String Algorithms and Data Structures
Markov Chains and HMMs

CS 199-225
Brad Solomon

April 16, 2026

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

Department of Computer Science

Learning Objectives

Introduce State Diagrams and Markov Chains

Identify how Markov chains can be used to:

- Estimate probabilities of sequences
- Identify more probable labels
- Predict future states

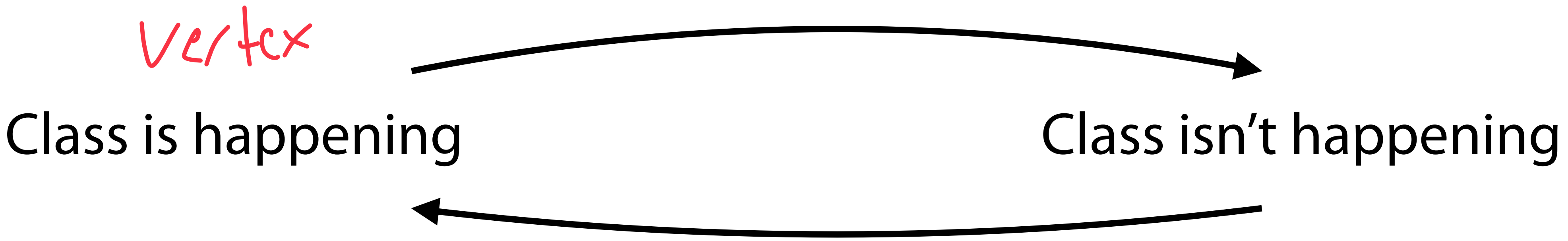
Define and determine stationary states

Introduce Hidden Markov Models and the Viterbi Algorithm

Modeling events with State Diagrams

↙ mp puzzle

A **state diagram** is a (usually weighted) directed graph where nodes are states and edges are transitions between them



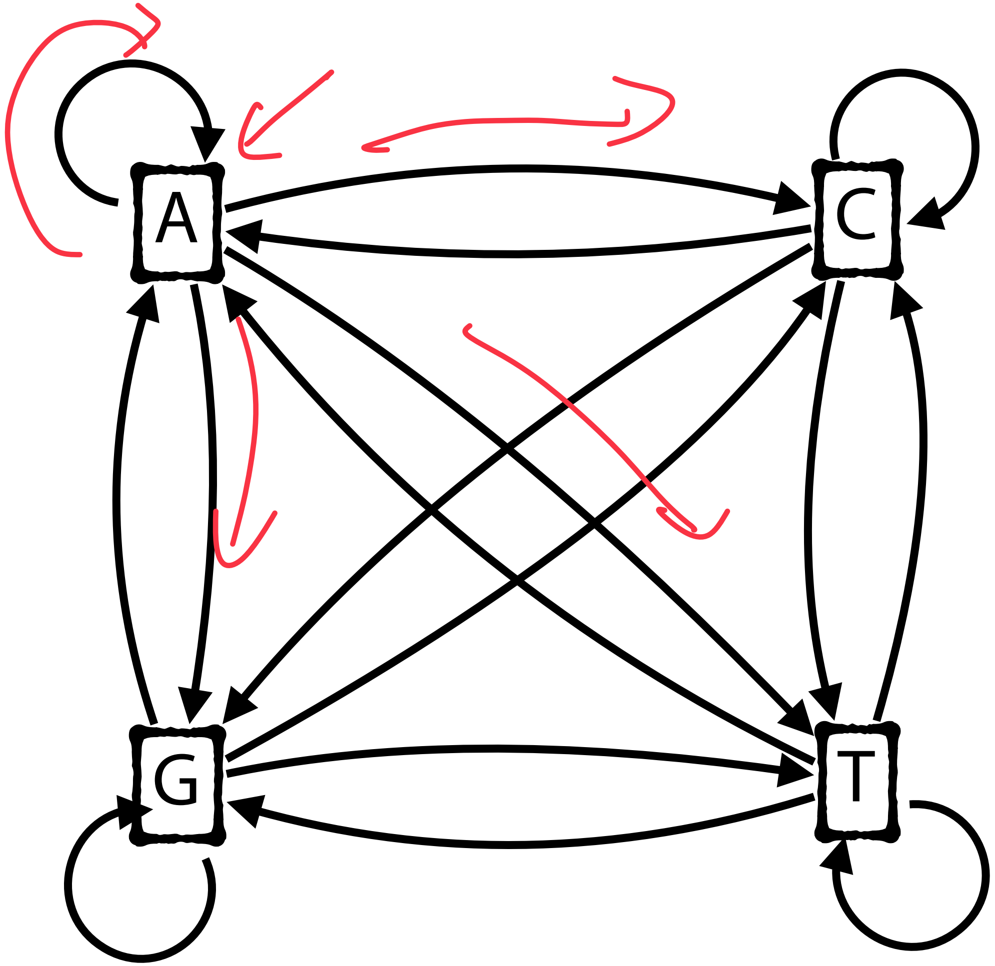
Vertex

Class is happening

Class isn't happening

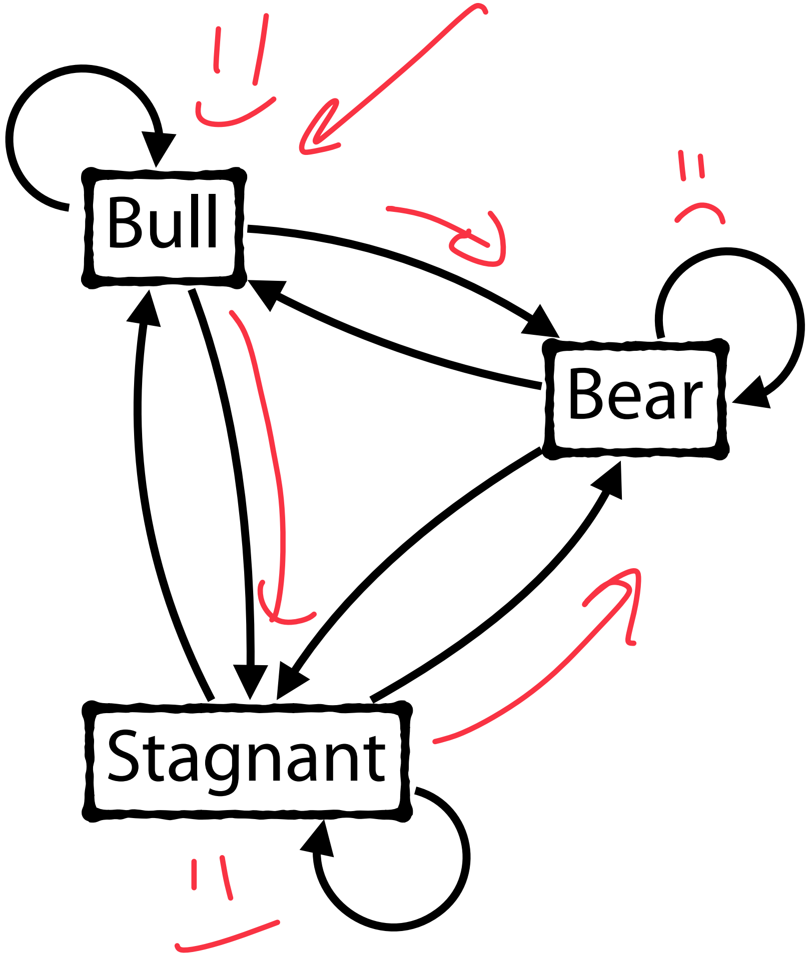
These diagrams are very useful in modeling many real world scenarios!

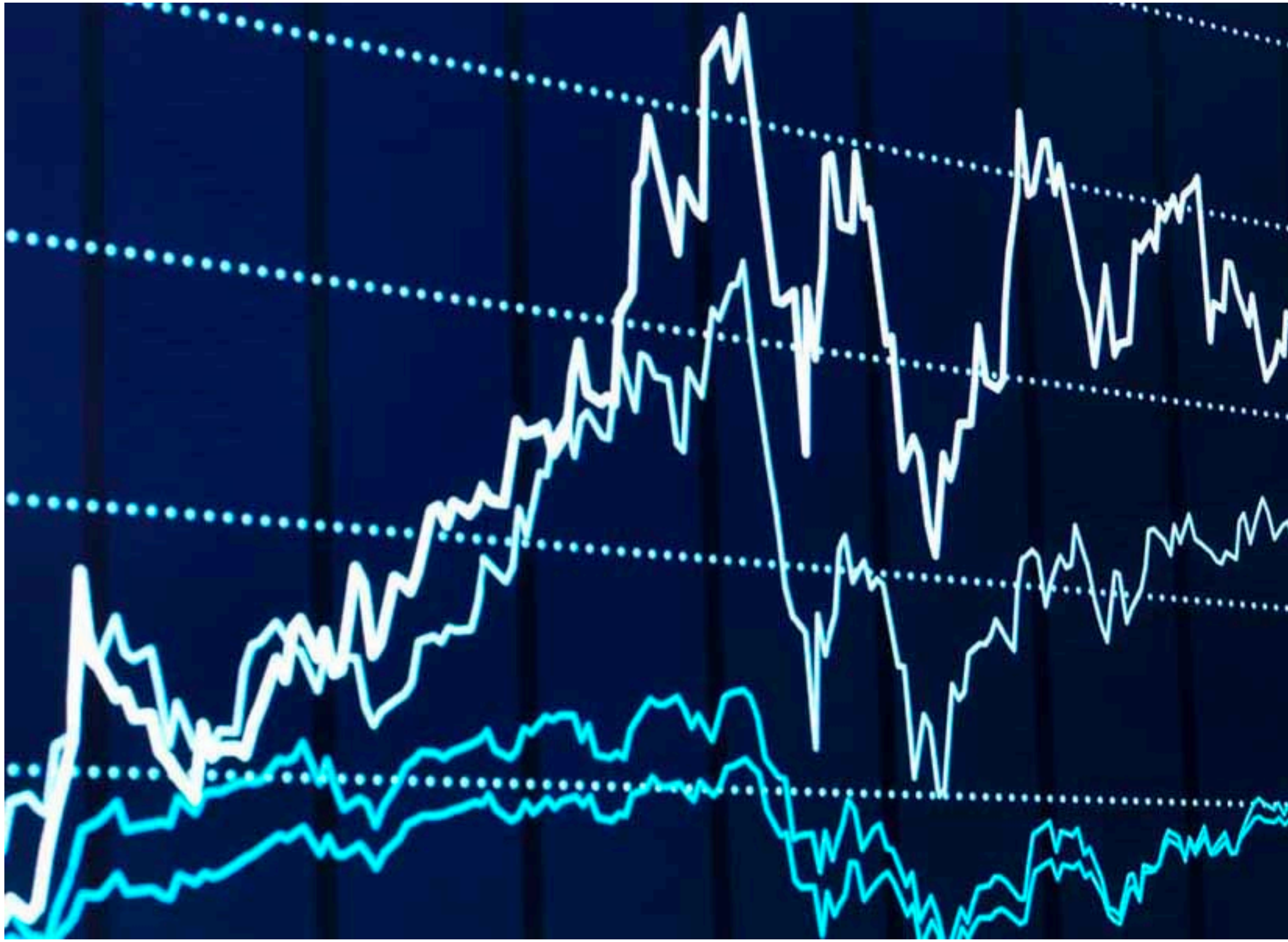
Sequence Modeling in Biology



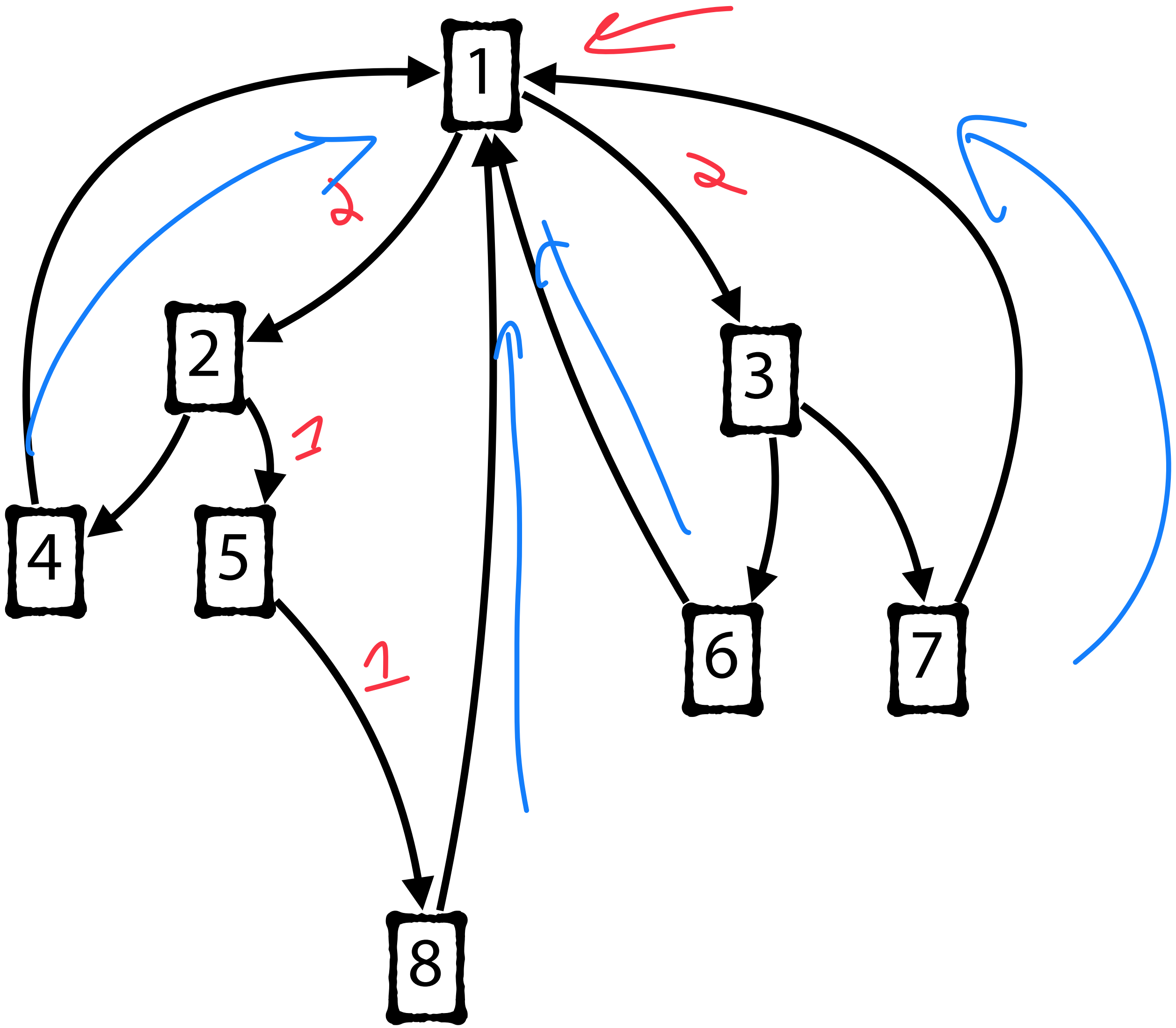
```
CATGACGTCGCGGACAACCCAGAATTGTCTTGAGCGATGGTAAGATCTAACCTCACTGC
CTGGGGCTTTACTGATGTCATACCGTCTTGCACGGGGATAGAATGACGGTGCCCGTGTC
ATTTTCTGAAAGTTACAGACTTCGATTAAAAAGATCGGACTGCGCGTGGGCCCGGAGAG
TTTTTCGACGTGTCAAGGACTCAAGGGAATAGTTTGGCGGGAGCGTTACAGCTTCAATT
CGATAAAATTCAACTACTGGTTTCGGCCTAATAGGTCACGTTTTATGTGAAATAGAGGG
CCCTGGGTGTTCTATGATAAGTCCTGCTTTATAACACGGGGCGGTTAGGTTAAATGACT
ATCCAAGCGCCCGCTAATTCTGTTCTGTTAATGTTCATACCAATACTCACATCACATTA
AGCCCAGTCGCAAGGGTCTGCTGCTGTTGTCGACGCCTCATGTTACTCCTGGAATCTAC
GGTTAAGGCGTGTGATCGACGATGCAGGTATACATCGGCTCGGACCTACAGTGGTCGAT
TCGCGGTTCGGCGCGTAGTTGAGTGCGATAACCCAACCGGTGGCAAGTAGCAAGAAGAC
AGACAACCTAACTAATAGTCTCTAACGGGGAATTACCTTTACCAGTCTCATGCCTCCAA
CAATGATATCGCCCACAGAAAGTAGGGTCTCAGGTATCGCATACGCCGCGCCCGGGTCC
GACAGTAGAGAGCTATTGTGTAATTCAGGCTCAGCATTCATCGACCTTTCCTGTTGTGA
TCTCGTCCGTAACGATCTGGGGGGCAAAACCGAATATCCGTATTCTCGTCCTACGGGTC
TGCGCGTGATCGTCAGTTAAGTTAAATTAATTCAGGCTACGGTAAACTTGTAGTGAGCT
ACGGGTTCGCTACAGATGAACTGAATTTATACACGGACAACTCATCGCCCATTTGGGCG
AAAGTGGCAGATTAGGAGTGCTTGATCAGGTTAGCAGGTGGACTGTATCCAACAGCGCA
CCAAAGCGTTGTAGTGGTCTAAGCACCCCTGAACAGTGGCGCCCATCGTTAGCGTAGTA
AGGTGCGACATGGGGCCAGTTAGCCTGCCCTATATCCCTTGCACACGTTCAATAAGAGG
TTTTTAAATTAGGATGCCGACCCCATCATTGGTAACTGTATGTTCATAGATATTTCTTC
AGCTGACACGCAAGGGTCAACAATAATTTCTACTATCACCCCGCTGAACGACTGTCTTT
CTTAGATTCGCGTCCTAACGTAGTGAGGGCCGAGTCATATCATAGATCAGGCATGAGAA
CACACGAGTTGTAAACAACTTGATTGCTATACTGTAGCTACCGCAAGGATCTCCTACAT
ATCTGGATCCGAGTCAGAAATACGAGTTAATGCAAATTTACGTAGACCGGTGAAAACAC
AGACCGTAGTCAGAAGTGTGGCGCGCTATTCGTACCGAACCGGTGGAGTATACAGAATT
AGGAGCTCGGTCCCCAATGCACGCCAAAAAAGGAATAAAGTATTCAAACTGCGCATGGT
CTATTATCCATCCGAACGTTGAACCTACTTCCTCGGCTTATGCTGTCCTCAACAGTATC
CGGCTGTGGATCTTAACGGCCACATTCTTAATTCCGACCGATCACCGATCGCCTTTCCT
ACTAAGTTATCCAGATCAAGGTTTGAACGGACTCGTATGACATGTGTGACTGAACCCGG
CTGTTTCAAGGCCTCTGCTTTGGTATCACTCAATATATTCAGACCAGACAAGTGGCAAA
CTAGGTATTCACGCAACCGTCGTAACATGCACTAAGGATAACTAGCGCCAGGGGGGCAT
AAAGACTACCCTATGGATTCCTTGGAGCGGGGACAATGCAGACCGGTTACGACACAATT
GGTATTATTAGCAAGACAATAAAGGACATTGCACAGAGACTTATTAGAATTCAACAAAC
GTGTTGGGTCGGGCAAGTCCCCGAAGCTCGGCCAAAAGATTCGCCATGGAACCGTCTGG
```

Market Trends in Economics





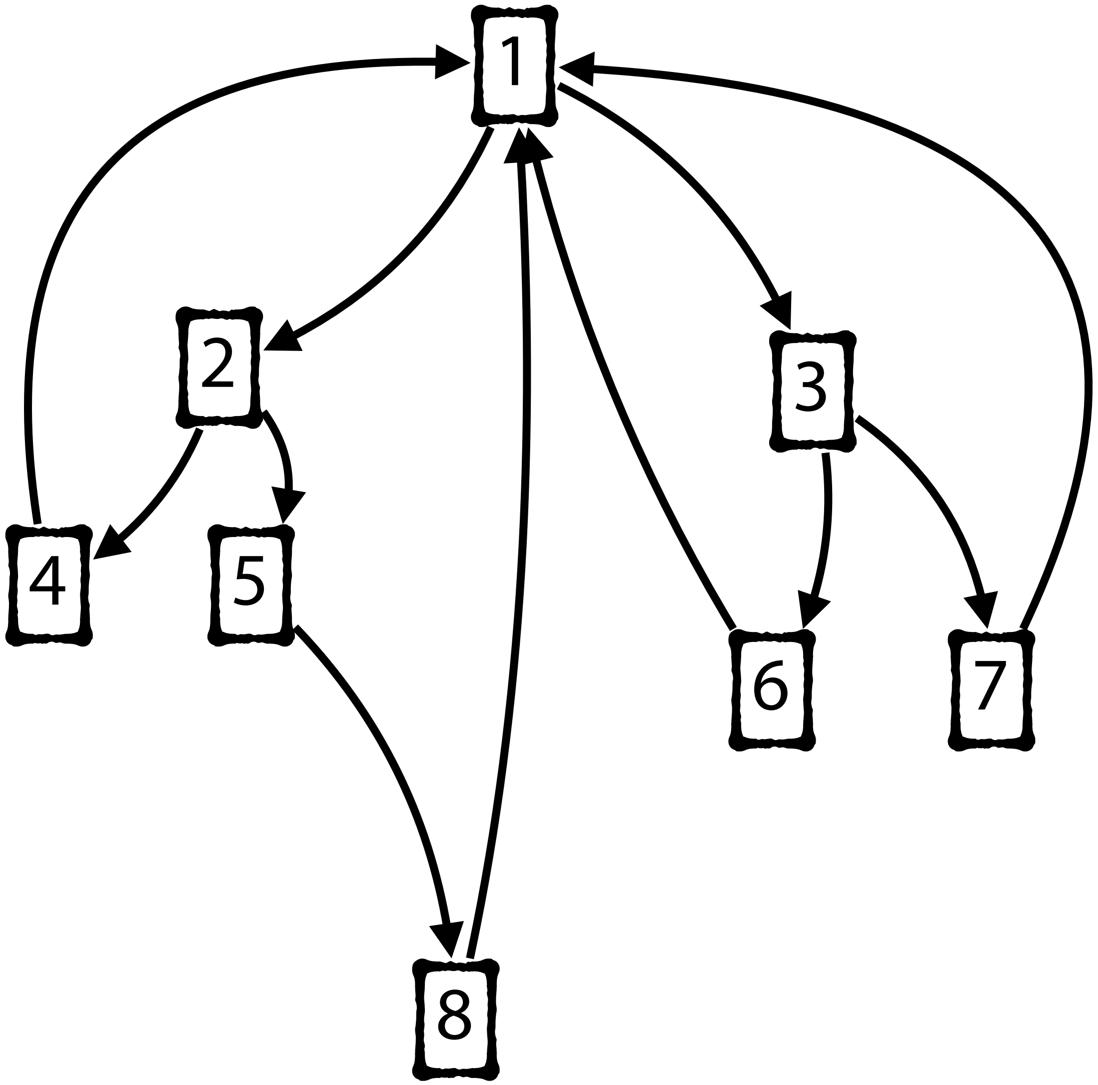
PageRank in Graphs



Equilibrium State

1: 4
2: 2
3: 2
4: 1
5: 1
6: 1
7: 1
8: 1

PageRank in Graphs



Equilibrium State

1: 4/13

2: 2/13

3: 2/13

4: 1/13

5: 1/13

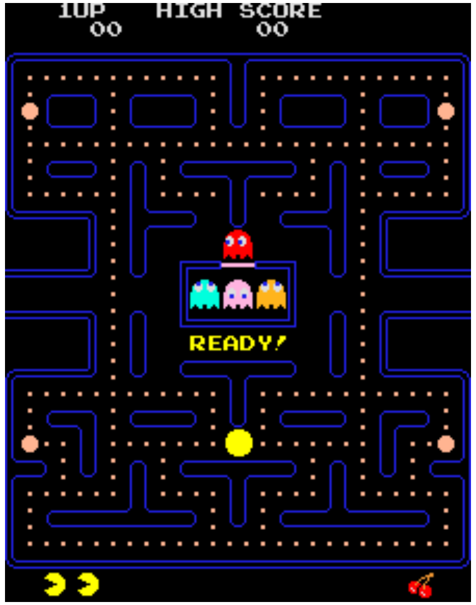
6: 1/13

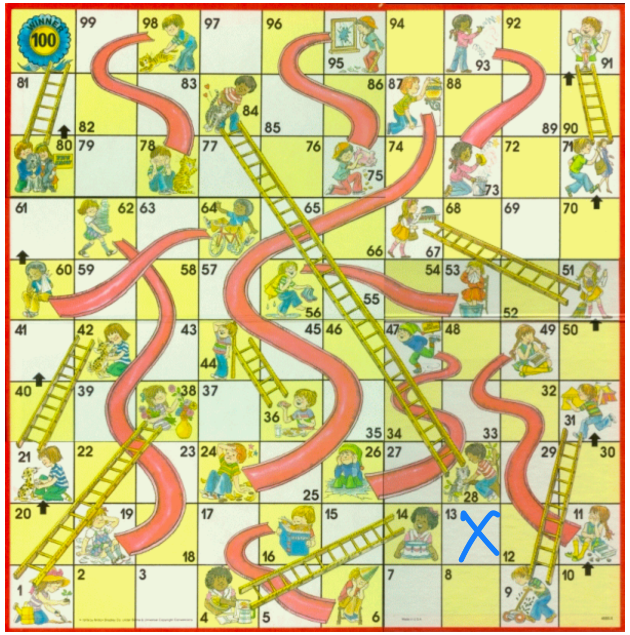
7: 1/13

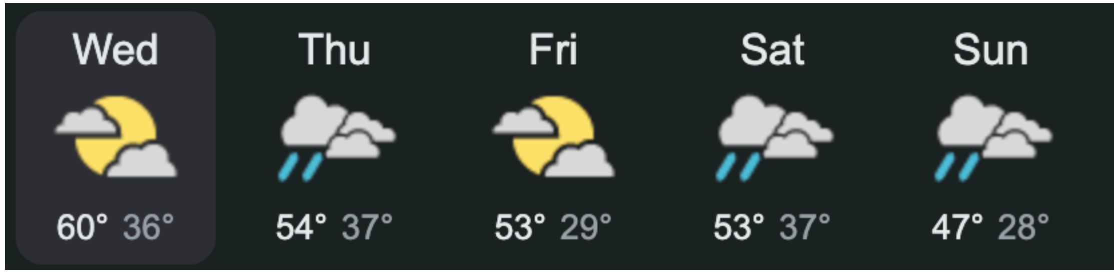
8: 1/13

Markov Assumption

The probability of the next state depends only on our current state





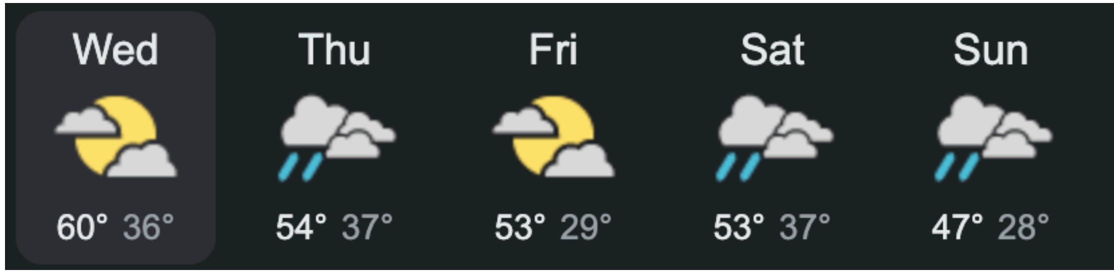


Markov Assumption

Probability of state x_k depends only on previous state x_{k-1}

5

Ex: Let $x = \{C, R, C, R, R\}$



$$P(x) = P(x_k, x_{k-1}, \ldots x_1)$$

$$= P(x_k \mid x_{k-1}, \ldots x_1) P(x_{k-1}, \ldots x_1)$$

$$= P(x_k \mid x_{k-1}, \ldots x_1) P(x_{k-1} \mid x_{k-2}, \ldots x_1) \ldots P(x_2 \mid x_1) P(x_1)$$

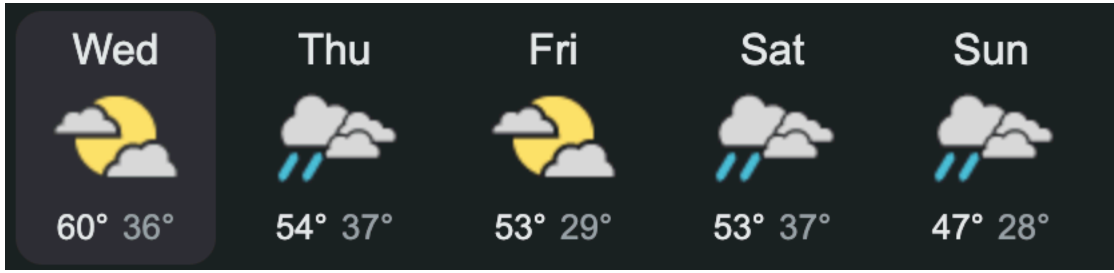
every things before me!

$$P(x) \approx$$

Markov Assumption

Probability of state x_k depends only on previous state x_{k-1}

Ex: Let $x = \{C, R, C, R, R\}$



Wed	Thu	Fri	Sat	Sun
60° 36°	54° 37°	53° 29°	53° 37°	47° 28°

$$P(x) = P(x_k, x_{k-1}, \ldots x_1)$$

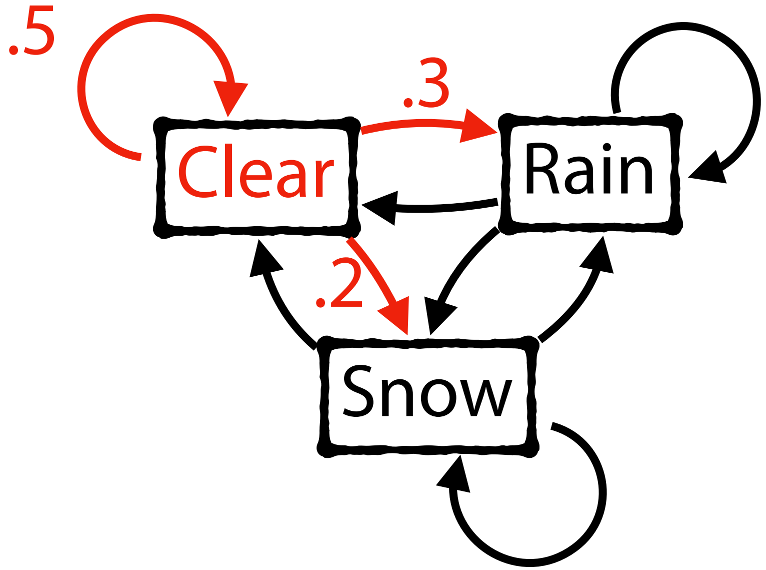
$$= P(x_k \mid x_{k-1}, \ldots x_1) P(x_{k-1}, \ldots x_1)$$

$$= P(x_k \mid x_{k-1}, \ldots x_1) P(x_{k-1} \mid x_{k-2}, \ldots x_1) \ldots P(x_2 \mid x_1) P(x_1)$$

$$P(x) \approx \underline{P(x_k \mid x_{k-1})} P(x_{k-1} \mid x_{k-2}) \underline{\ldots} \underline{P(x_2 \mid x_1)} \underline{P(x_1)}$$

Markov Chain

A **finite Markov Chain** has a set of states S and a finite matrix M

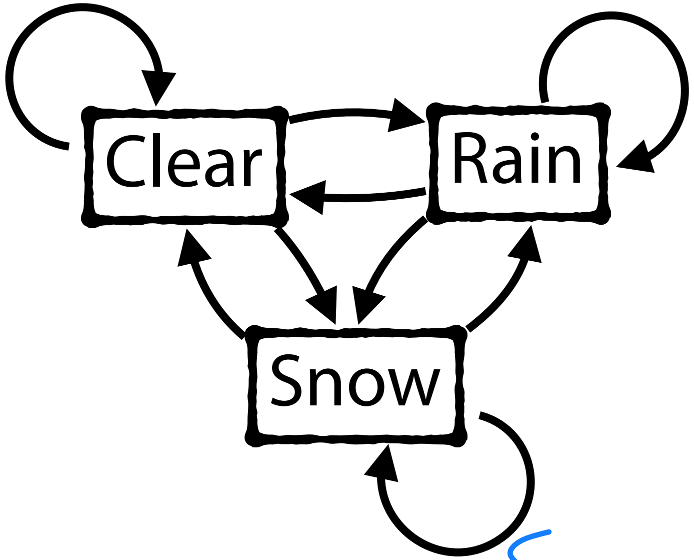


$$S = \{Clear,\ Rain,\ Snow\}$$

$$M = \begin{array}{c|ccc} & C & R & S \\ \hline C & \underline{.5} & .3 & .2 \\ R & \underline{.5} & \underline{.4} & .1 \longleftarrow \\ S & .2 & .1 & \enclose{circle}{.7} \end{array}$$

Markov Chain

Given a Markov Chain and an initial state, all subsequent states can be represented either as **a series of random states** or a transition probability.



$$M = \begin{array}{cccc} & & & S \\ & .5 & .3 & .2 \\ & .5 & .4 & .1 \\ S & .2 & .1 & \boxed{.7} \end{array}$$

We can compute probability directly:

$X_0 = \text{Clear}$ p = 1

0,5

$X_1 = \text{Clear}$

0.2

$X_2 = \text{Snow}$

0,7

$X_3 = \text{Snow}$

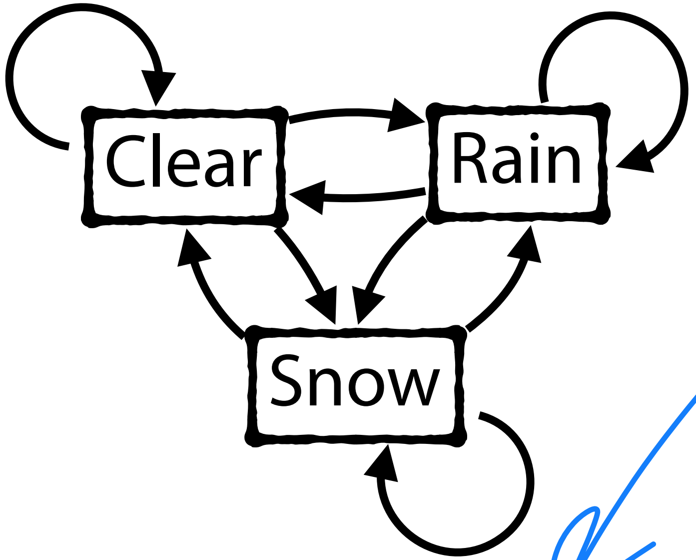
0,7

$X_4 = \text{Snow}$

Multiply all this

Markov Chain

Given a Markov Chain and an initial state, all subsequent states can be represented either as a series of random states or a **transition probability.**



We compute a new transition probability:

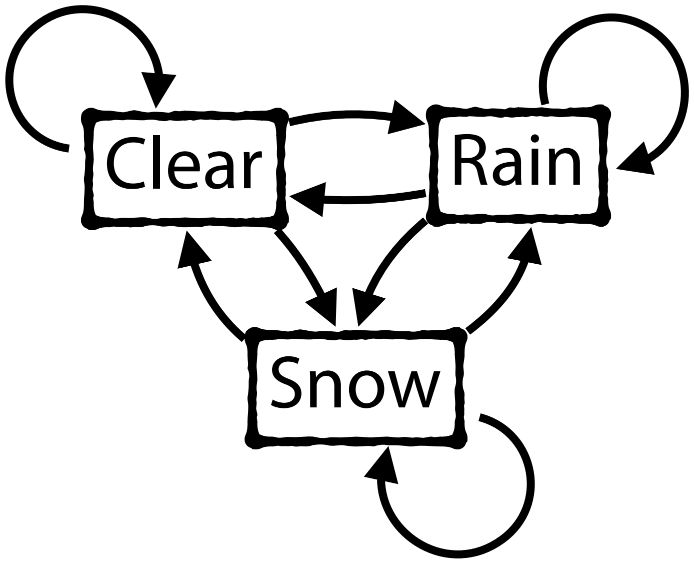
$$M_0 = (.4 \quad .3 \quad .3) \qquad \begin{pmatrix} .5 \\ .5 \\ .2 \end{pmatrix}$$

$$M_{1c} = .4 * .5 + .3 * .5 + .3 * .2 = 0.41$$

$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix}$$

Markov Chain

Given a Markov Chain and an initial state, all subsequent states can be represented either as a series of random states or a **transition probability.**



$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix}$$

We compute a new transition probability:

$$M_0 = (.4 \quad .3 \quad .3)$$

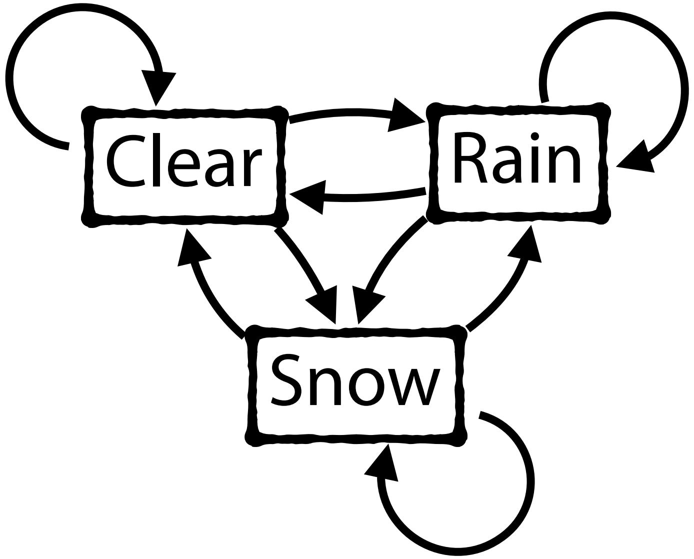
$$M_{1c} = .4 * .5 + .3 * .5 + .3 * .2 = 0.41$$

$$M_{1r} = .4 * .3 + .3 * .4 + .3 * .1 = 0.27$$

$$M_{1s} = .4 * .2 + .3 * .1 + .3 * .7 = 0.32$$

Markov Chain

Given a Markov Chain and an initial state, all subsequent states can be represented either as a series of random states or a **transition probability.**



$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix}$$

We compute a new transition probability:

$$M_0 = (.4 \quad .3 \quad .3)$$

$$M_1 = (.41 \quad .27 \quad .32)$$

$$M_2 = (.404 \quad .263 \quad .333)$$

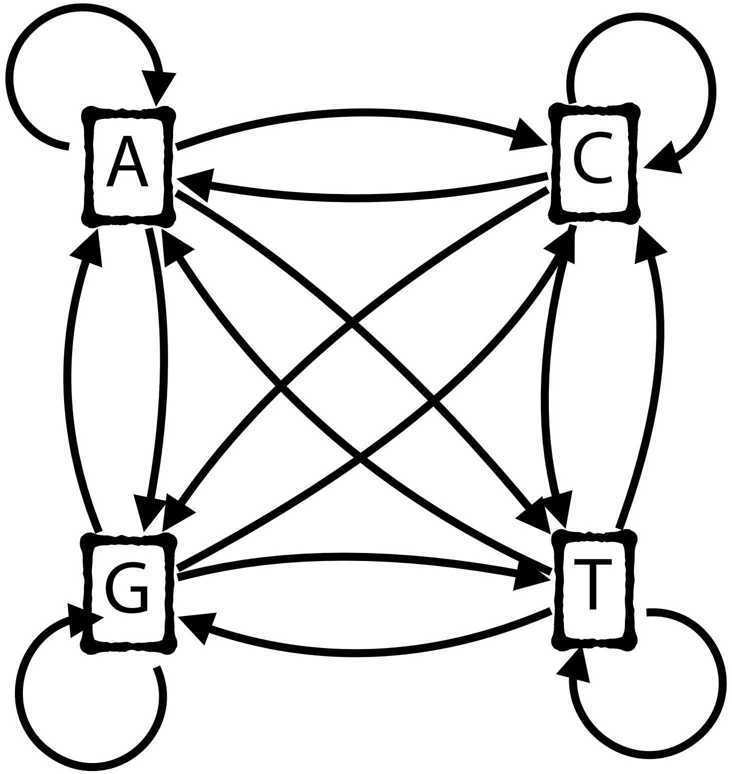
$$M_3 = (.401 \quad .259 \quad .340)$$

$$M_\infty = (0.345, \ 0.256, \ 0.349)$$

Steady State!

Markov Chain in Sequencing

Given a set of sequences, we can construct a model of transitions



next

P(A | A) = # times AA occurs / # times AX occurs

P(C | A) = # times AC occurs / # times AX occurs

P(G | A) = # times AG occurs / # times AX occurs

P(T | A) = # times AT occurs / # times AX occurs

P(A | C) = # times CA occurs / # times CX occurs

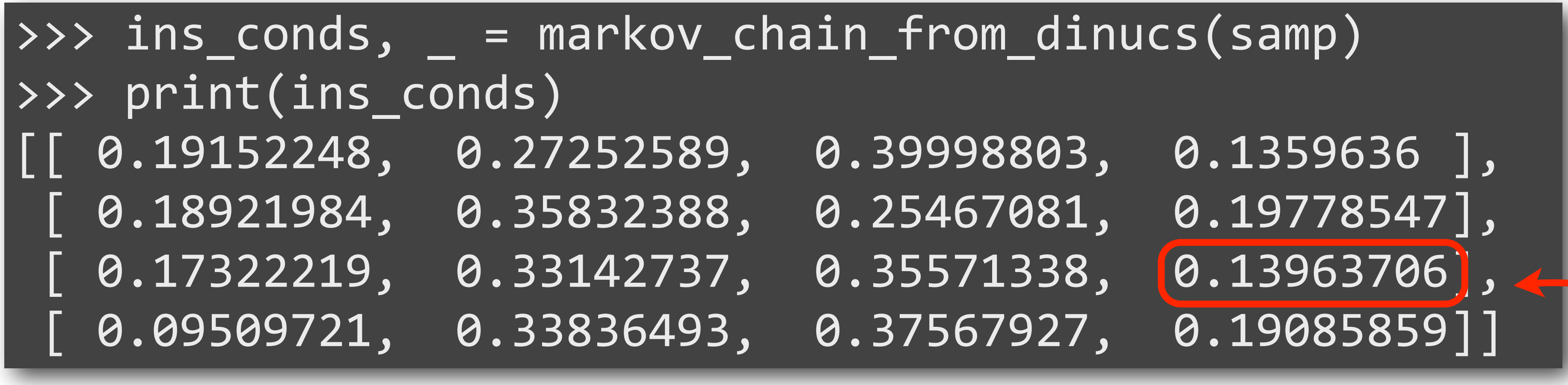
(etc)

where X is any base

Example by Ben Langmead

Markov Chain in Sequencing

Given a set of sequences, we can construct a model of transitions

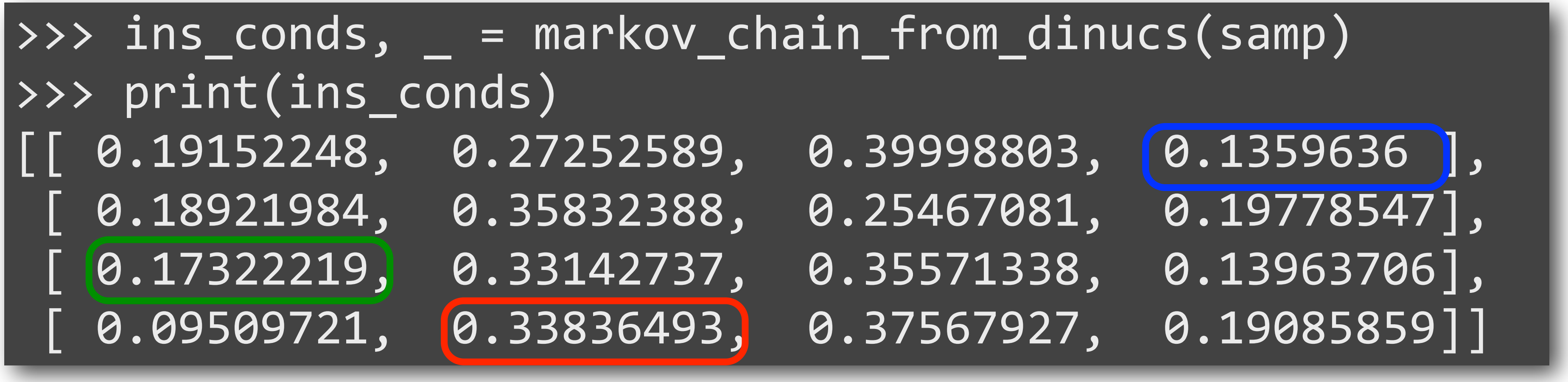


```
>>> ins_conds, _ = markov_chain_from_dinucs(samp)
>>> print(ins_conds)
[[ 0.19152248,  0.27252589,  0.39998803,  0.1359636 ],
 [ 0.18921984,  0.35832388,  0.25467081,  0.19778547],
 [ 0.17322219,  0.33142737,  0.35571338,  0.13963706],
 [ 0.09509721,  0.33836493,  0.37567927,  0.19085859]]
```

$P(T|G)$

Example by Ben Langmead

Markov Chain in Sequencing



```
>>> ins_conds, _ = markov_chain_from_dinucs(samp)
>>> print(ins_conds)
[[ 0.19152248,  0.27252589,  0.39998803,  0.1359636 ],
 [ 0.18921984,  0.35832388,  0.25467081,  0.19778547],
 [ 0.17322219,  0.33142737,  0.35571338,  0.13963706],
 [ 0.09509721,  0.33836493,  0.37567927,  0.19085859]]
```

x_{i-1}: A, C, G, T

A C G T

x_i

$x = \text{GATC}$

$$P(x) = P(x_4 \mid x_3)\, P(x_3 \mid x_2)\, P(x_2 \mid x_1)\, P(x_1)$$

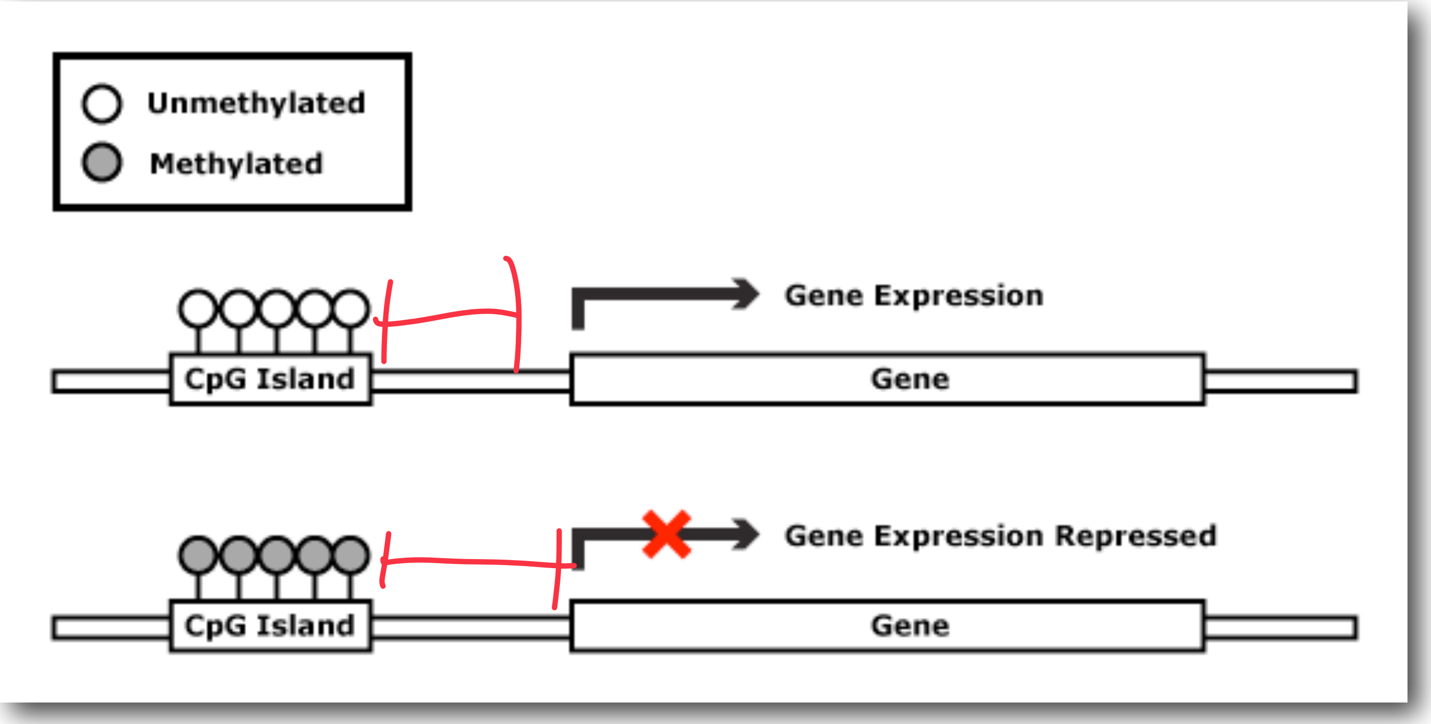
$$P(x) = P(C \mid T)\, P(T \mid A)\, P(A \mid G)\, P(G) = 0.33836493 * 0.1359636 * 0.17322219 * 0.25 = 0.001992$$

Example by Ben Langmead

Markov Chain in Sequencing

We can use this same approach to predict a *label* in our sequences as well

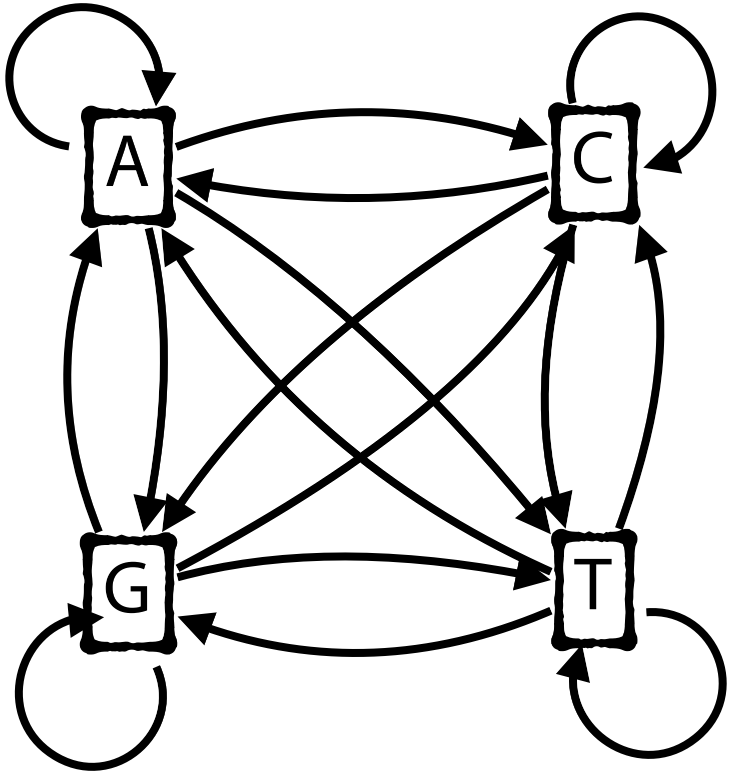
CpG island: part of the genome where CG occurs particularly frequently



Example by Ben Langmead

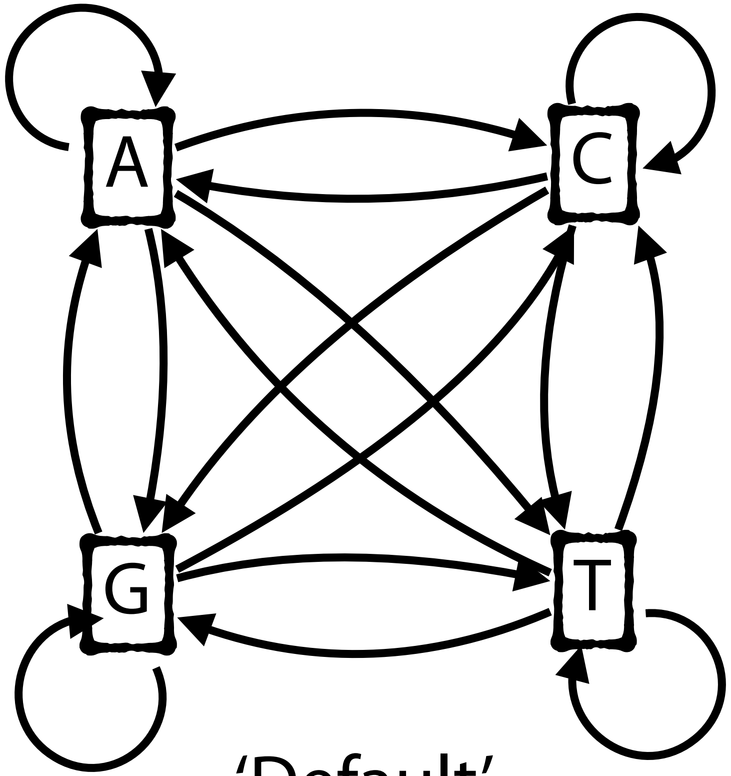
Markov Chain in Sequencing

To predict a *label* of a sequencing region, make a Markov chain for both!



CpG Island

↳ C & G more likely!

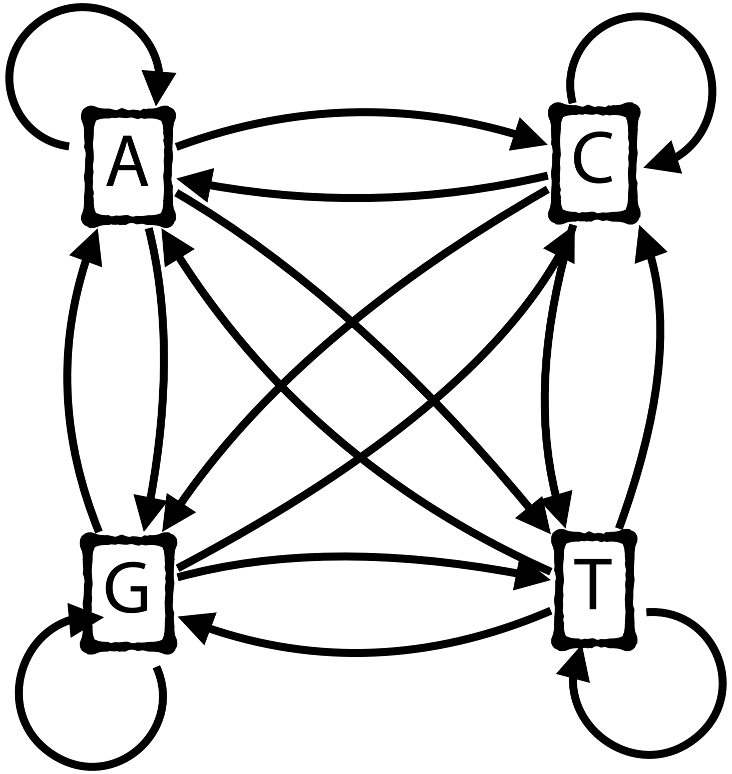


‘Default’

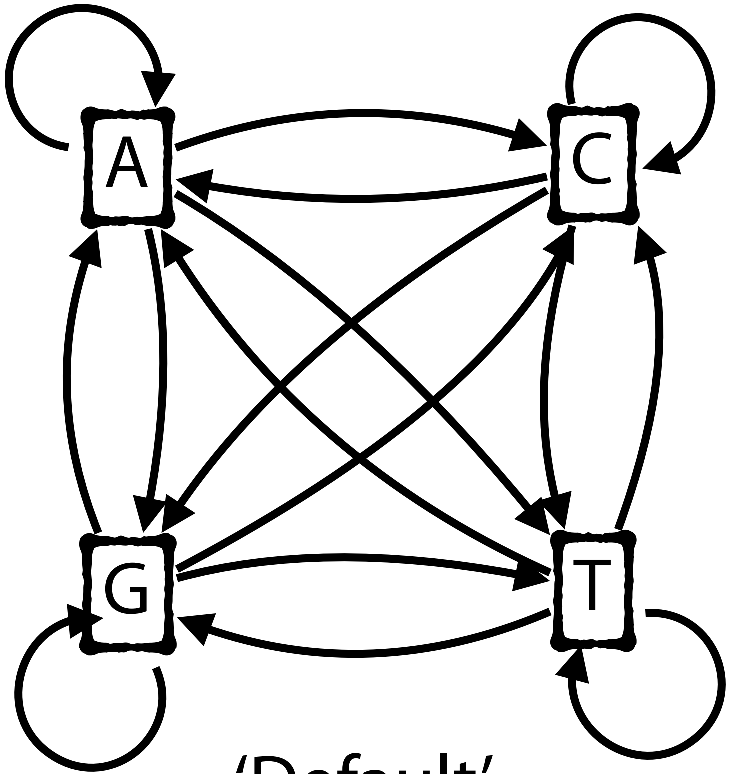
Example by Ben Langmead

Markov Chain in Sequencing

To predict a *label* of a sequencing region, make a Markov chain for both!



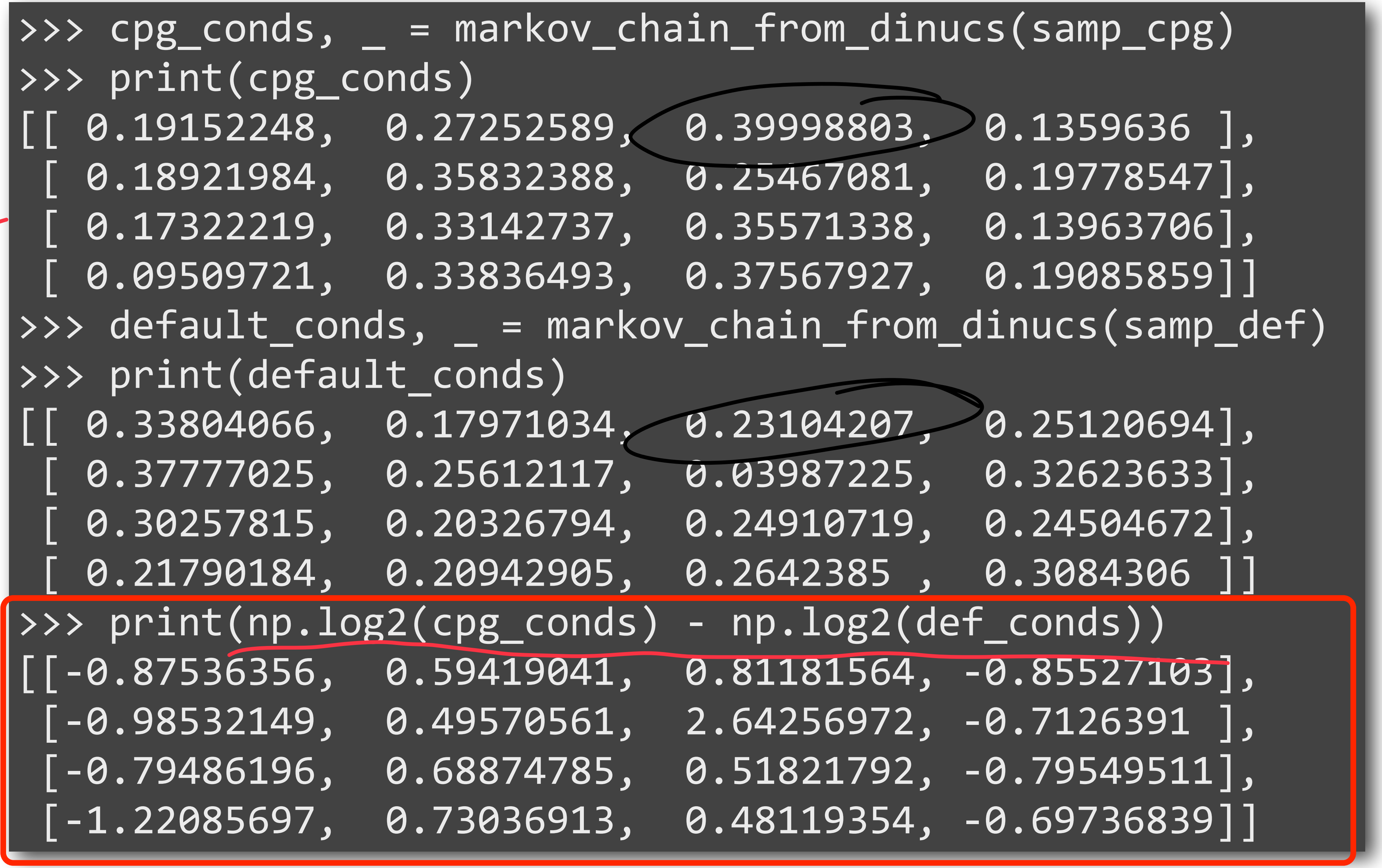
CpG Island



'Default'

Use *ratio:*
$$\frac{P(x) \text{ from CpG model}}{P(x) \text{ from Default model}}$$

Example by Ben Langmead



```
>>> cpg_conds, _ = markov_chain_from_dinucs(samp_cpg)
>>> print(cpg_conds)
[[ 0.19152248,  0.27252589,  0.39998803,  0.1359636 ],
 [ 0.18921984,  0.35832388,  0.25467081,  0.19778547],
 [ 0.17322219,  0.33142737,  0.35571338,  0.13963706],
 [ 0.09509721,  0.33836493,  0.37567927,  0.19085859]]
>>> default_conds, _ = markov_chain_from_dinucs(samp_def)
>>> print(default_conds)
[[ 0.33804066,  0.17971034,  0.23104207,  0.25120694],
 [ 0.37777025,  0.25612117,  0.03987225,  0.32623633],
 [ 0.30257815,  0.20326794,  0.24910719,  0.24504672],
 [ 0.21790184,  0.20942905,  0.2642385 ,  0.3084306 ]]
>>> print(np.log2(cpg_conds) - np.log2(def_conds))
[[-0.87536356,  0.59419041,  0.81181564, -0.85527103],
 [-0.98532149,  0.49570561,  2.64256972, -0.7126391 ],
 [-0.79486196,  0.68874785,  0.51821792, -0.79549511],
 [-1.22085697,  0.73036913,  0.48119354, -0.69736839]]
```

CpG

A
C
G
T

Default

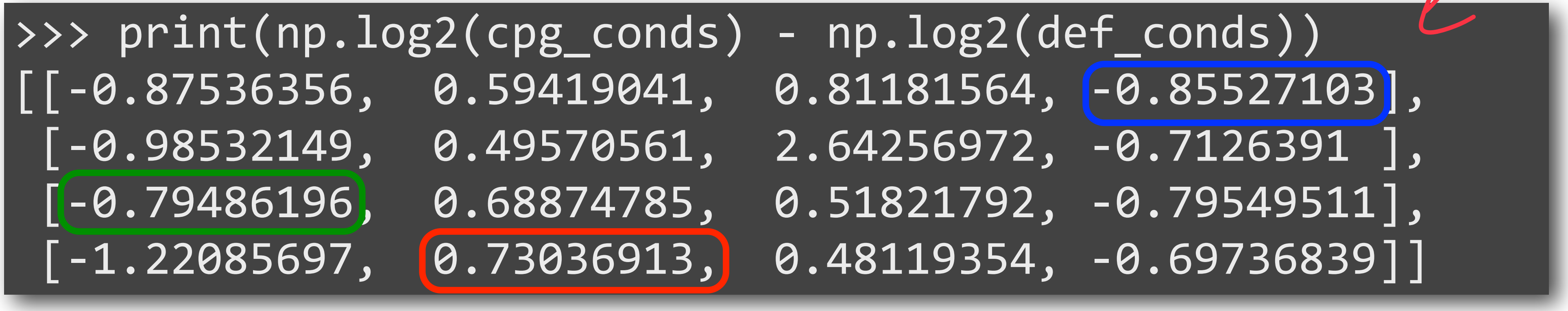
A
C
G
T

Log ratio

A
C
G
T

A **C** **G** **T**

Markov Chain in Sequencing



```
>>> print(np.log2(cpg_conds) - np.log2(def_conds))
[[-0.87536356,  0.59419041,  0.81181564, -0.85527103],
 [-0.98532149,  0.49570561,  2.64256972, -0.7126391 ],
 [-0.79486196,  0.68874785,  0.51821792, -0.79549511],
 [-1.22085697,  0.73036913,  0.48119354, -0.69736839]]
```

x_{i-1}

A
C
G
T

A C G T

x_i

x = GATC

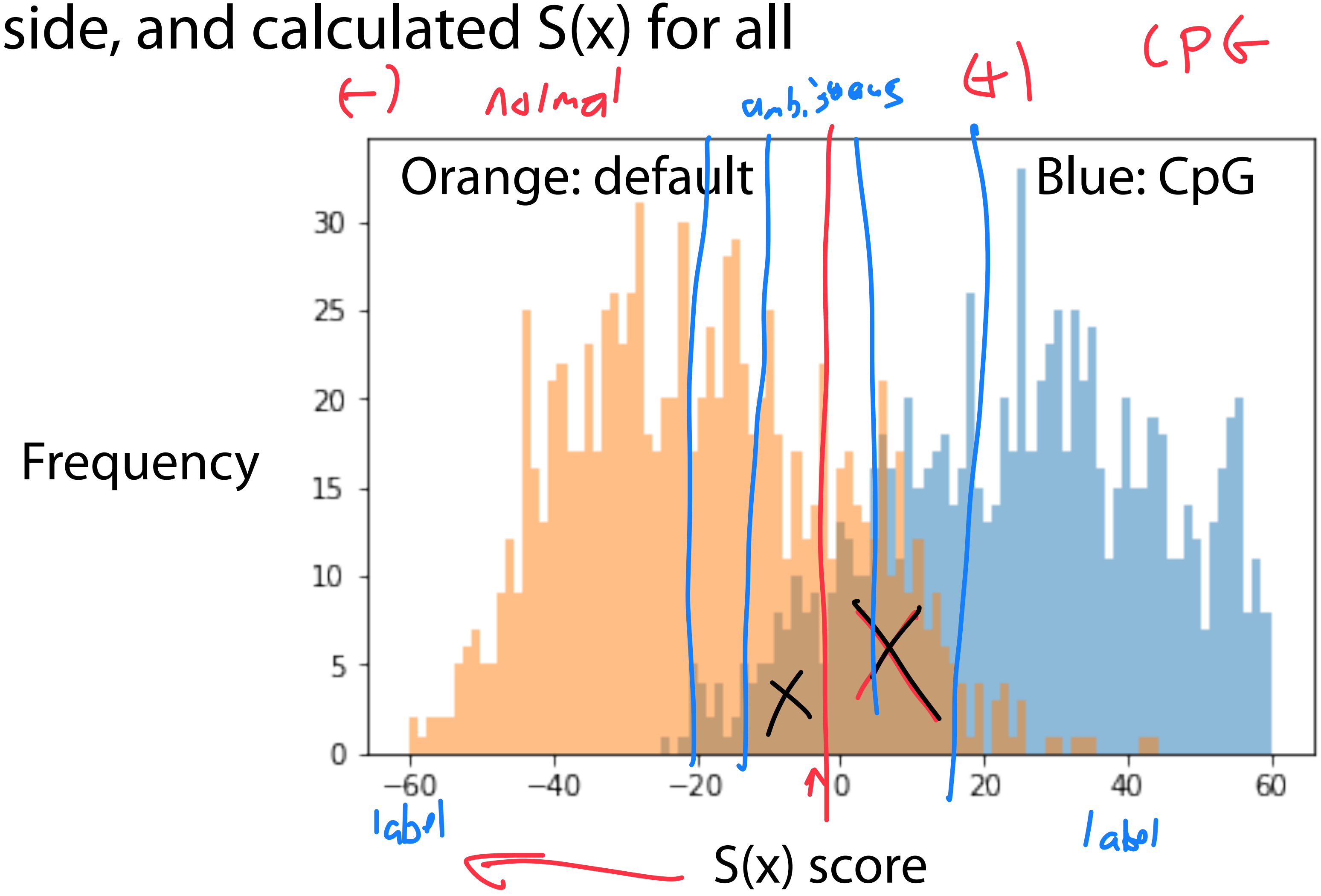
$$P(x) = P(x_4 \mid x_3)\,P(x_3 \mid x_2)\,P(x_2 \mid x_1)\,P(x_1)$$

$$P(x) = P(C \mid T)\,P(T \mid A)\,P(A \mid G)\,P(G) = 0.73036913 + -0.85527103 + -0.79486196 = -0.919763$$

Example by Ben Langmead

Markov Chain in Sequencing

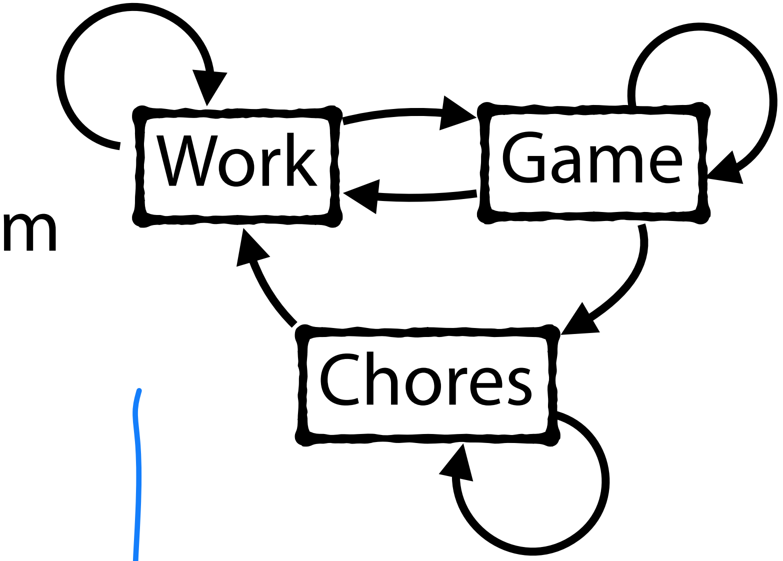
Drew 1,000 100-mers from inside CpG islands and another 1,000 from outside, and calculated S(x) for all



Markov Chain Matrix

If I'm working at time 0, what is probability that I'm working at time t?

Claim: $Pr(X_t = v \mid X_0 = u) = M^t[u, v]$

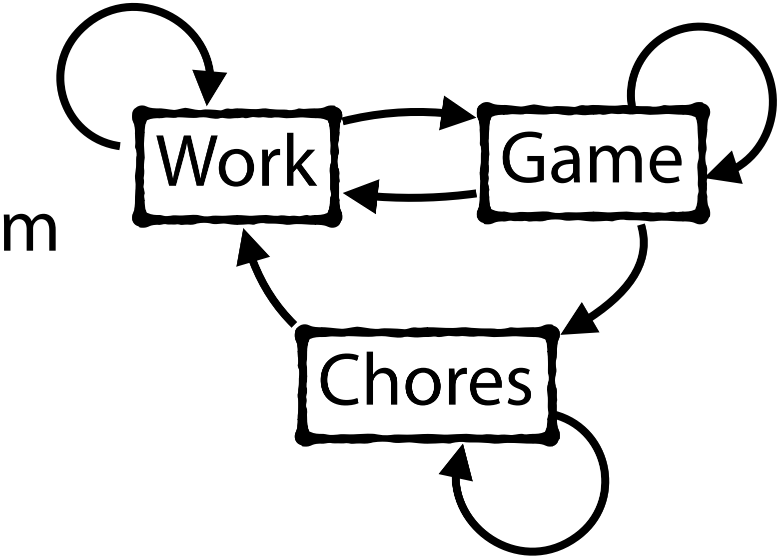


$$M = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & \cancel{0} & .5 \end{pmatrix}$$

Base Case:

T=1:

Markov Chain Matrix



If I'm working at time 0, what is probability that I'm working at time t?

Claim: $Pr(X_t = v \mid X_0 = u) = M^t[u,\ v]$

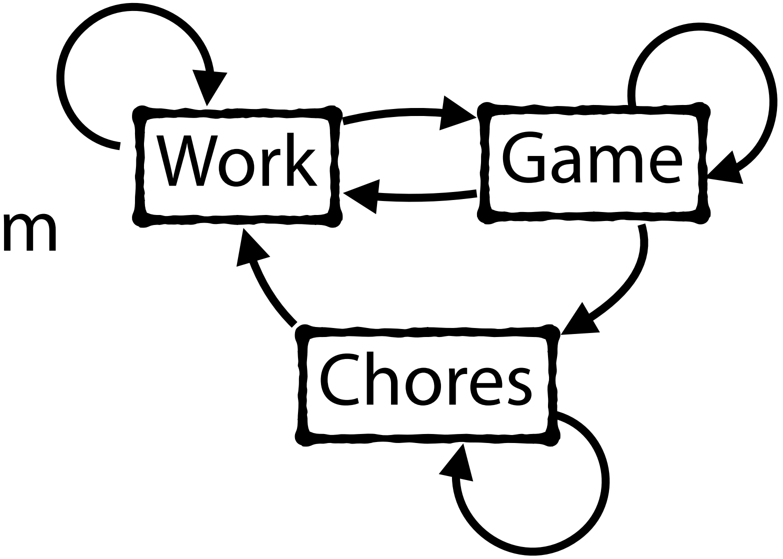
Base Case:

✓

T=1: $Pr(\underline{X_1 = W} \mid \underline{X_0 = W}) = M^1[W,\ W] = 0.4$

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

Markov Chain Matrix



If I'm working at time 0, what is probability that I'm working at time t?

Claim: $Pr(X_t = v \mid X_0 = u) = M^t[u,\ v]$

Base Case:

T=2: $Pr(X_2 = W \mid X_0 = W) =$ 0.16

0.4 0.4

$= Pr(X_2 = W \mid X_1 = W) * Pr(X_1 = W \mid X_0 = W)$

0.1

$+ Pr(X_2 = W \mid X_1 = G) * Pr(X_1 = G \mid X_0 = W)$ 0.6 — 0.06

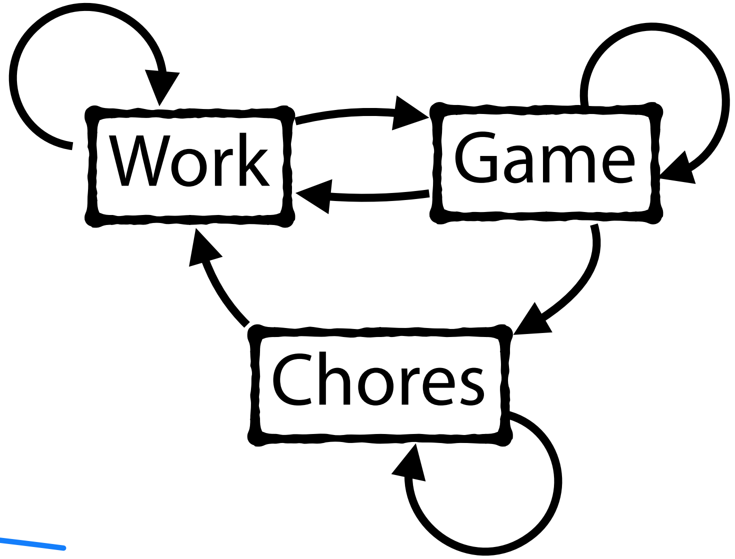
$+ Pr(X_2 = W \mid X_1 = C) * Pr(X_1 = C \mid X_0 = W)$

0.5 0

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

$$M^2 = \begin{pmatrix} .22 & .6 & .18 \\ .25 & .42 & .33 \\ .45 & 0.3 & .25 \end{pmatrix}$$

Markov Chain Matrix



Claim: $Pr(X_t = v \,|\, X_0 = u) = M^t[u,\ v]$

Induction:

Assume $Pr(X_{t-1} = v \,|\, X_0 = u) = M^{t-1}[u,\ v]$.

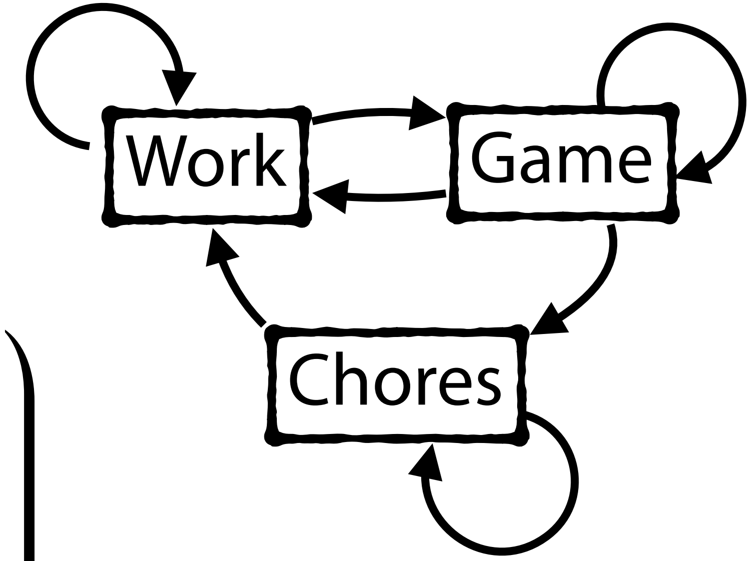
Show holds for $Pr(X_t = w \,|\, X_0 = u) = M^t[u,\ w]$

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

By Markov Assumption — trivial!

The same logic (and math) for finding T=2 applies here

Markov Chain Matrix



What happens as $t \rightarrow \infty$?

$$M = \begin{pmatrix} .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{pmatrix} \qquad M^3 = \begin{pmatrix} .238 & .492 & .270 \\ .307 & .402 & .291 \\ .335 & .450 & .215 \end{pmatrix}$$

$$M^{10} = \begin{pmatrix} .2940 & .4413 & .2648 \\ .2942 & .4411 & .2648 \\ .2942 & .4413 & .2648 \end{pmatrix}$$

$$M^{60} = \begin{pmatrix} .2941 & .4412 & .2647 \\ .2941 & .4412 & .2647 \\ .2941 & .4412 & .2647 \end{pmatrix}$$

Markov Chain Stationary Distribution

A probability vector π is called a **stationary distribution** for a Markov Chain if it satisfies the stationary equation: $\pi = \pi M$

$$M = \begin{array}{c} \\ \left(\begin{array}{ccc} W & G & C \\ .4 & .6 & 0 \\ .1 & .6 & .3 \\ .5 & 0 & .5 \end{array}\right) \end{array}$$

$$\pi[W] = .4\pi[W] + .1\pi[G] + .5\pi[C]$$

$$\pi[S] = .6\pi[W] + .6\pi[G] + 0\pi[C]$$

$$\pi[E] = 0\pi[W] + .3\pi[G] + .5\pi[C]$$

$$W + G + C = 1$$

Markov Chain Stationary Distribution

A probability vector π is called a **stationary distribution** for a Markov Chain if it satisfies the stationary equation: $\pi = \pi M$

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

$$\pi[W] = .4\pi[W] + .1\pi[G] + .5\pi[C]$$

$$\pi[S] = .6\pi[W] + .6\pi[G] + 0\pi[C]$$

$$\pi[E] = 0\pi[W] + .3\pi[G] + .5\pi[C]$$

$$0.6W = 0.1G + 0.5C$$

$$0.4G = 0.6W$$

$$0.5C = 0.3G$$

$$W + G + C = 1$$

$$W = 10/34 = 0.291$$

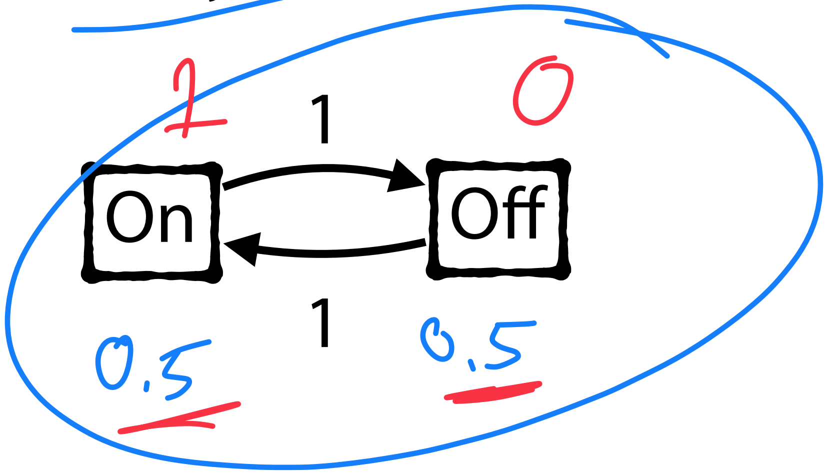
$$G = 15/34 = 0.441$$

$$C = 9/34 = 0.265$$

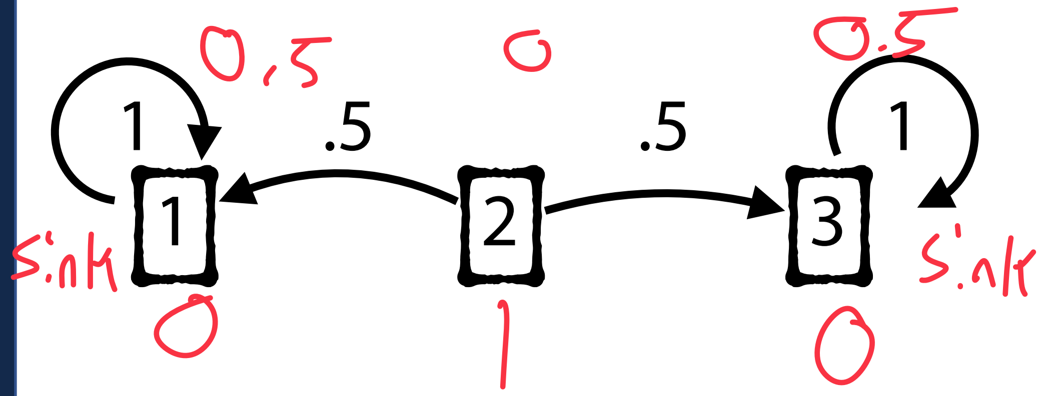
$$M^{60} = \begin{pmatrix} .2941 & .4412 & .2647 \\ .2941 & .4412 & .2647 \\ .2941 & .4412 & .2647 \end{pmatrix}$$

Markov Chain Stationary Distribution

Stationary distributions can be calculated using the system of equation (and that all probabilities sum to 1). **But not every Markov Chain has a steady state (and some have infinitely many)!**



One steady state only if (0.5, 0.5)



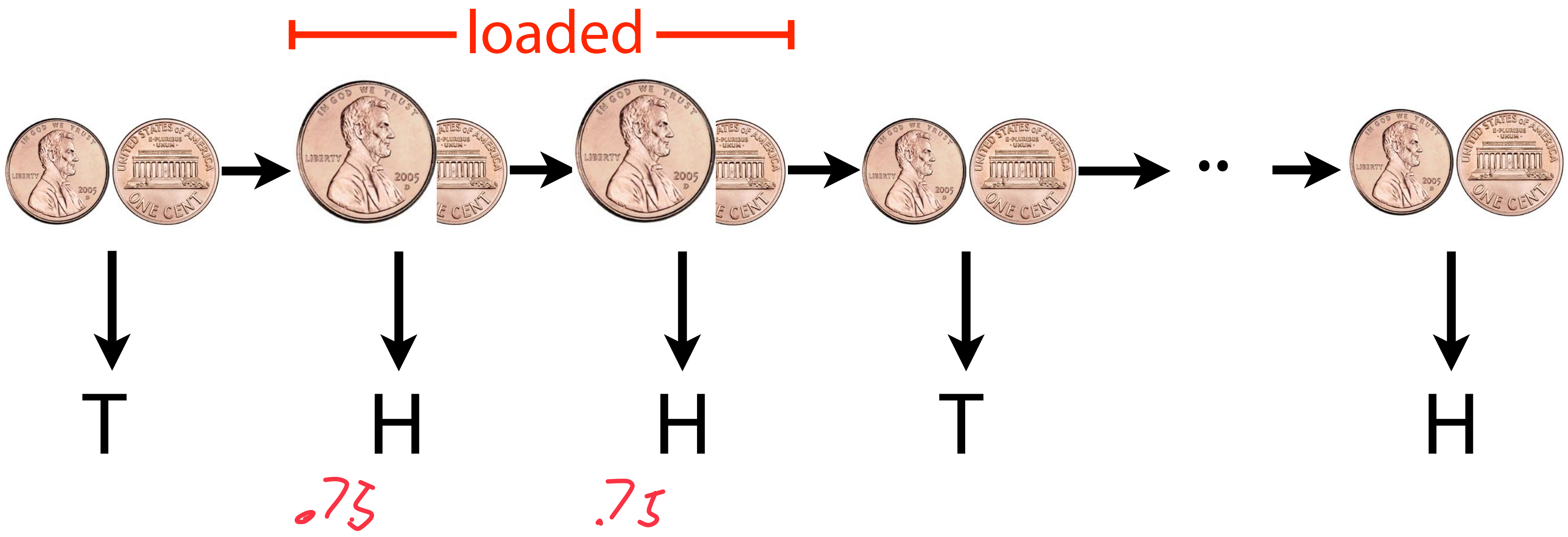
∞ steady states!

Hidden Markov Models

In the real world, we often don't know the underlying markov chain!

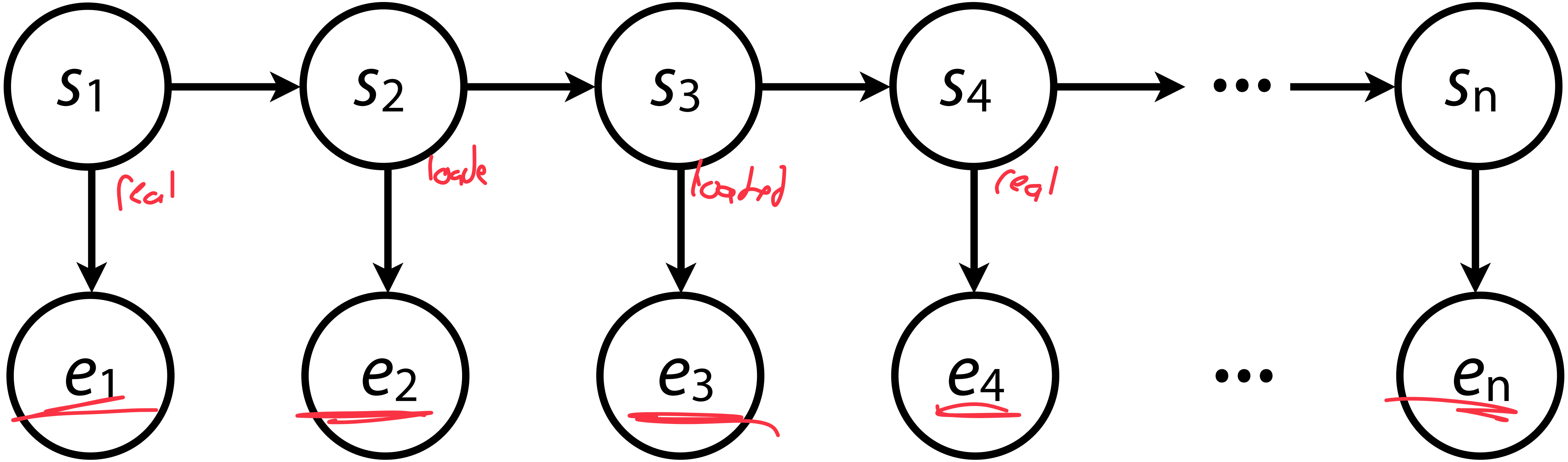
Instead, we have observations that can be used to predict our current state.

Ex: Repeated coin flips but *sometimes* I cheat and use a fixed coin.



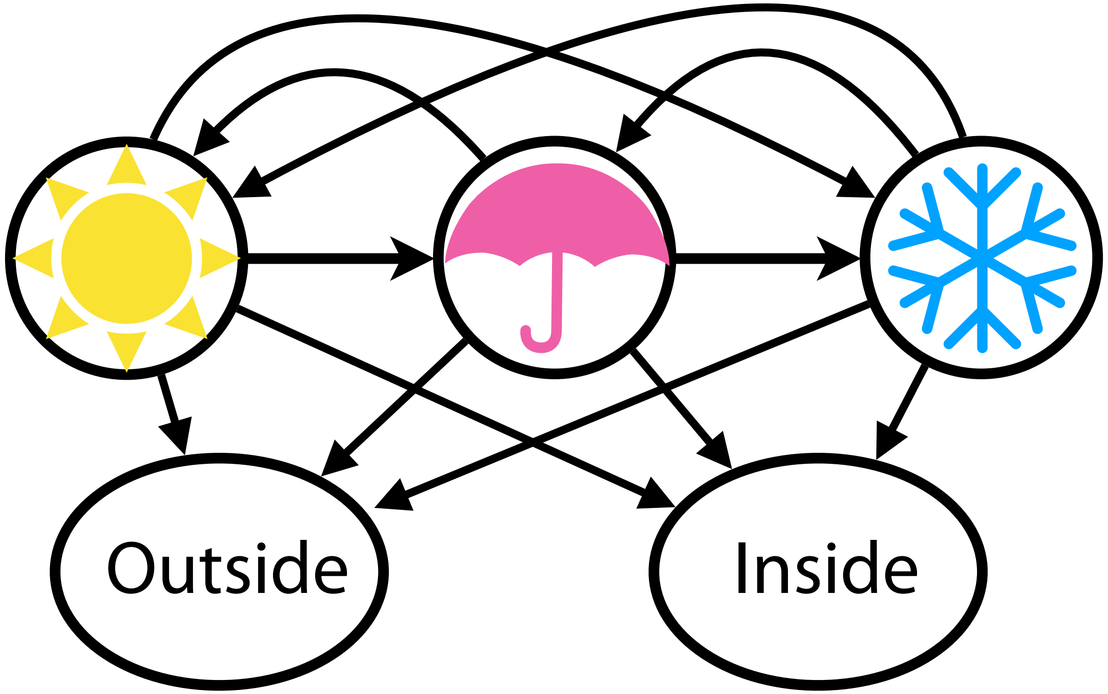
Hidden Markov Models

Unobserved States



Observed Emissions

Hidden Markov Models



$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix} \qquad E = \begin{pmatrix} \underline{.8} & .2 \\ .3 & \underline{.7} \\ \underline{.5} & \underline{.5} \end{pmatrix}$$

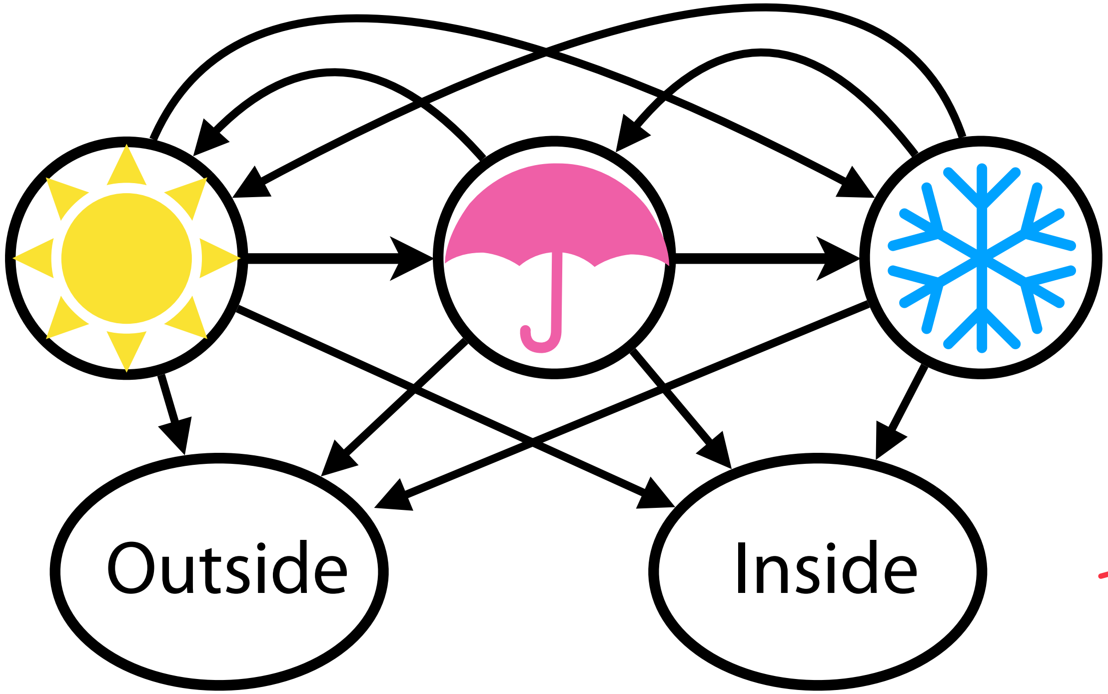
Pr(e = {O, I, O} | s = {C, R, S})?

observed emissions given underlying states

starting state

Pr(e={O, I, O}, s={C, R, S} | π_C =0.4)?

Hidden Markov Models



$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix} \quad E = \begin{pmatrix} .8 & .2 \\ .3 & .7 \\ .5 & .5 \end{pmatrix}$$

Pr(e = {O, I, O} | s = {C, R, S})?

Pr(O|C)*P(I|R)*P(O|S)

Pr(e={O, I, O}, s={C, R, S} | π_C =0.4)?

specific

Value is probability

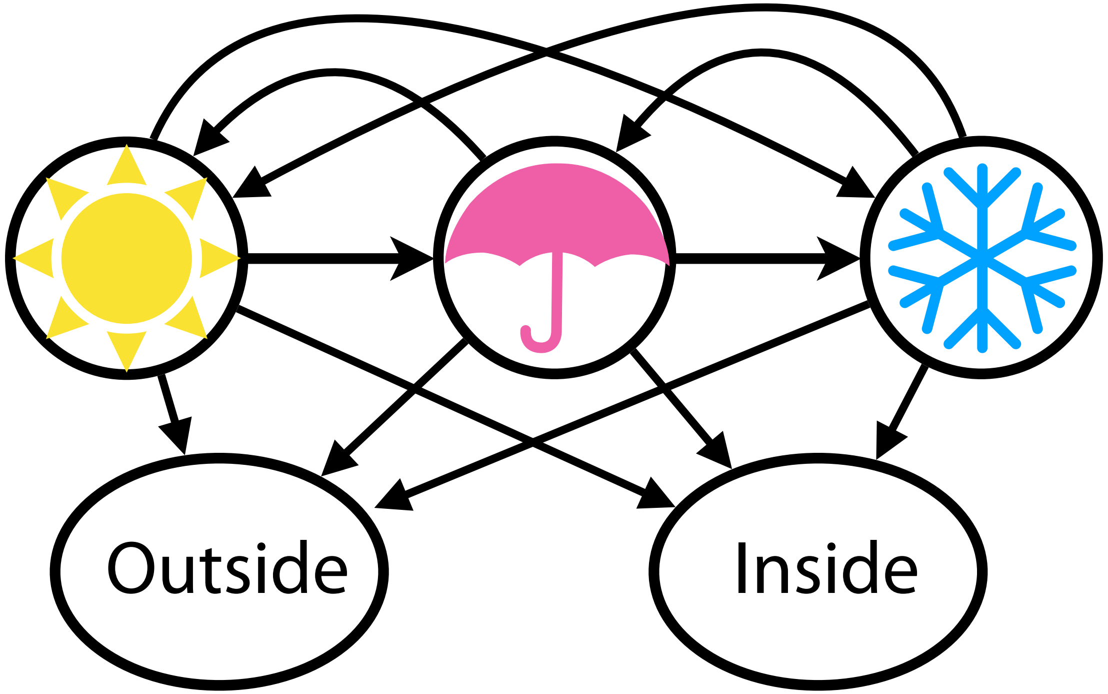
P(T_0=C)*P(T_1=R | T_0 = C)*P(T_2=S| T_1=R)] * [P(O|C)*P(I|R)*P(O|S)]

transitions in unknown states

emission probabilities

.4*.3*.1*.8*.7*.5 = 0.00336

Hidden Markov Models



$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix} \quad E = \begin{pmatrix} .8 & .2 \\ .3 & .7 \\ .5 & .5 \end{pmatrix}$$

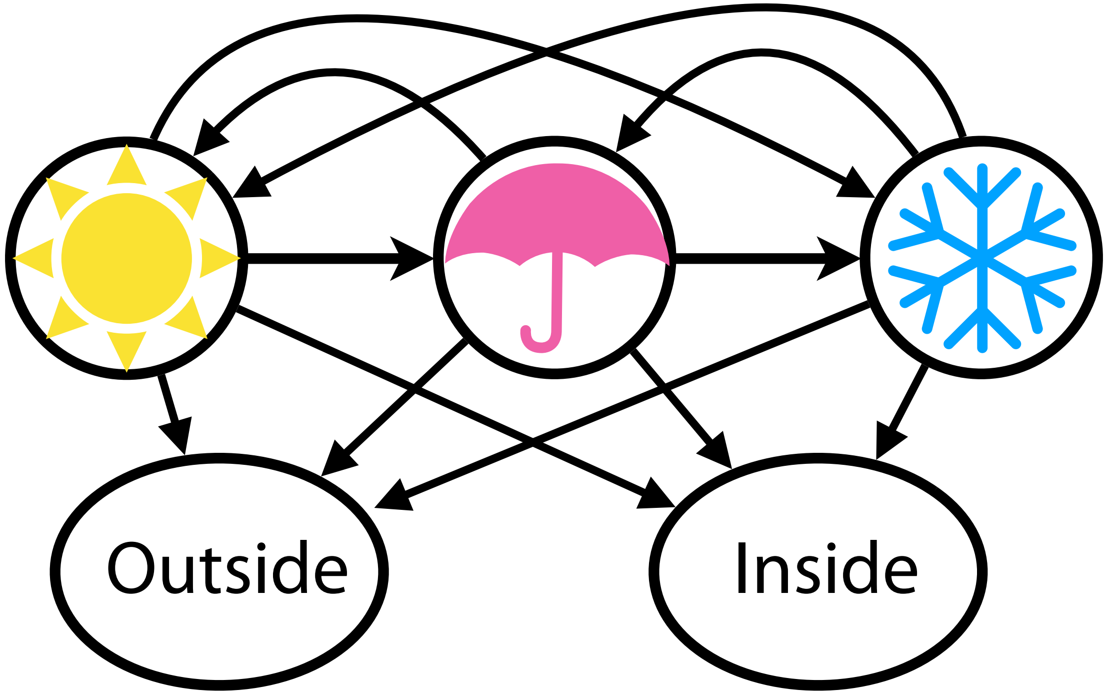
Pr(e={O, I, O})?

Prob of states

All starting states

• Prob Emission

Hidden Markov Models



$$M = \begin{pmatrix} .5 & .3 & .2 \\ .5 & .4 & .1 \\ .2 & .1 & .7 \end{pmatrix} \quad E = \begin{pmatrix} .8 & .2 \\ .3 & .7 \\ .5 & .5 \end{pmatrix}$$

If I go outside for three days, what was the most likely weather?

Viterbi Algorithm

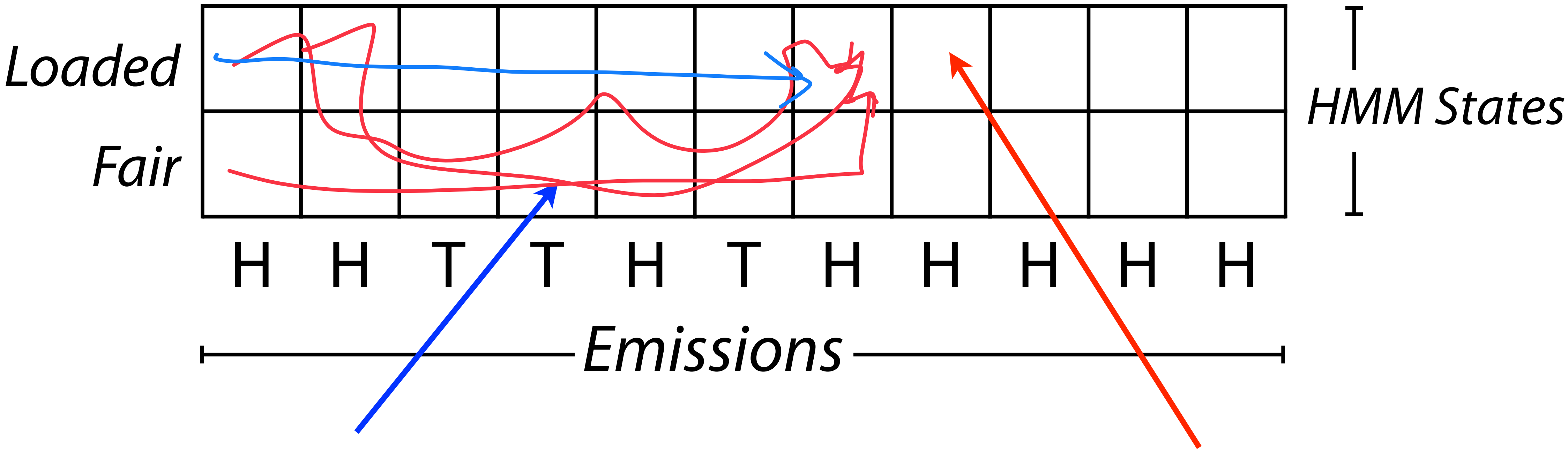
We can brute force all possible combinations…

… or we can use the Markov Assumption with Dynamic Programming

$$M = \begin{pmatrix} .6 & .4 \\ .4 & .6 \end{pmatrix} \qquad E = \begin{pmatrix} .8 & .2 \\ .5 & .5 \end{pmatrix}$$

Example by Ben Langmead

Viterbi Algorithm

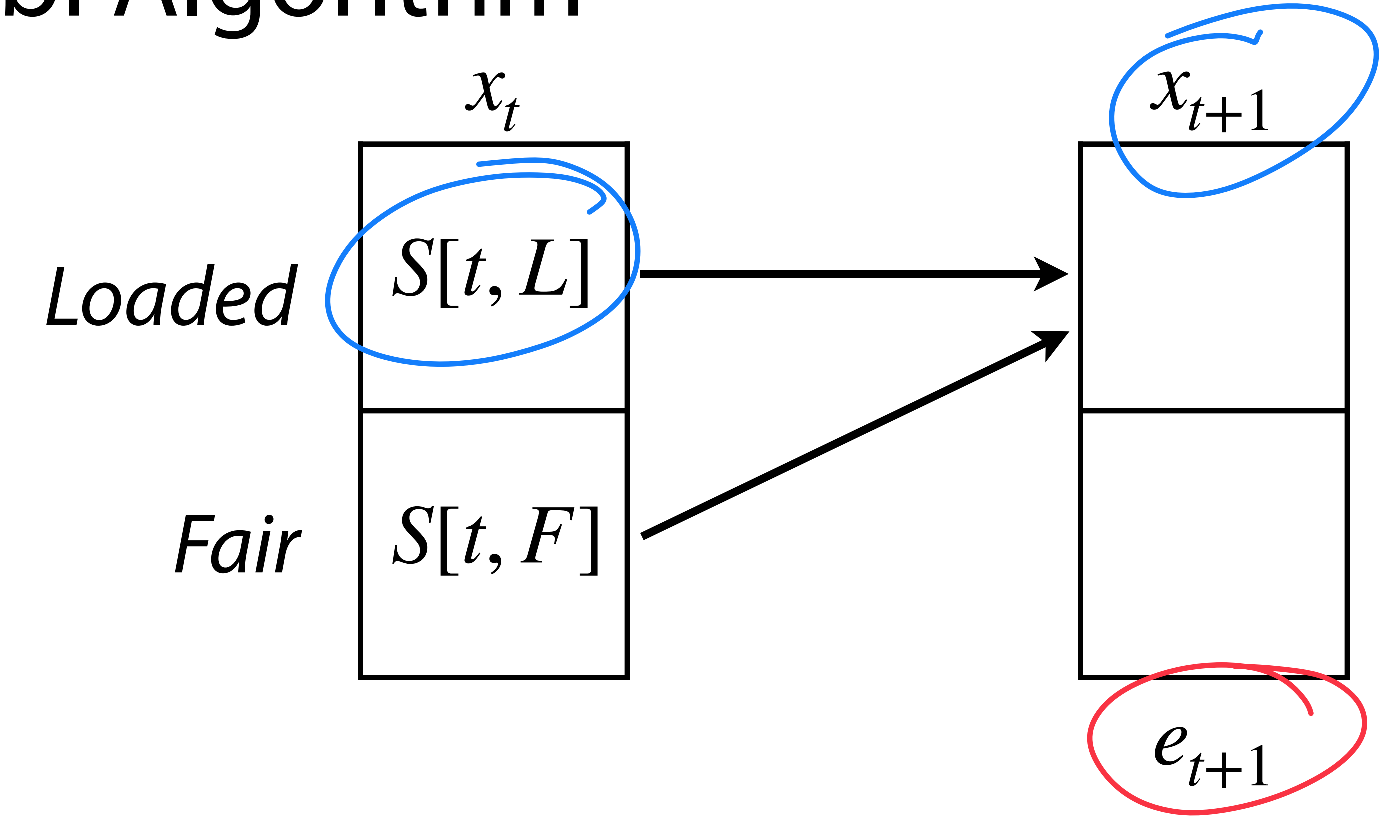


$\max_{x} P(xF, HHTT)$

$\max_{x} P(xL, HHTTHTHH)$

S[i, k] = greatest joint probability of observing the length-*i* prefix of *e* and any sequence of states ending in state *k*

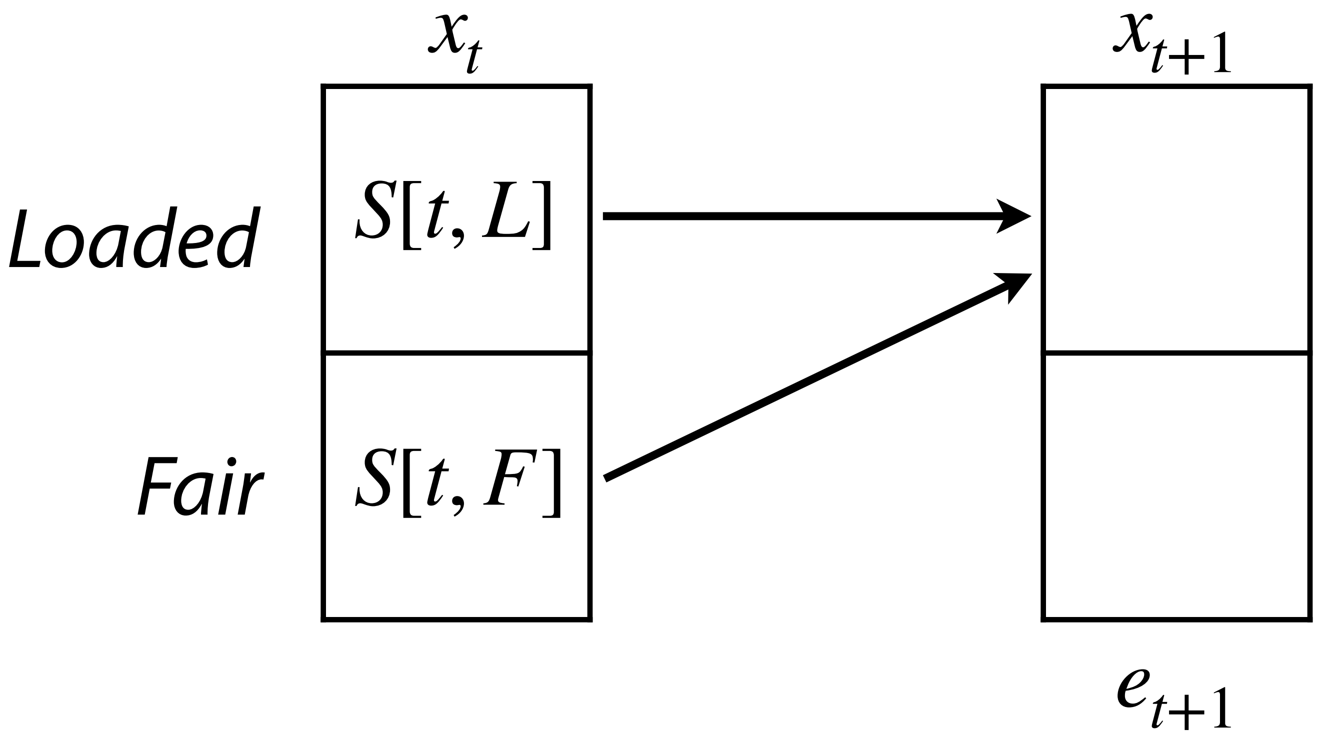
Viterbi Algorithm



$$S[t+1, L] = S[t,L] * M[L,L] \cdot Pr(e_{t+1} | L)$$

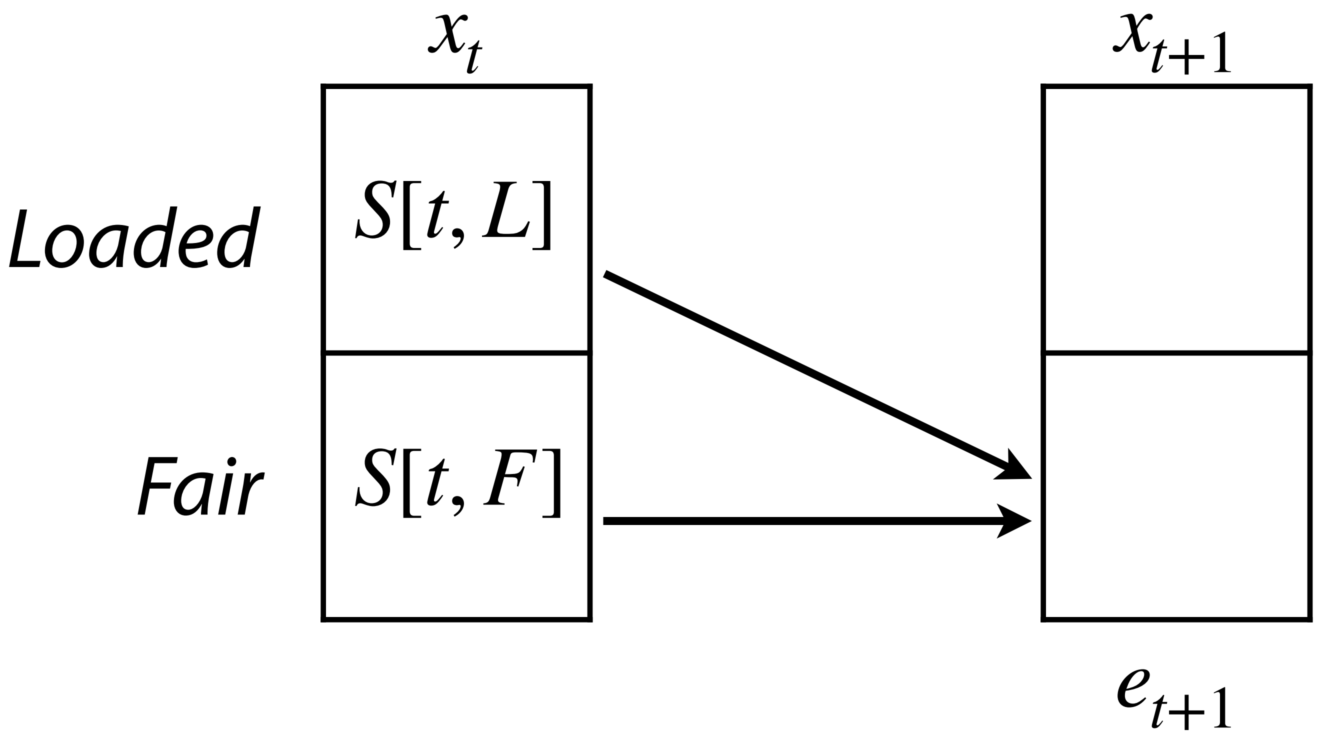
$$S[t,F] * M[L,F] \cdot Pr(e_{t+1} | F)$$

Viterbi Algorithm



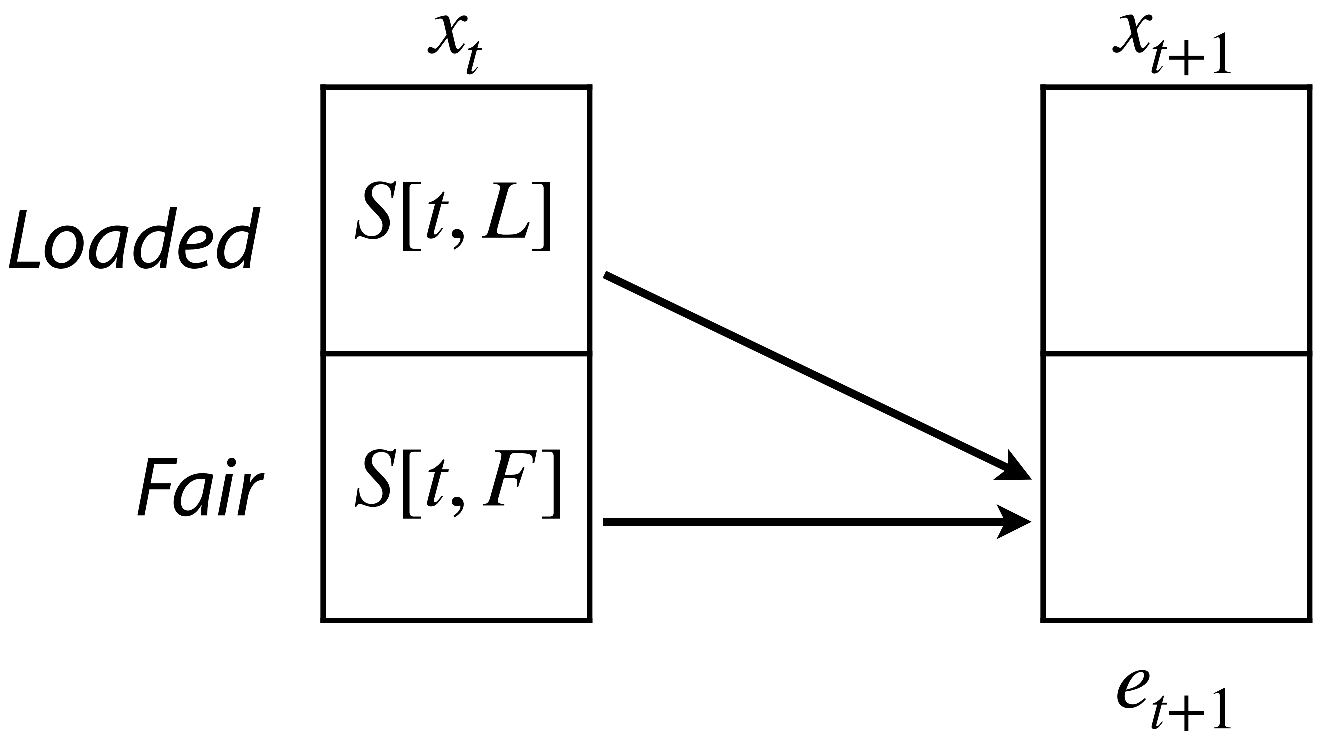
$$S[t+1,\ L] = \ max \begin{cases} S[t,\ L] * M[L\,|\,L] * E[e_{t+1}\,|\,L] \\ S[t,\ F] * M[L\,|\,F] * E[e_{t+1}\,|\,L] \end{cases}$$

Viterbi Algorithm



$$S[t+1,\ F] =$$

Viterbi Algorithm



x_t

x_{t+1}

Loaded

$S[t, L]$

Fair

$S[t, F]$

e_{t+1}

$$S[t+1,\ F] = \ max \begin{cases} S[t,\ L] * M[F \mid L] * E[e_{t+1} \mid F] \\ S[t,\ F] * M[F \mid F] * E[e_{t+1} \mid F] \end{cases}$$

Viterbi Algorithm

Loaded											
Fair											
Observations	H	H	T	T	H	T	H	H	H	H	H

$$M = \begin{pmatrix} .6 & .4 \\ .4 & .6 \end{pmatrix}$$

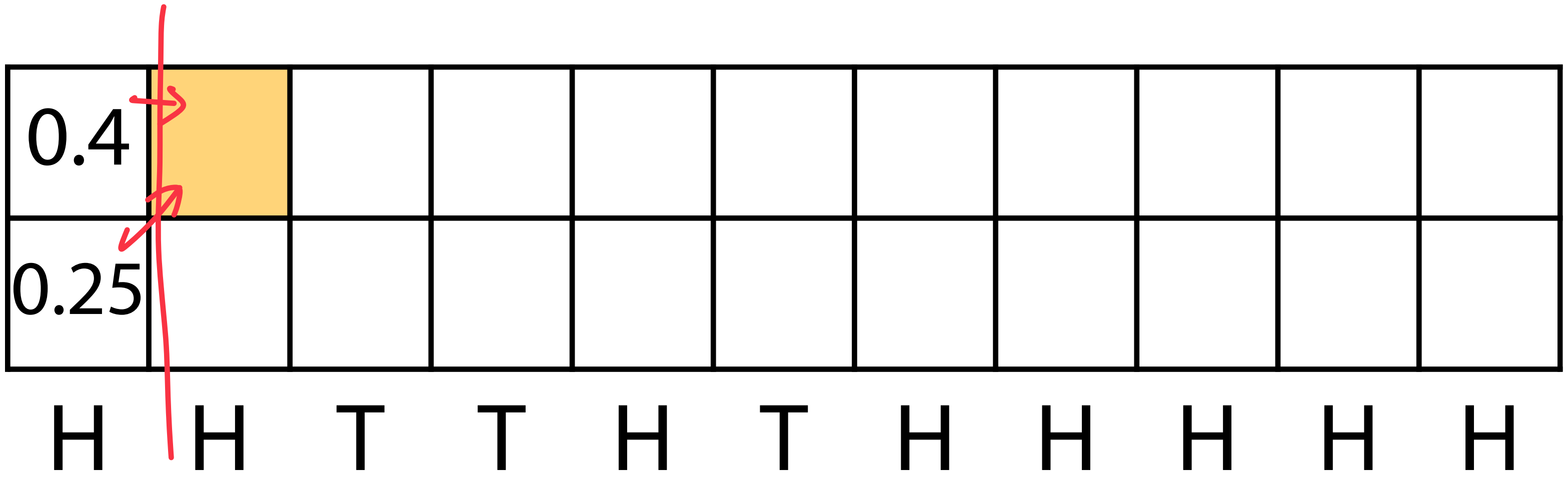
$$E = \begin{pmatrix} .8 & .2 \\ .5 & .5 \end{pmatrix}$$

Assume we start with Fair/Loaded with equal probability

$$S[0, L] = \underline{0.5} \cdot \boldsymbol{E(H \mid L)} \qquad S[0, F] = \underline{0.5} \cdot \boldsymbol{E(H \mid F)}$$

$$= 0.5 \cdot \boldsymbol{0.8} \qquad = 0.5 \cdot \boldsymbol{0.5}$$

Viterbi Algorithm



	H	H	T	T	H	T	H	H	H	H	H
Loaded	0.4										
Fair	0.25										

$$M = \begin{pmatrix} \underline{.6} & .4 \\ .4 & \underline{.6} \end{pmatrix}$$

$$E = \begin{pmatrix} \boxed{.8} & .2 \\ .5 & .5 \end{pmatrix}$$

$$S[1, L] = \overset{s[0,L)}{0.4} \cdot \overset{M[L,L)}{.6} \cdot \overset{e(H|L)}{.8}$$

Viterbi Algorithm

Loaded	0.4	0.19									
Fair	0.25										
	H	H	T	T	H	T	H	H	H	H	H

$$M = \begin{pmatrix} .6 & .4 \\ .4 & .6 \end{pmatrix}$$

$$E = \begin{pmatrix} .8 & .2 \\ .5 & .5 \end{pmatrix}$$

$S[1,\, F] =$

Viterbi Algorithm

	H	H	T	T	H	T	H	H	H	H	H
Loaded	0.4	0.19									
Fair	0.25	0.08									

$$M = \begin{pmatrix} .6 & .4 \\ .4 & .6 \end{pmatrix}$$

$$E = \begin{pmatrix} .8 & .2 \\ .5 & .5 \end{pmatrix}$$

Viterbi Algorithm

These get small very fast— use log_2 scaling

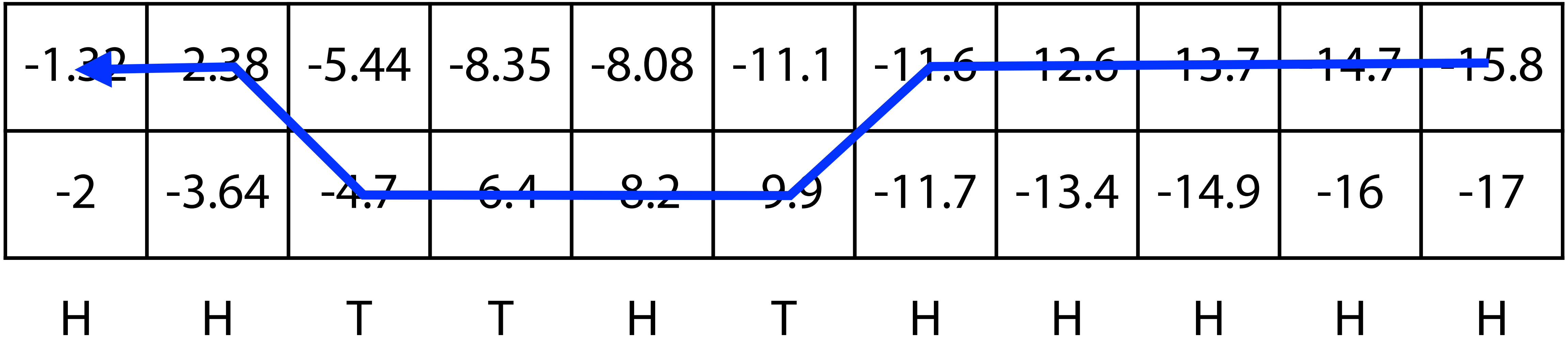
-1.32	-2.38	-5.44	-8.35	-8.08	-11.1	-11.6	-12.6	-13.7	-14.7	-15.8
-2	-3.64	-4.7	-6.4	-8.2	-9.9	-11.7	-13.4	-14.9	-16	-17
H	H	T	T	H	T	H	H	H	H	H

Traceback: Same as edit distance!

Start from largest value and remember ‘where I came from’

Viterbi Algorithm

These get small — now log_2 scaled

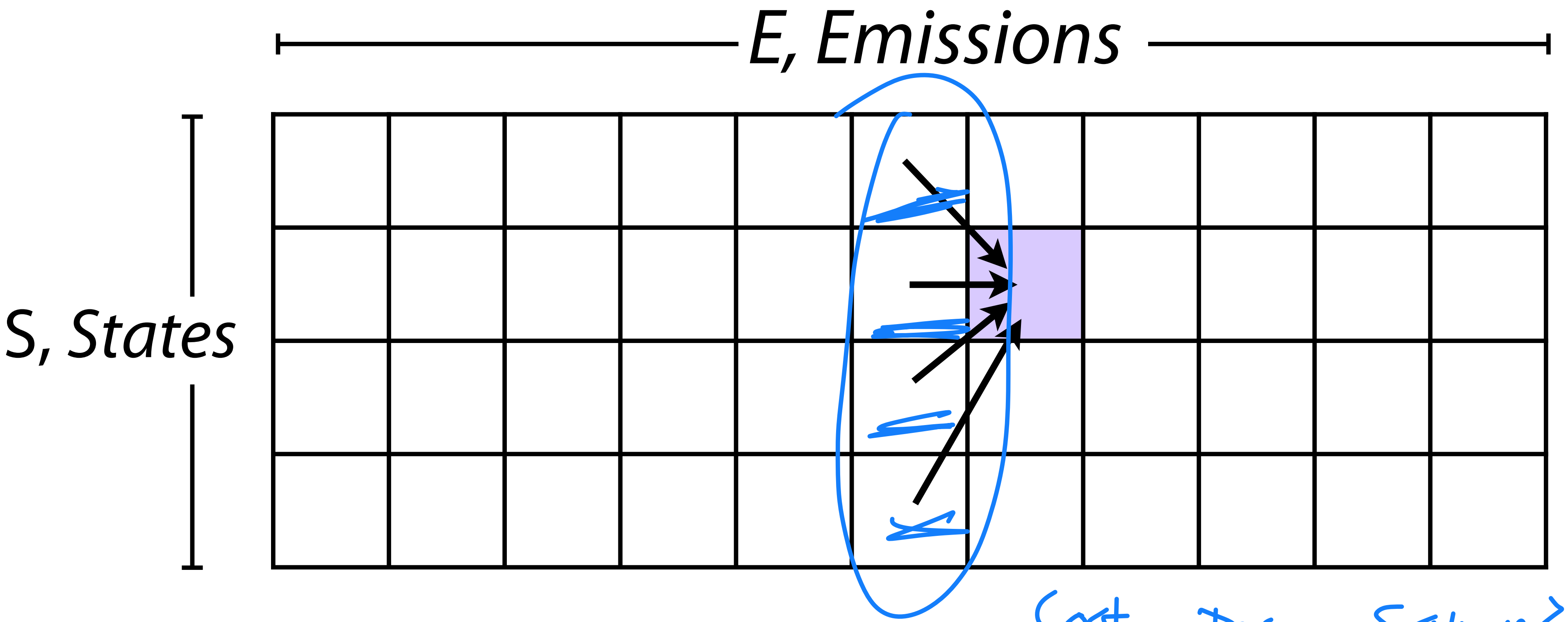


-1.32	-2.38	-5.44	-8.35	-8.08	-11.1	-11.6	-12.6	-13.7	-14.7	-15.8
-2	-3.64	-4.7	-6.4	-8.2	-9.9	-11.7	-13.4	-14.9	-16	-17
H	H	T	T	H	T	H	H	H	H	H

Traceback: Same as edit distance!

Start from largest value and remember 'where I came from'

Viterbi Algorithm



$O(E \cdot S^2)$

Cost per square?

$O(S)$

What is running time?

$S \cdot E$ is # of squares

*