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**: \( \pi_{s,k} \)
  - (row) vector \( \pi_{s,k} \) i.e. \( \pi_{s,k}(s') \) for all states \( s' \)

- **Note**: this is a discrete probability distribution
  - so we have \( \pi_{s,k} : S \rightarrow [0,1] \)
  - recall instead \( \Pr_s : \Sigma_{\text{Path}(s)} \rightarrow [0,1] \), where \( \Sigma_{\text{Path}(s)} \subseteq 2^{\text{Path}(s)} \)
Transient distributions

k=0:

k=1:

k=2:

k=3:
Computing transient probabilities

- **Transient state probabilities:**
  - $\pi_{s,k}(s') = \sum_{s'' \in S} \pi_{s,k-1}(s'') \cdot P(s'',s')$
  - (i.e. look at incoming transitions, into $s'$)

- **Computation of transient state distribution:**
  - $\pi_{s,0}$ 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
  - $\pi_{s,k} = \pi_{s,k-1} \cdot P$

- i.e. successive vector–matrix multiplications
Computing transient probabilities

\[ \begin{align*}
\pi_{s0,0} &= [1, 0, 0, 0, 0] \\
\pi_{s0,1} &= [0, \frac{1}{2}, 0, \frac{1}{2}, 0] \\
\pi_{s0,2} &= [\frac{1}{4}, 0, \frac{1}{8}, \frac{1}{2}, \frac{1}{8}, 0] \\
\pi_{s0,3} &= [0, \frac{1}{8}, 0, \frac{5}{8}, \frac{1}{8}, \frac{1}{8}] \\
\end{align*} \]

\[ P = \begin{pmatrix}
0 & 0.5 & 0 & 0.5 & 0 & 0 \\
0.5 & 0 & 0.25 & 0 & 0.25 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 \\
\end{pmatrix} \]
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 probabilities of \( k \)-step transition probabilities
  - i.e. \( P^k(s,s') = \pi_{s,k}(s') \)

- A possible optimisation: 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
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:** $\pi_s = \lim_{k \to \infty} \pi_{s,k}$
  – where $\pi_{s,k}$ 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:

\[
\begin{align*}
\pi_{s0,0} &= [1,0,0,0,0,0,0] \\
\pi_{s0,1} &= [0,\frac{1}{2},0,\frac{1}{2},0,0] \\
\pi_{s0,2} &= \left[\frac{1}{4},0,\frac{1}{8},\frac{1}{2},\frac{1}{8},0\right] \\
\pi_{s0,3} &= \left[0,\frac{1}{8},0,\frac{5}{8},\frac{1}{8},\frac{1}{8}\right] \\
\vdots \\
\pi_{s0} &= \left[0,0,\frac{1}{12},\frac{2}{3},\frac{1}{6},\frac{1}{12}\right]
\end{align*}
\]
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 \(E \subseteq V \times V\) are edges
  – \(V = S\) and \(E = \{(s,s') \text{ s.t. } P(s,s') > 0\}\)
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

\[
\begin{array}{c}
\text{S_0} \\
\downarrow 1 \\
\text{S_1}
\end{array}
\]

- 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 $\pi$

• These probabilities can be computed as the unique solution of the linear equation system:

$$\pi \cdot P = \pi \quad \text{and} \quad \sum_{s \in S} \pi(s) = 1$$
**Steady-state – Balance equations**

- Known as **balance equations**

\[
\pi \cdot P = \pi \quad \text{and} \quad \sum_{s \in S} \pi(s) = 1
\]

- That is:

- \( \pi(s') = \sum_{s \in S} \pi(s) \cdot P(s, s') \)

- \( \sum_{s \in S} \pi(s) = 1 \)

balance the probability of leaving and entering a state \( s' \)

normalisation
Steady-state – Example

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

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0.01 & 0.01 & 0.98 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \approx \begin{bmatrix} 0.332215, & 0.335570, \\ 0.003356, & 0.328859 \end{bmatrix}$$
Steady-state – Example

- Let \( x = \pi \)
- Solve: \( x \cdot P = x, \quad \Sigma_s x(s) = 1 \)

\[
P = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0.01 & 0.01 & 0.98 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

\[
x \approx [0.332215, 0.335570, 0.003356, 0.328859]
\]

Long-run percentage of time spent in the state “try”
\[
\approx 33.6\%
\]

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:

\[ \lim_{k \to \infty} \prod_{s,k} (s') = \lim_{n \to \infty} \prod_{s,n} (s') \]

• In general does not exist, but this limit does:

\[ \lim_{n \to \infty} \frac{1}{n} \cdot \sum_{k=1}^{n} \prod_{s,k} (s') \] (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:

\[ \pi \cdot P = \pi \quad \text{and} \quad \sum_{s \in S} \pi(s) = 1 \]
Steady–state – General case

- General case: reducible DTMC
- there are multiple solutions of steady–state equation
  \[ \pi \cdot P = \pi \quad \text{and} \quad \sum_{s \in S} \pi(s) = 1 \]
- 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 $\pi_s$

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

- $\pi_s$ depends on initial state $s$

\[
\begin{align*}
\pi_{s3} &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\
\pi_{s4} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\
\pi_{s2} &= \pi_{s5} = \begin{bmatrix} 0, 0, \frac{1}{2}, 0, 0, \frac{1}{2} \end{bmatrix} \\
\pi_{s0} &= \begin{bmatrix} 0, 0, \frac{1}{12}, \frac{2}{3}, \frac{1}{6}, \frac{1}{12} \end{bmatrix} \\
\pi_{s1} &= \ldots
\end{align*}
\]
Qualitative properties

- **Quantitative properties:**
  - “what is the probability of event A?”

- **Qualitative properties:**
  - “the probability of event A is 1” (“almost surely A”)
  - or: “the probability of event A is > 0” (“possibly A”)

- For finite DTMCs, qualitative 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:
  \[
  \Pr_{s_0}(s_0s_1s_2\ldots | \exists i \geq 0, \exists \text{BSCC } T \text{ such that } \forall j \geq i \ s_j \in T \text{ and } \forall \ s \in T \ s_k = s \text{ for infinitely many } k) = 1
  \]
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...”
• $\Pr_{s_0}(s_0 s_1 s_2... \mid \forall i \geq 0 \exists j \geq i s_j \in B)$
  – 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: $\bigcap \bigcup_{n \geq 0 \text{ and } m \geq n} C_m$
  – where $C_m$ is the union of all cylinder sets $\text{Cyl}(s_0 s_1...s_m)$ for finite paths $s_0 s_1...s_m$ such that $s_m \in B$
Qualitative repeated reachability

- \( \Pr_{s_0}(s_0s_1s_2... | \forall i \geq 0 \exists j \geq i \ s_j \in B) = 1 \)
  \( \Pr_{s_0}(\text{“always eventually B”}) = 1 \)

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…”
- $\text{Pr}_{s_0}( s_0 s_1 s_2 \ldots \ | \ \exists \ i \geq 0 \ \forall \ j \geq i \ s_j \in B )$
  - 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…”

• $\Pr_{s_0} (s_0 s_1 s_2 ... \mid \exists i \geq 0 \ \forall j \geq i \ s_j \in B)$
  – 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… $FG B = \neg GF (S \setminus B)$
Qualitative persistence

\[ \Pr_{s_0}(s_0s_1s_2\ldots \mid \exists i \geq 0 \ \forall j \geq i \ s_j \in B) = 1 \]
\[ \Pr_{s_0}(\text{“eventually forever } B\text{”}) = 1 \]

if and only if

\[ T \subseteq B \text{ for each BSCC } T \text{ that is reachable from } s_0 \]

Example:
\[ B = \{ s_2, s_3, s_4, s_5 \} \]
Aside: Infinite–state Markov chains

• Infinite–state random walk

![Diagram of random walk](image)

• Value of probability $p$ does affect qualitative properties
  
  – $\text{ProbReach}(s, \{s_0\}) = 1$ if $p \leq 0.5$
  
  – $\text{ProbReach}(s, \{s_0\}) < 1$ 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