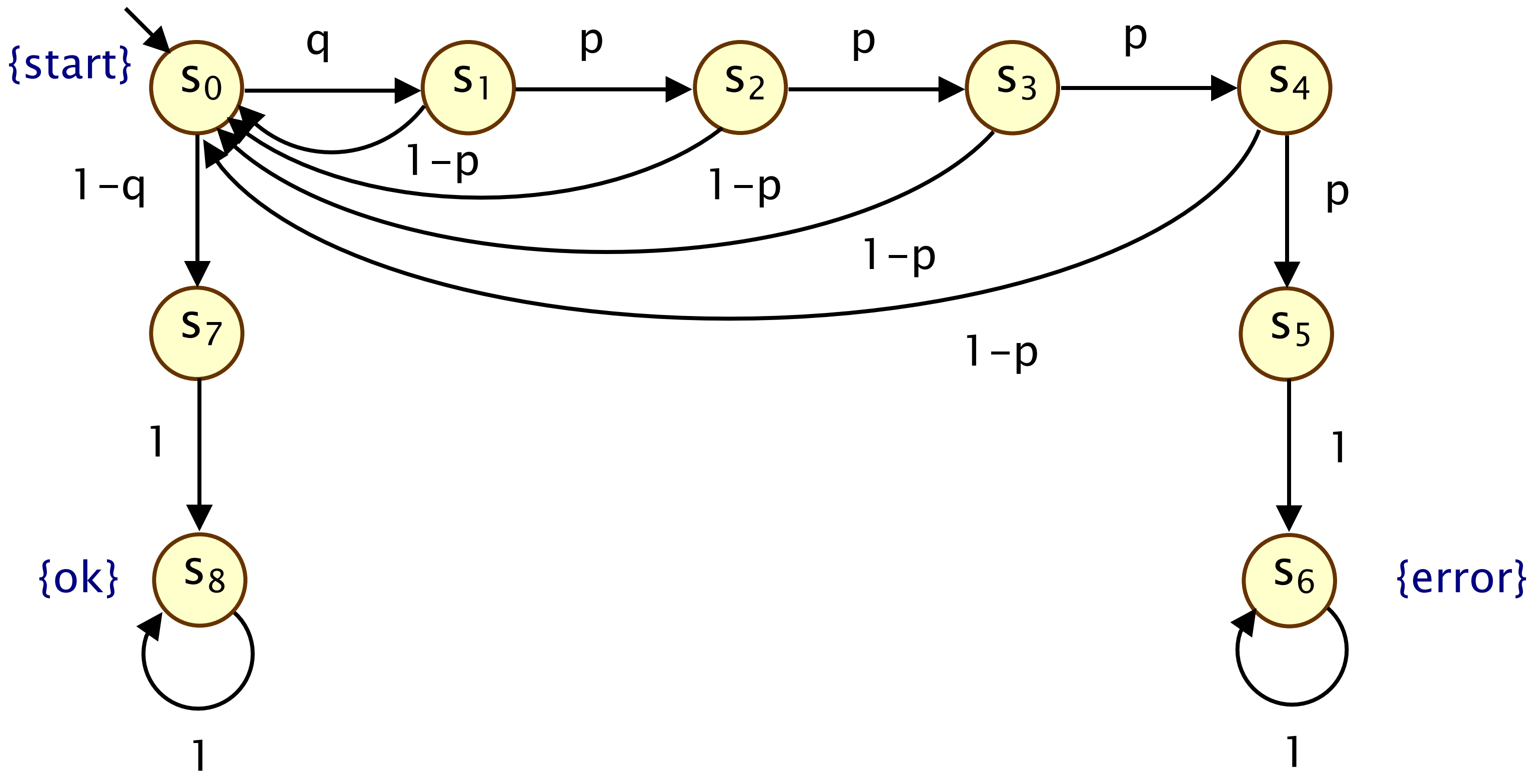


- With probability 1, a BSCC will be reached and all of its states visited infinitely often

- Formally:
  – Pr_{s0} ($s_0 s_1 s_2 \ldots$ | ∃ i≥0, ∃ BSCC T such that
  ∀ j≥i s_j ∈ T and
  ∀ s∈T s_k = s for infinitely many k) = 1

Zeroconf example

- 2 BSCCs: $\{s_6\}$, $\{s_8\}$
- Thus, probability of trying to acquire a new address infinitely often (i.e., visiting {start} i.o.) is 0



Repeated reachability

- Repeated reachability: GF B
  - "always eventually…", "infinitely often…"
- $Pr_{s0}(s_0 s_1 s_2 \ldots \mid \forall\ i \geq 0\ \exists\ j \geq i\ s_j \in B)$
  - 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: $\bigcap_{n \geq 0} \bigcup_{m \geq n} C_m$
  - where C_m is the union of all cylinder sets $Cyl(s_0 s_1 \ldots s_m)$ for finite paths $s_0 s_1 \ldots s_m$ such that $s_m \in B$

Qualitative repeated reachability

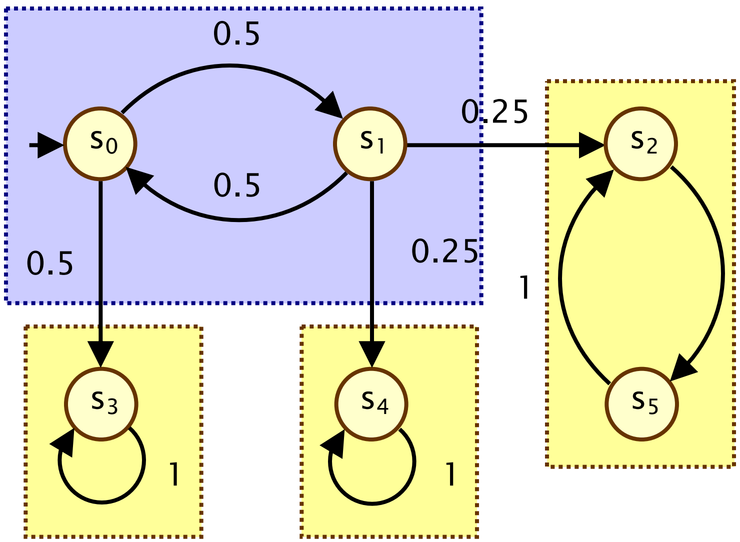
- Pr_{s0} ($s_0 s_1 s_2 \ldots$ | $\forall$ $i \geq 0$ $\exists$ $j \geq i$ $s_j \in B$) = 1
  Pr_{s0} ("always eventually B") = 1

  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…"
- Pr_{s0} ($s_0 s_1 s_2 \ldots \mid \exists\, i \geq 0\ \forall\, j \geq i\ s_j \in B$)
  - where $B \subseteq S$ is a set of states

- e.g. "what is the probability of the leader election algorithm reaching, and staying in, a stable state?"

- e.g. "what is the probability that an irrecoverable error occurs?"

- Is this measurable? Yes…

Persistence

- Persistence properties:
  - "eventually forever…"
- Pr_{s0} ($s_0 s_1 s_2 \ldots \mid \exists\ i \geq 0\ \forall\ j \geq i\ s_j \in B$)
  - where $B \subseteq S$ is a set of states

- e.g. "what is the probability of the leader election algorithm reaching, and staying in, a stable state?"

- e.g. "what is the probability that an irrecoverable error occurs?"

- Is this measurable? Yes… $FG\ B = \neg\ GF\ (S \backslash B)$

Qualitative persistence

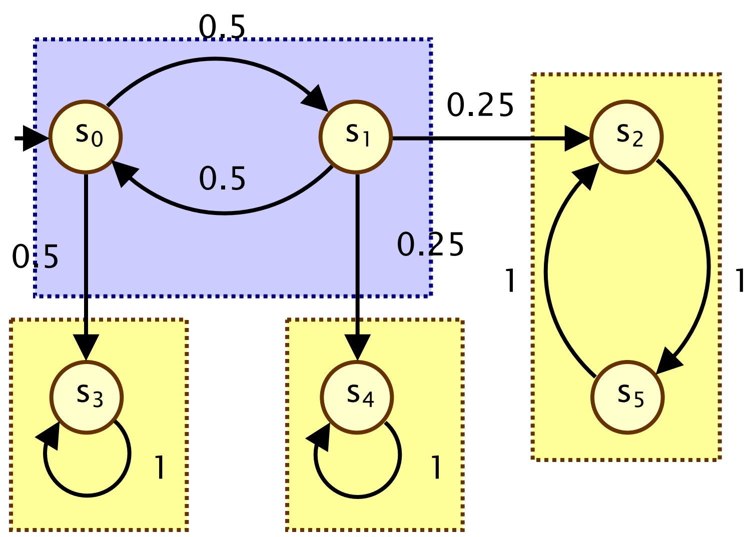
- Pr_{s0} ($s_0s_1s_2\ldots \mid \exists\ i \geq 0\ \forall\ j \geq i\ s_j \in B$) = 1
  Pr_{s0} ("eventually forever B") = 1

    if and only if

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

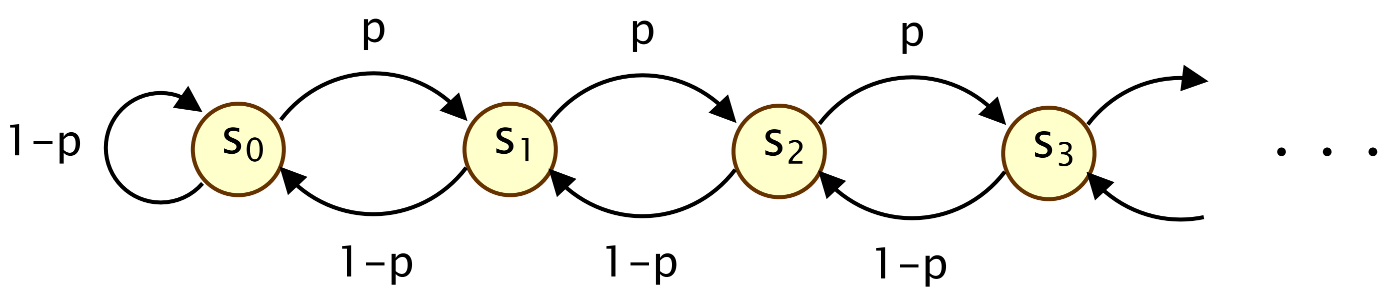
Example:

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



Aside: Infinite-state Markov chains

- Infinite-state random walk



- Value of probability p does affect qualitative properties
  - ProbReach(s, $\{s_0\}$) $= 1$ if $p \leq 0.5$
  - ProbReach(s, $\{s_0\}$) < 1 if $p > 0.5$
- (not comprehensively studied in this course)

Summing up…

- Transient state probabilities
  - successive vector-matrix multiplications

- Long-run/steady-state probabilities
  - requires graph analysis
  - irreducible case: solve linear equation system
  - reducible case: steady-state for sub-DTMCs + reachability

- Qualitative properties
  - repeated reachability
  - persistence