CSE 312

# Foundations of Computing II

## Lecture 26: Markov Chain

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

**Rachel Lin, Hunter Schafer**

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Alex Tsun's and Anna Karlin's slides for 312 20su and 20au

## So far, a single-shot random process

Random Process → **Outcome Distribution** $D$

**Today:**
See a very special type of DTSP called **Markov Chains**

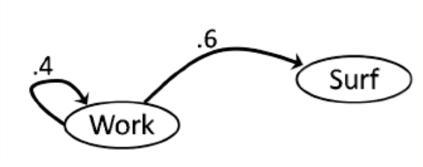
## Many-step random process

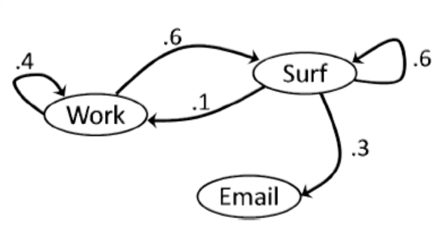
Random Process 1 → **Outcome Distribution** $D_1$ → Random Process 2 → **Outcome Distribution** $D_2$ → Random Process 3 → **Outcome Distribution** $D_3$ …

**Definition:** A **discrete-time stochastic process** (DTSP) is a sequence of random variables $X^{(0)}, X^{(1)}, X^{(2)}, \ldots$ where $X^{(t)}$ is the value at time $t$.

# A boring day in pandemic

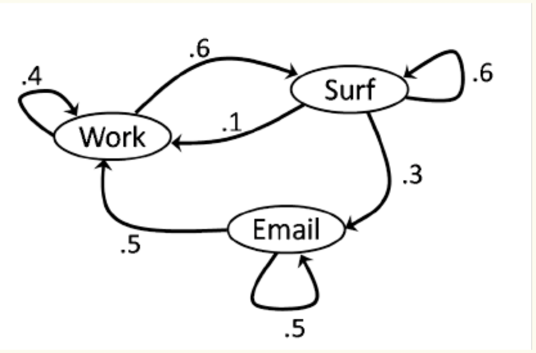
**t = 0**


This graph defines a **Markov Chain**

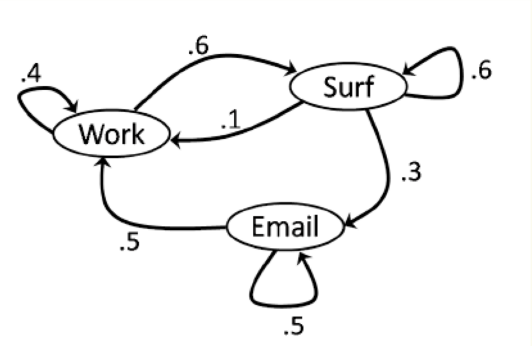
**Special Property: The random process at each step is the same!**




For ANY $t \geq 0$,
if I was working at time t, then at t+1
with probability 0.4 I continue working
with probability 0.6, I move to surfing, and
with probability 0, I move to emailing

**This is called History Independent (similar to memoryless)**

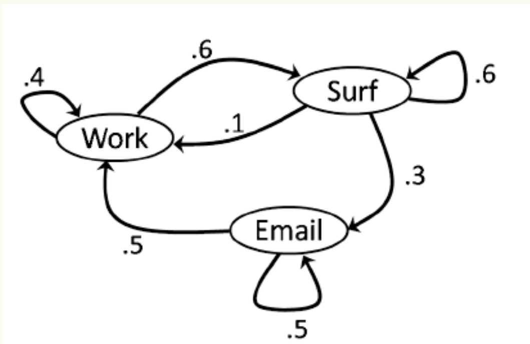
# A boring day in pandemic




**Many interesting questions?**

1. What is the probability that I work at time 1?
2. What is the probability that I work at time 2?
3. What is the probability that I work at time t=100?
4. What is the probability that I work at time $t \to \infty$? Does it always converge?

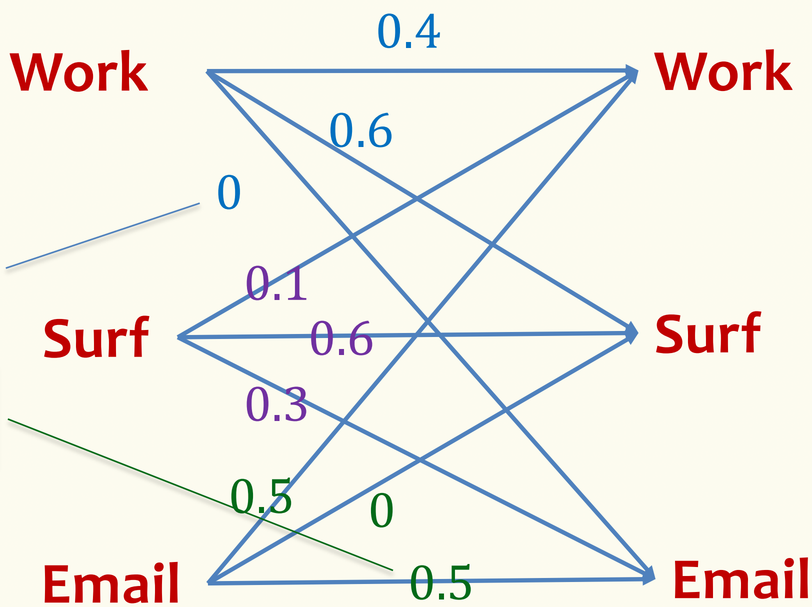
# Formalizing Markov Chain




$X^{(t)}$ my state at t $X^{t+1}$ my state at t+1

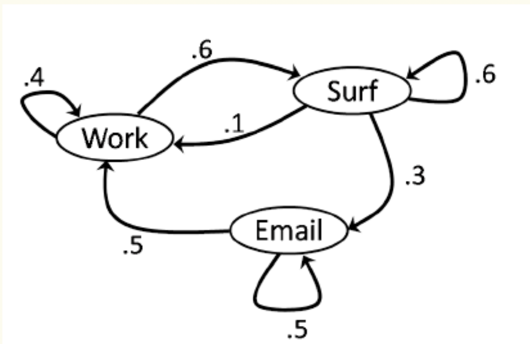
$p_{WE} = \mathrm{P}(X^{(t+1)} = E \mid X^t = W) = 0$

$p_{EE} = \mathrm{P}(X^{(t+1)} = E \mid X^t = E) = 0.5$

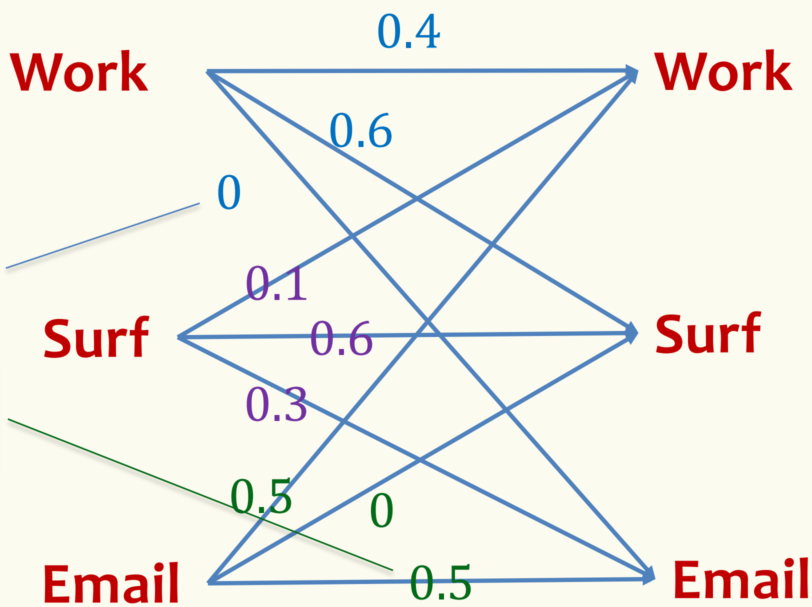



1. What is the probability that I work at time 1? $p_W^{(1)} = P(X^{(1)} = W)$
2. What is the probability that I work at time 2? $p_W^{(2)} = P(X^{(2)} = W)$

# Formalizing Markov Chain




$\boldsymbol{X^{(t)}}$ my state at t $\boldsymbol{X^{t+1}}$ my state at t+1




$p_{WE} = \mathrm{P}(X^{(t+1)} = E \mid X^t = W) = 0$

$p_{EE} = \mathrm{P}(X^{(t+1)} = E \mid X^t = E) = 0.5$

3. What is the prob that I work at t+1 = 100?

**By LTP:** $p_W^{(t+1)} = P(X^{(t+1)} = W) = \Sigma_{U \in \{W,S,E\}} \mathrm{P}\big(X^{(t+1)} = W \big| \mathrm{X}^{(t)} = U\big)\, \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$

# Vectors and Matrixes

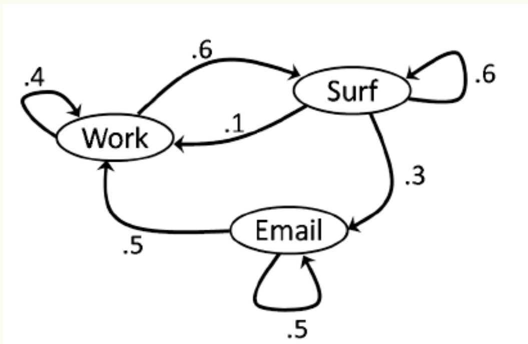
$$\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \\ c_7 & c_8 & c_9 \end{bmatrix}$$

# Vectors and Matrixes

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1p} + \cdots + a_{mn}b_{np} \end{pmatrix}$$

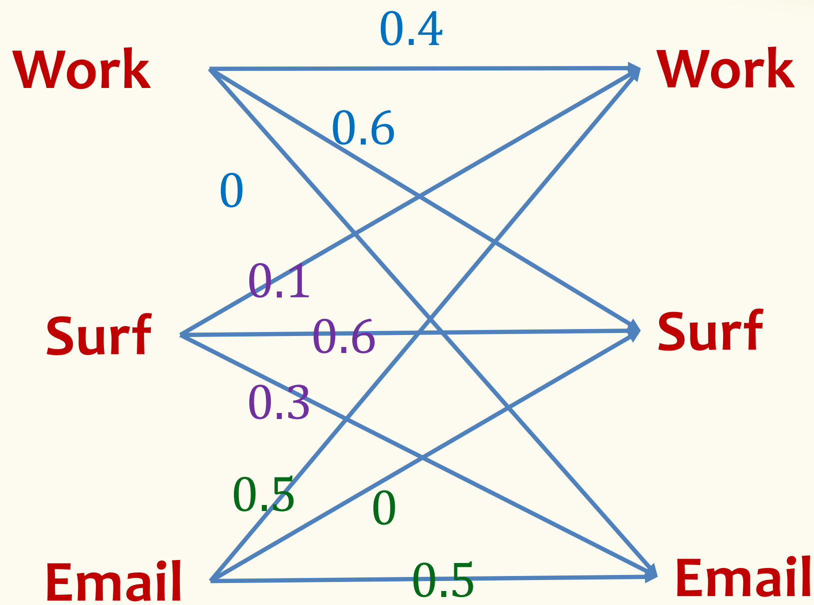
# Transition Matrix




$$P = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

$$X^{(t)} = (p_w^{(t)} \quad p_S^{(t)} \quad p_E^{(t)})$$

$\boldsymbol{X^{(t)}}$ my state at t $\boldsymbol{X^{t+1}}$ my state at t+1




By LTP: $p_W^{(t+1)} = P(X^{(t+1)} = W) = \Sigma_{U \in \{W,S,E\}} \mathrm{P}\left(X^{(t+1)} = W \middle| \mathrm{X}^{(t)} = U\right) \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$

➔ $X^{(t+1)} = X^{(t)} P$

## Transition Matrix

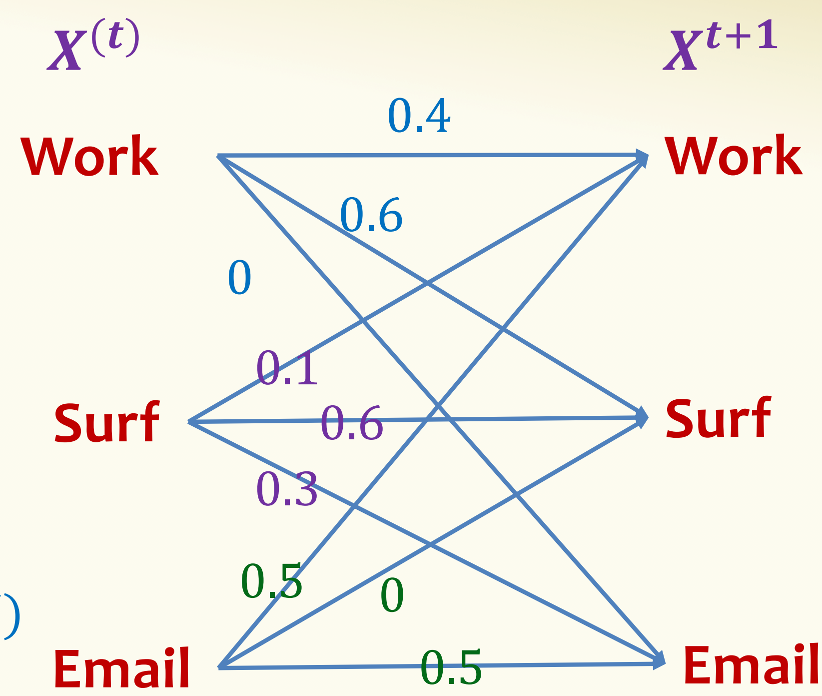
$$P = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

$$X^{(t)} = (p_w^{(t)} \quad p_S^{(t)} \quad p_E^{(t)})$$

LTP: $p_W^{(t+1)} = P\left(X^{(t+1)} = W\right)$

$= \Sigma_{U\in\{W,S,E\}} \mathrm{P}\left(X^{(t+1)} = W \middle| \mathrm{X}^{(t)} = U\right) \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$

➔ $X^{(t+1)} = X^{(t)} P$




$$\left(p_w^{(t+1)} \quad p_S^{(t+1)} \quad p_E^{(t+1)}\right) = (p_w^{(t)} \quad p_S^{(t)} \quad p_E^{(t)}) \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

## Transition Matrix

$$P = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

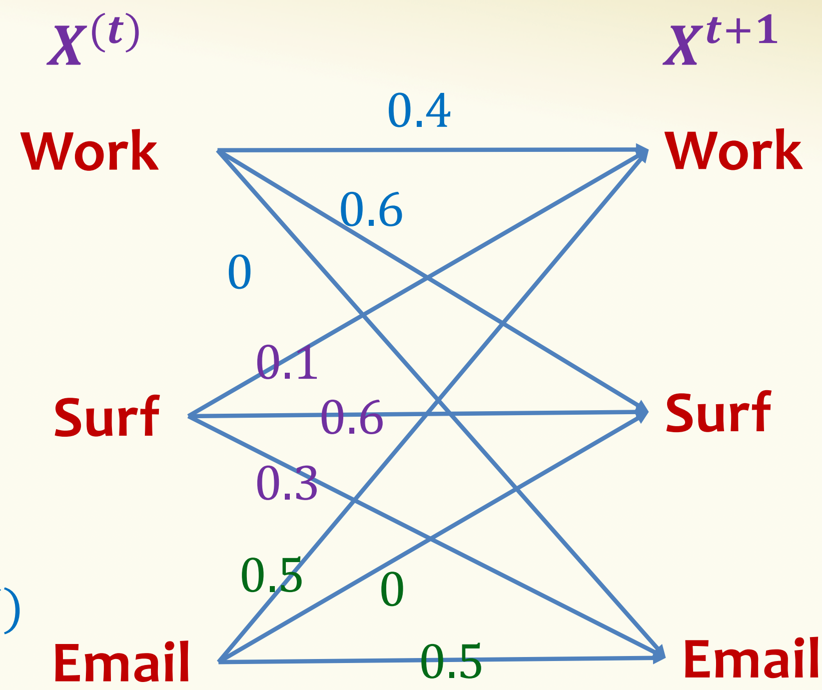
$$X^{(t)} = (p_w^{(t)} \quad p_S^{(t)} \quad p_E^{(t)})$$

LTP: $p_W^{(t+1)} = P\left(X^{(t+1)} = W\right)$

$= \Sigma_{U \in \{W,S,E\}} \mathrm{P}\left(X^{(t+1)} = W \middle| \mathrm{X}^{(t)} = U\right) \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$

➔ $X^{(t+1)} = X^{(t)} P$

➔ $X^{(t)} = X^{(0)} P^t$




3. What is the prob that I work at t = 100?

Closed formula: $p_W^{(t)} = X^{(t)}[1] = (X^{(0)} P^t)\ [1]$

# Transition Matrix

$$P = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

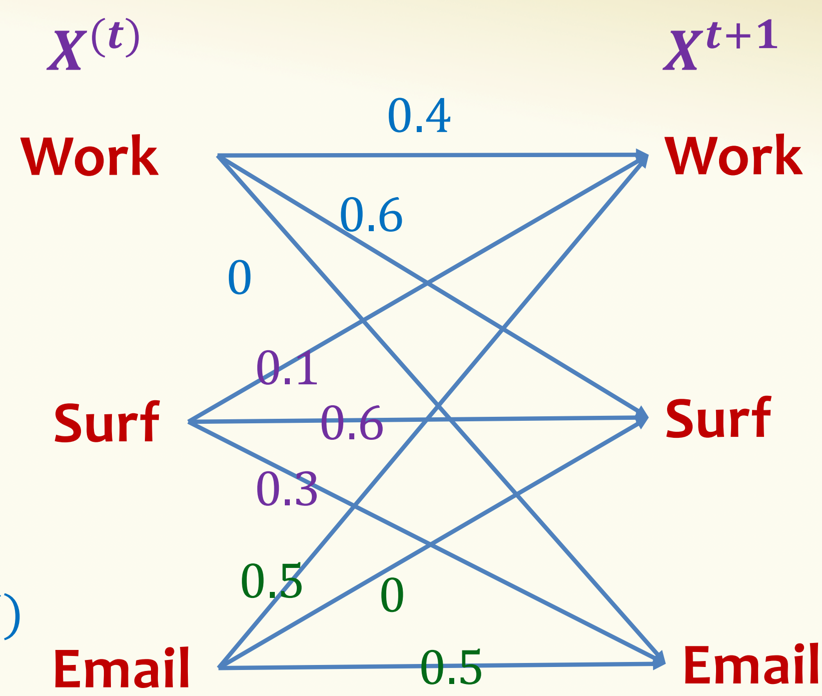
$$X^{(t)} = (p_w^{(t)} \quad p_S^{(t)} \quad p_E^{(t)})$$

LTP: $p_W^{(t+1)} = P\left(X^{(t+1)} = W\right)$

$= \Sigma_{U \in \{W,S,E\}} \mathrm{P}\left(X^{(t+1)} = W \middle| \mathrm{X}^{(t)} = U\right) \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$

➔ $X^{(t+1)} = X^{(t)} P$

➔ $X^{(t)} = X^{(0)} P^t$




4. What is the probability that I work at time $t \to \infty$?
Does it always converge?

Poll: A. Yes it converges (orderly universe)
B. No, it does not converge (anarchic universe)

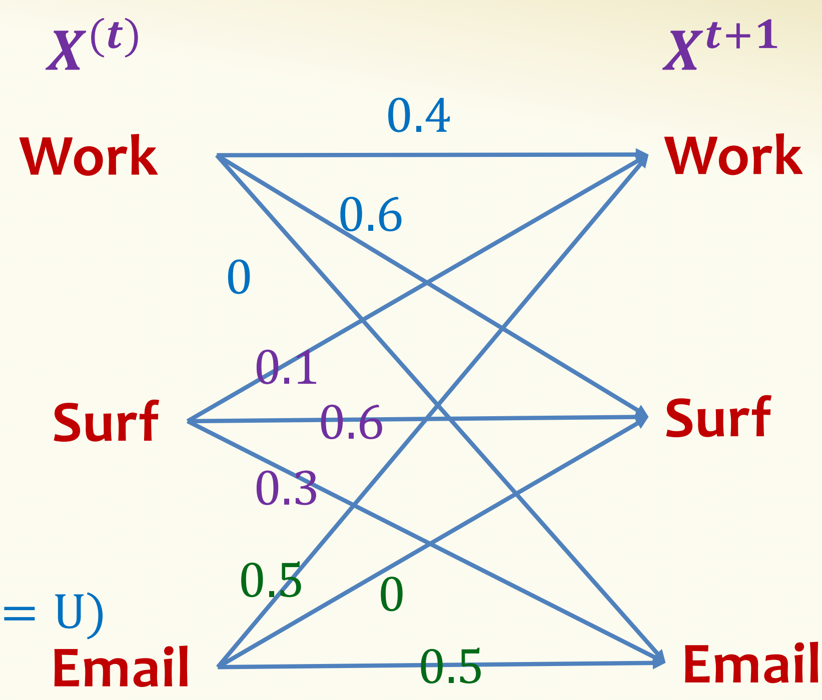
# Stationary Distribution

$$P = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

$$X^{(t)} = (p_w^{(t)} \quad p_S^{(t)} \quad p_E^{(t)})$$

LTP: $p_W^{(t+1)} = P\left(X^{(t+1)} = W\right)$

$= \Sigma_{U \in \{W,S,E\}} \mathrm{P}\left(X^{(t+1)} = W \middle| \mathrm{X}^{(t)} = U\right) \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$




➔ $X^{(t+1)} = X^{(t)}\ P$

➔ $X^{(t)} = X^{(0)}\ P^t$

4. What is the probability that I work at time $t \rightarrow \infty$?
Does it always converge?

➔ As t $\rightarrow \infty$, $X^{(t)} \rightarrow \pi$ the stationary distribution

# Solving for Stationary Distribution

$$P = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix}$$

*Constraints on Stationary Distribution*

- $\pi \;=\; \pi P$
- $\pi[1] + \pi[2] + \pi[3] = 1$

➔ $\pi[1] = \frac{15}{34}, \quad \pi[2] = \frac{10}{34}, \quad \pi[3] = \frac{9}{34}$

LTP: $p_W^{(t+1)} = P\left(X^{(t+1)} = W\right)$

$= \Sigma_{U \in \{W,S,E\}} \mathrm{P}\left(X^{(t+1)} = W \middle| \mathrm{X}^{(t)} = U\right) \mathrm{P}(\mathrm{X}^{(t)} = \mathrm{U})$

➔ $X^{(t+1)} = X^{(t)}\, P$

➔ $X^{(t)} = X^{(0)}\, P^t$

➔ As $t \to \infty$, $X^{(t)} \to \pi$ the stationary distribution

## General Markov Chain

- A set of $n$ states $\{1, 2, 3, \dots\ n\}$
- The state at time t $X^{(t)}$
- A transition matrix P, dimension $n \times n$

$$P[i,j] = \Pr\left(X^{(t+1)} = j \middle| \mathrm{X}^{(t)} = i\right)$$

- Transition: LTP ➔ $X^{(t+1)} = X^{(t)}\ P \implies X^{(t)} = X^{(0)}\ P^t$
- A stationary distribution $\pi$ is the solution to:

$$\pi \;=\; \pi P,\ \text{normalized so that } \Sigma_{i \in [n]} \pi[i] = 1$$

## The Fundamental Theorem of Markov Chain

If a Markov Chain is Irreducible and aperiodic, then it has a unique stationary distribution.

Moreover, $t \rightarrow \infty$, for all $i, j, P^t[i, j] \rightarrow \pi[j]$