

Finite State Automata and Markov Chains

Reading

- E.A. Lee and S.A. Seshia, Introduction to Embedded Systems: CPS Approach, Second Edition, MIT Press, 2017
 - Book:
https://ptolemy.berkeley.edu/books/leeseshia/releases/Lee_Seshia_DigitalV2_2.pdf
 - Chapter 3
- Not exactly a standard DFA chapter, has a dynamical system bias, but similar to MDPs

A Simple Dynamics Model

- Suppose a car is moving in a straight line at $v \text{ m/s}$
- How much will the car have travelled after $T \text{ s}$?
$$vT \text{ m}$$
- Suppose the car's position at time 0 is p_0 and at time T is p_T
$$p_T = p_0 + vT$$
- Suppose every T seconds velocity jumps up by $a \text{ m/s}$
- How do we adapt the model (for discrete times when velocity is changed)?

$$p_{kT} = p_{(k-1)T} + v_{(k-1)T}T$$

$$v_{kT} = v_{(k-1)T} + a$$

– where $k = 1, 2, \dots$

A Simple Dynamics Model, cont'd

- Discrete-time position/velocity model

$$p_{kT} = p_{(k-1)T} + v_{(k-1)T}T$$

$$v_{kT} = v_{(k-1)T} + a$$

- This is a discrete-time model where $\mathbf{x} = [p, v]^T$, $u_k = a$, so

$$f([x_1, x_2], u) = \begin{bmatrix} x_1 + x_2 T \\ x_2 + u \end{bmatrix}$$

- In this case, f is linear, so the system can also be written as

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k$$

– where $\mathbf{A} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

- Note that we implicitly dropped the T in the subscript
 - It is redundant, since k is chosen for a sampling rate of T

General Model Form

- In its most general form, the model can be written as

$$\begin{aligned}\mathbf{x}_{k+1} &= f(\mathbf{x}_k, \mathbf{u}_k) \\ \mathbf{u}_k &= h(\mathbf{x}_k)\end{aligned}$$

- This model has the Markov property, i.e., the current state depends only on the previous state and control
 - It doesn't matter how we got to the previous state
- Given f , one needs to design a controller $\mathbf{u}_k = h(\mathbf{y}_k)$
 - E.g., to navigate the track as fast as possible
 - In standard control, how do we pick the controls $\mathbf{u}_k$?
 - Minimize a cost function (surprise, surprise), e.g.,

$$J = \mathbf{x}_{k+H}^T \mathbf{Q} \mathbf{x}_{k+H} + \sum_{j=0}^{H-1} \mathbf{x}_{k+j}^T \mathbf{Q} \mathbf{x}_{k+j} + \mathbf{u}_{k+j}^T \mathbf{R} \mathbf{u}_{k+j} + \mathbf{x}_{k+j}^T \mathbf{N} \mathbf{u}_{k+j}$$

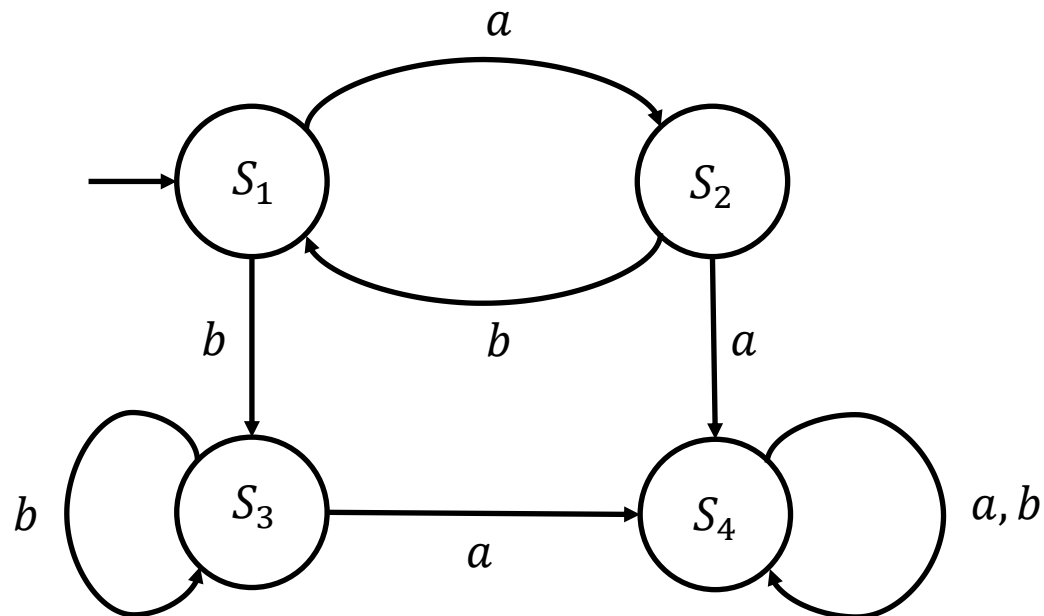
- where H is a time horizon, $\mathbf{Q}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{N}$ are user-defined matrices

Reinforcement learning modeling and cost function

- Historically, RL theory has been based on finite state models
 - The f formulation is infinite-state
 - However, deep RL is increasingly (and surprisingly) able to work in infinite-state settings
 - RL models (often) also have the Markov property
- Unlike optimal control, RL doesn't minimize a cost function
 - It maximizes a reward function
 - Mathematically, there is no difference
 - Maybe RL researchers are young and optimistic O.o

Finite State Machines

- One of the fundamental models in computer science
- Also known as deterministic finite automata (DFA)
- Historically used to model computer programs
 - DFAs are a simple model but have served us well



Finite State Machines Formalization

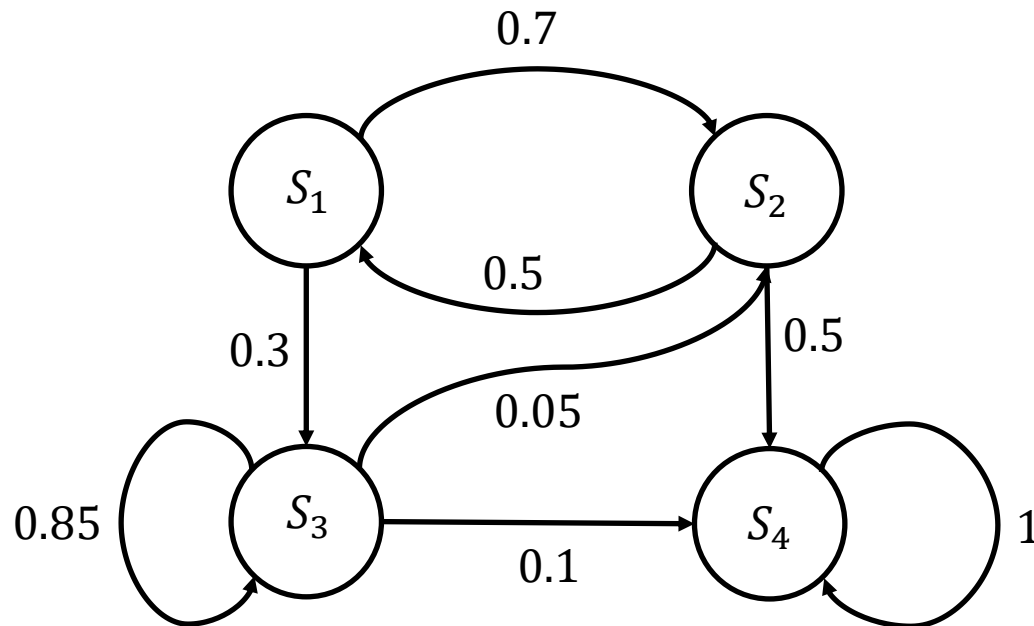
- A DFA is a tuple (A, S, S_0, δ, F) , where
 - A is the input alphabet
 - S is the finite set of states
 - S_0 is the initial state
 - $\delta: S \times A \rightarrow S$ is the transition function
 - F is the (possibly empty) set of final (accepting) states
- For each state and input pair S and A , $\delta(S, A)$ outputs exactly one state
 - Hence the deterministic in the name
 - e.g., $\delta(S_1, a) = S_2$
- In a non-deterministic FA (NFA), δ can output 0 or more values
 - Every NFA can be converted to an equivalent DFA

DFA Expressivity

- DFAs are one of the simplest models of computation
 - E.g., simpler than pushdown automata, Turing machines
- At the same time, many problems are just extremely large DFAs
 - E.g., games are for the most part (very large) DFAs
 - E.g., in chess, every position is a state and every input (move/action) causes a transition to exactly one state
- Classical RL was actually developed for stochastic models, not deterministic
 - More expressive than DFAs
- To get there, we need to talk about Markov chains first

Markov Chains

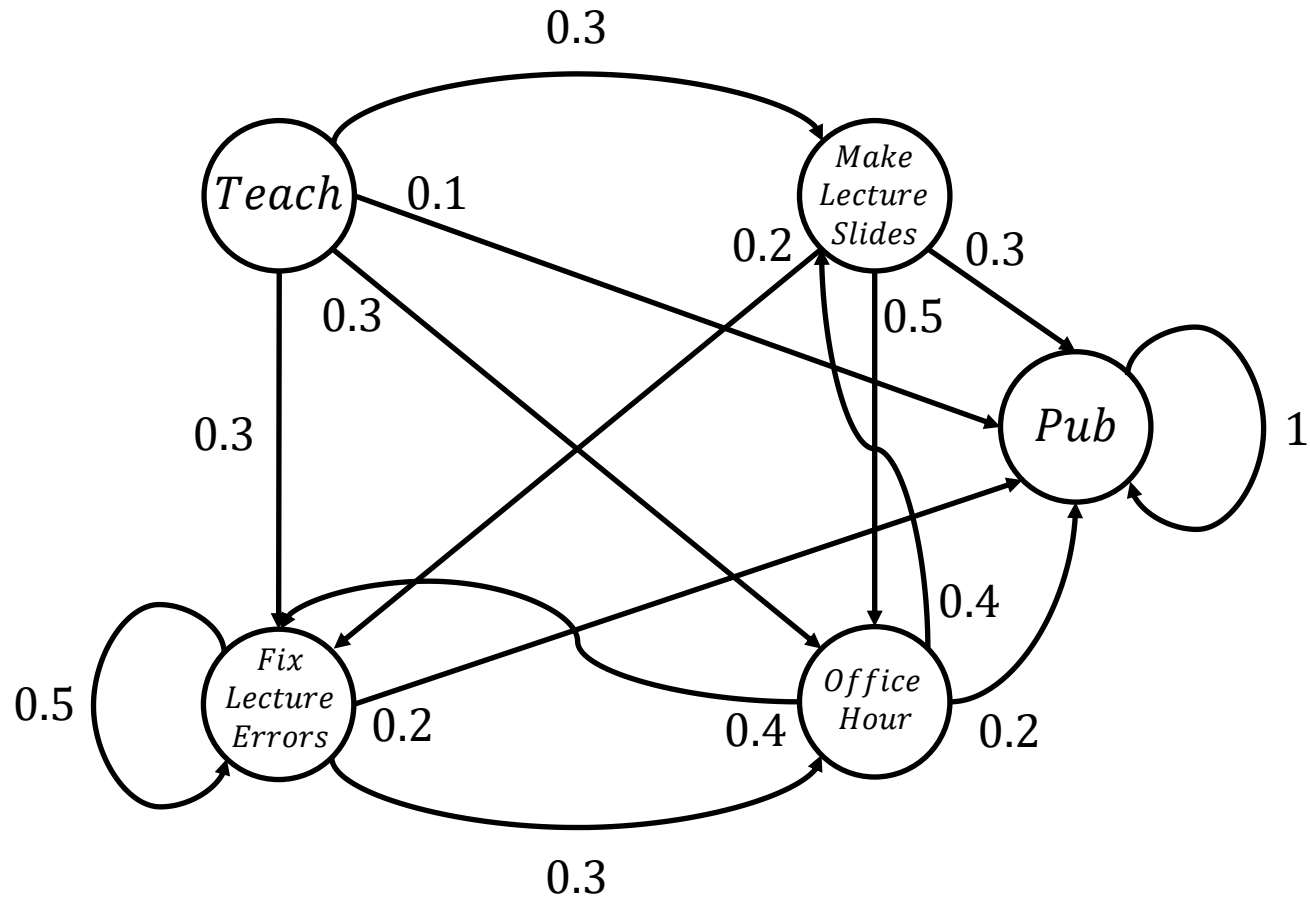
- Markov chains are effectively probabilistic automata
 - Formulation can be made more general, but we'll only need the finite-state version



- Each transition has an associated probability
 - E.g., probability of going from S_1 to S_2 is 0.7

Workday Example

- Markov chain describing my workday



Formalization

- A Markov Chain is a tuple (S, P, η) , where
 - S is the finite set of states
 - $P: S \times S \rightarrow \mathbb{R}$ is the probabilistic transition function
 - $\eta: S \rightarrow \mathbb{R}$ is the initial state distribution
- Called Markov chain because the probability of the current state is determined only by the previous state

$$\mathbb{P}[S_t | S_{t-1}, S_{t-2}, \dots, S_0] = \mathbb{P}[S_t | S_{t-1}] = P(S_{t-1}, S_t)$$

– where S_t denotes the state after t steps

– Examples:

$$\begin{aligned}\mathbb{P}[S_0 = \textit{Teach}] &= \eta(\textit{Teach}) \\ \mathbb{P}[S_t = \textit{Pub} | S_{t-1} = \textit{Office Hour}, S_0 = \textit{Teach}] &= \\ \mathbb{P}[S_t = \textit{Pub} | S_{t-1} = \textit{Office Hour}] &= \\ &= 0.2\end{aligned}$$

Examples

- What is the probability that I am at *Pub* two steps after *Teach*?
 - Need to look at all possible ways to get to *Pub* in two steps
- Formally:

$$\begin{aligned}\mathbb{P}[S_2 = \textit{Pub} | S_0 = \textit{Teach}] = & \\ & \mathbb{P}[S_1 = \textit{Pub}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] + \\ & \mathbb{P}[S_1 = \textit{Office Hour}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] + \\ & \mathbb{P}[S_1 = \textit{Fix Lecture Errors}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] + \\ & \mathbb{P}[S_1 = \textit{Make Lecture Slides}, S_2 = \textit{Pub} | S_0 = \textit{Teach}]\end{aligned}$$

- Summing through all possibilities is called marginalization
- Recall the definition of conditional probability:

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A, B]}{\mathbb{P}[B]}$$

Examples

- What is the probability that I am at *Pub* two steps after *Teach*?
 - Need to look at all possible ways to get to *Pub* in two steps
- Formally:

$$\begin{aligned}\mathbb{P}[S_2 = \textit{Pub} | S_0 = \textit{Teach}] = & \\ & \mathbb{P}[S_1 = \textit{Pub}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] + \\ & \mathbb{P}[S_1 = \textit{Office Hour}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] + \\ & \mathbb{P}[S_1 = \textit{Fix Lecture Errors}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] + \\ & \mathbb{P}[S_1 = \textit{Make Lecture Slides}, S_2 = \textit{Pub} | S_0 = \textit{Teach}]\end{aligned}$$

- Summing through all possibilities is called marginalization
- Probabilities are:

$$\begin{aligned}\mathbb{P}[S_1 = \textit{Pub}, S_2 = \textit{Pub} | S_0 = \textit{Teach}] = & \\ \mathbb{P}[S_2 = \textit{Pub} | S_1 = \textit{Pub}, S_0 = \textit{Teach}] \mathbb{P}[S_1 = \textit{Pub} | S_0 = \textit{Teach}] = & \\ \mathbb{P}[S_2 = \textit{Pub} | S_1 = \textit{Pub}] \mathbb{P}[S_1 = \textit{Pub} | S_0 = \textit{Teach}] = & 0.1\end{aligned}$$

Examples, cont'd

- What is the probability that I am at *Pub* two steps after *Teach*?
 - Need to look at all possible ways to get to *Pub* in two steps
- All possible paths
 - *Pub, Pub*
 - *Office Hour, Pub*
 - *Fix Lecture Errors, Pub*
 - *Make Lecture Slides, Pub*
- Probabilities are:
 - $\mathbb{P}[S_1 = \text{Office Hour}, S_2 = \text{Pub} | S_0 = \text{Teach}] = 0.06$
 - $\mathbb{P}[S_1 = \text{Fix Lecture Errors}, S_2 = \text{Pub} | S_0 = \text{Teach}] = 0.06$
 - $\mathbb{P}[S_1 = \text{Make Lecture Slides}, S_2 = \text{Pub} | S_0 = \text{Teach}] = 0.09$
- Total probability is $0.1 + 0.06 + 0.06 + 0.09 = 0.31$

Linear Algebra Aside

- Suppose we are given a square matrix $A \in \mathbb{R}^{n \times n}$
- A vector v is said to be an eigenvector of A if
$$Av = \lambda v$$
 - Where $\lambda \in \mathbb{R}$ is a corresponding eigenvalue
- The matrix A has n eigenvectors, v_i
 - And n corresponding eigenvalues, λ_i
- If eigenvalues are distinct, the eigenvectors form a basis in $\mathbb{R}^n$
 - i.e., any $x \in \mathbb{R}^n$ can be written as a linear combination
$$x = c_1 v_1 + \cdots + c_n v_n$$
- There may be repeated eigenvalues
- A is full rank iff $\lambda_i \neq 0$ for all i

Transition Matrix

- We can store all transition probabilities in a matrix $\mathbf{P}$
- Entry P_{ij} denotes the probability of going from state i to j
- E.g., let states be ordered:
Teach, Office Hour, MLS, FLE, Pub
- The transition matrix becomes:

$$\mathbf{P} = \begin{bmatrix} 0 & 0.3 & 0.3 & 0.3 & 0.1 \\ 0 & 0 & 0.4 & 0.4 & 0.2 \\ 0 & 0.5 & 0 & 0.2 & 0.3 \\ 0 & 0.3 & 0 & 0.5 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- What properties does $\mathbf{P}$ have?

Transition Matrix Properties

- Each row must sum up to 1
 - Why?
 - For each state, transition probabilities must sum up to 1
- Has an eigenvalue of 1
 - Why? What is the corresponding eigenvector?
 - Pick any row, p_j^T
 - Let $\mathbf{1} \in \mathbb{R}^{|S|}$ be a vector of all ones
 - What is $p_j^T \mathbf{1}$?
 - 1! So $\mathbf{1}$ is an eigenvector

Transition Matrix Properties, cont'd

- Let $\boldsymbol{\eta}_t$ represent the probabilities the system is in any given state at time t

– E.g., $\boldsymbol{\eta}_t = [1 \ 0 \ 0 \ 0 \ 0]^T$ means the state is *Teach*

- What happens if we multiply $\boldsymbol{\eta}_t^T \mathbf{P}$?

$$[1 \ 0 \ 0 \ 0 \ 0] \begin{bmatrix} 0 & 0.3 & 0.3 & 0.3 & 0.1 \\ 0 & 0 & 0.4 & 0.4 & 0.2 \\ 0 & 0.5 & 0 & 0.2 & 0.3 \\ 0 & 0.3 & 0 & 0.5 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = [0 \ 0.3 \ 0.3 \ 0.3 \ 0.1]$$

- We get the distribution of states after one step, i.e., $\boldsymbol{\eta}_{t+1}^T$
 - What happens if we multiply $\boldsymbol{\eta}_t^T \mathbf{P} \mathbf{P}$?

Transition Matrix Properties, cont'd

- What happens if we multiply $\eta_t^T P P$?

$$\eta_t^T P P = \eta_{t+1}^T P = \eta_{t+2}^T$$

- Now suppose you are given η_0
 - The distribution at time 0
- How do you express η_t as a function of η_0 and P ?

$$\eta_t^T = \eta_0^T P^t$$

- Can quickly compute state distributions over time
- What does this expression remind you of?
 - It's a linear system!

$$\begin{aligned} x_{k+1} &= A x_k \\ \eta_{t+1} &= P^T \eta_t \\ &= (P^T)^{t+1} \eta_0 \end{aligned}$$

Linear Algebra Aside, cont'd

- Suppose a square matrix A has eigenvalues $\lambda_1, \dots, \lambda_n$
- What are the eigenvalues of A^2 ?
- Take any eigenvalue λ_i and corresponding eigenvector $\mathbf{v}_i$

$$\begin{aligned} A A \mathbf{v}_i &= A \lambda_i \mathbf{v}_i \\ &= \lambda_i^2 \mathbf{v}_i \end{aligned}$$

- In general, the eigenvalues of A^k are

$$\lambda_1^k, \dots, \lambda_n^k$$

– The eigenvectors are the same as those of A

Linear System Solution

- Consider a general discrete-time linear system

$$\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0$$

- Suppose $\mathbf{A}$ has distinct eigenvalues for simplicity
- Recall that the eigenvectors of $\mathbf{A}$ form a basis in $\mathbb{R}^n$, so

$$\mathbf{x}_0 = a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n$$

- Then

$$\mathbf{A}^k \mathbf{x}_0 = a_1 \lambda_1^k \mathbf{v}_1 + \cdots + a_n \lambda_n^k \mathbf{v}_n$$

- Under what conditions does $\mathbf{x}_k$ converge to $\mathbf{0}$?

$$|\lambda_i| < 1, \text{ for all } i$$

Transition Matrix Linear Systems

- Consider the transition matrix linear system

$$\boldsymbol{\eta}_{t+1} = (\boldsymbol{P}^T)^{t+1} \boldsymbol{\eta}_0$$

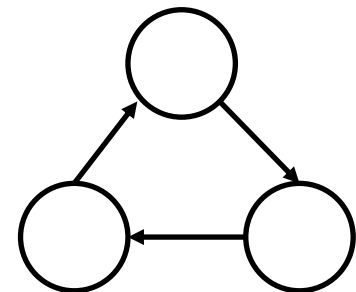
- We know that $\mathbf{1}$ is an eigenvector of $\boldsymbol{P}$
 - Also known as a right eigenvector
 - However, we are now interested in left eigenvectors
 - AKA eigenvectors of $\boldsymbol{P}^T$

$$\boldsymbol{v}^T \boldsymbol{P} = (\boldsymbol{P}^T \boldsymbol{v})^T$$

- It turns out that $\boldsymbol{P}$ also has a left eigenvalue of 1
 - Left and right eigenvalues are the same for square matrices
 - Eigenvectors may be different
- This means the system never converges to 0
 - But what does it converge to?

Transition Matrix Linear Systems, cont'd

- Consider a row vector μ such that
$$\mu P = \mu$$
- Then μ is an eigenvector corresponding to eigenvalue 1
 - There could be more than 1 such vectors
- What graph property determines whether there is a unique μ ?
 - There is one μ per closed communication class
 - i.e., loop in the graph that cannot be left
 - Formally, having one such class is known as irreducibility
- Another requirement is aperiodicity
 - If you have a periodic graph, you will never converge

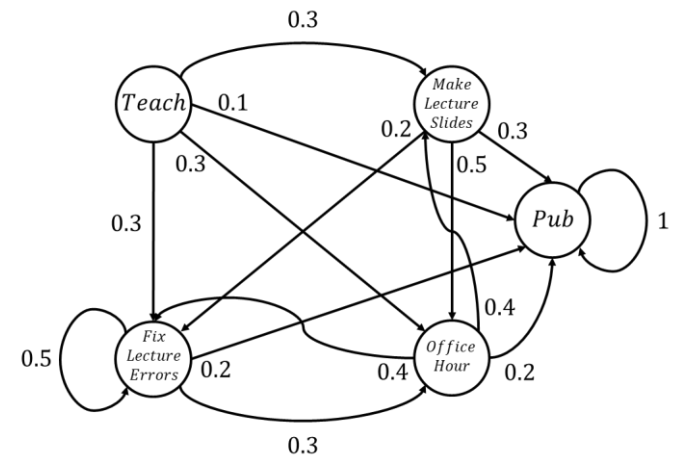


Stationary Distribution

- If you have an aperiodic, irreducible Markov chain, then there is a unique μ such that

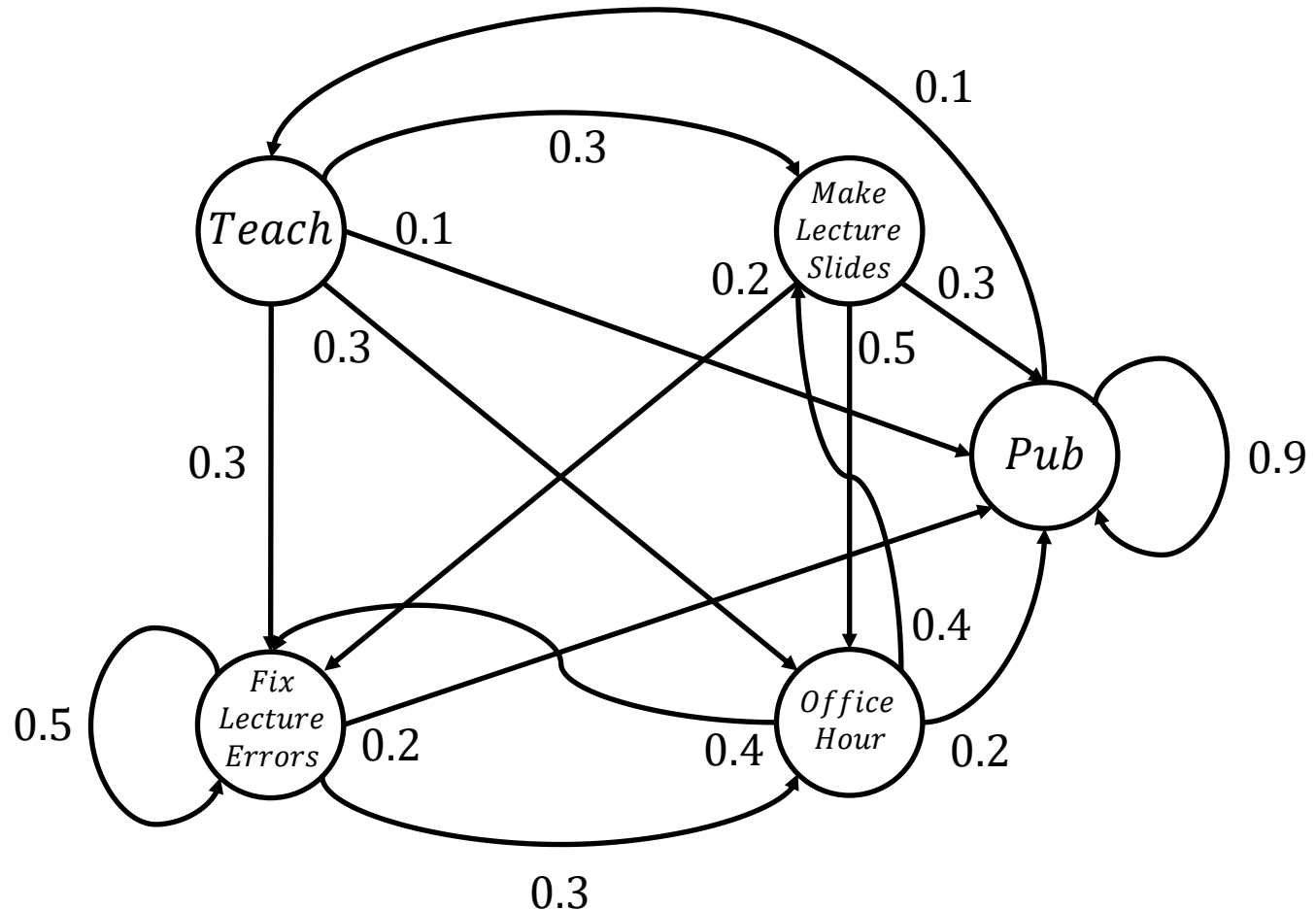
$$\mu P = \mu$$

- This is known as the stationary distribution
 - Each element of μ denotes the *proportion* of time spent in that state in the long run
- What is μ for the workday example?
 - It is $[0 \ 0 \ 0 \ 0 \ 1]$
 - Pub* is an absorbing state
 - With probability 1, every trace eventually gets to *Pub*
- Teach* is a transient state
 - You cannot go back to it



Workday Example, revisited

- Suppose I add a new transition from *Pub* to *Teach*



Stationary Distribution for Revisited Workday Example

- Stationary distribution is hard to derive by simply looking at the graph anymore
- Two ways of finding μ
 - Can either find left eigenvalues and eigenvectors of P
 - Which eigenvalue does μ correspond to?
1
 - Might need to normalize eigenvector
 - Or just compute P^t for a big t and then compute $\eta_0 P^t$
 - Recall μ is the same for any initial η_0
- For the revisited example
$$\mu = [0.067 \quad 0.086 \quad 0.054 \quad 0.13 \quad 0.66]$$
- Still spending most time in *Pub*, but other states are also visited infinitely often