Markov Models

Introduction to
Artificial Intelligence

COS302

Michael L. Littman

Fall 2001

Administration

Hope your midterm week is going well.

Breaking the Bank

Assistant professor: $20k/year.

How much total cash?

20+20+20+20+... = infinity!

Nice, eh?

Discounted Rewards

Idea: The promise of future payment is not worth *quite* as much as payment now.

- Inflation / investing
- Chance of game “ending”

Ex. $10k next year might be worth $10k x 0.9 today.

Infinite Sum

- Assuming a discount rate of 0.9, how much does the assistant professor get in total?

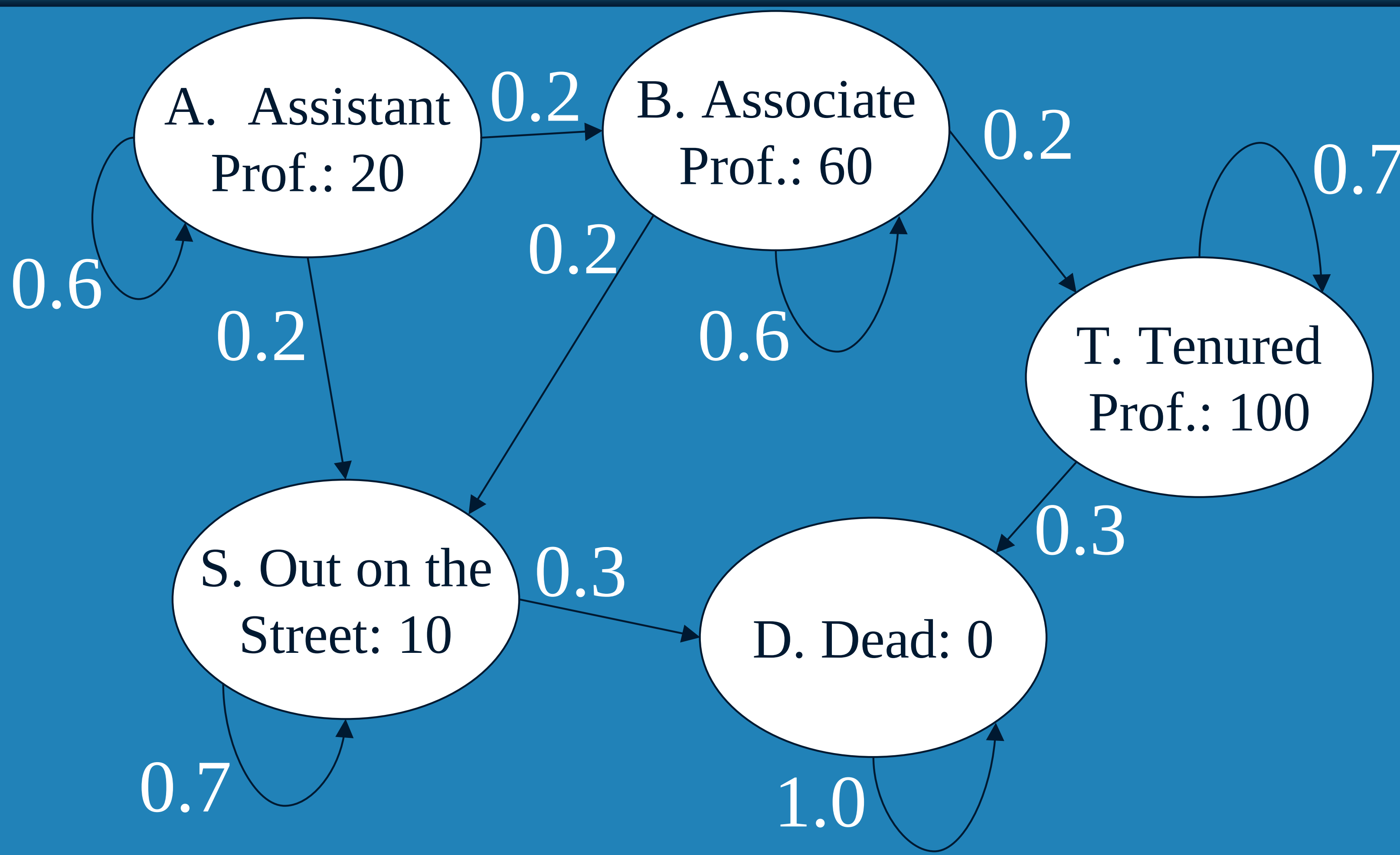
$20 + .9\ 20 + .9^2\ 20 + .9^3\ 20 + \ldots$

$= 20 + .9\ (20 + .9\ 20 + .9^2\ 20 + \ldots)$

$x = 20 + .9\ x$

$x = 20/(.1) = 200$

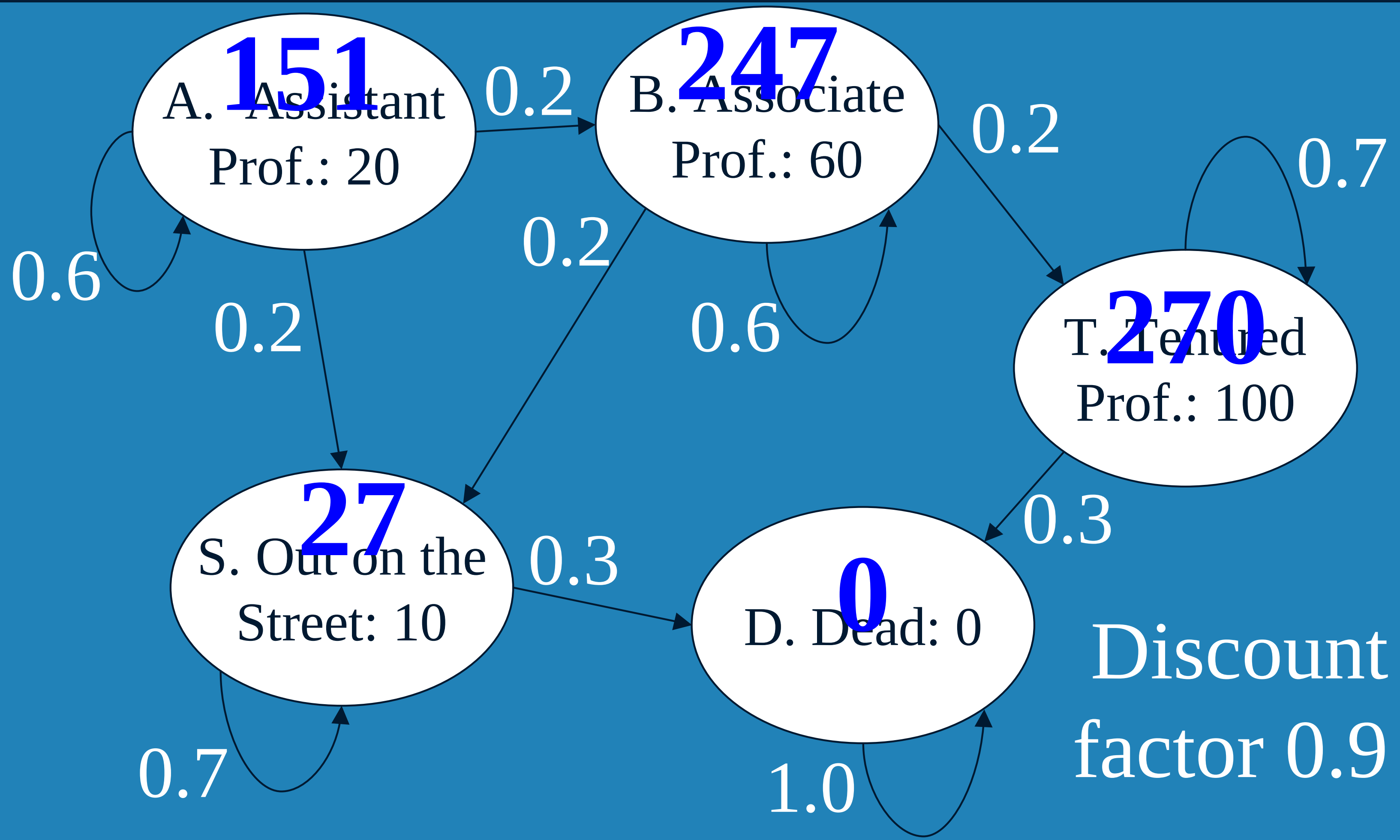
Academic Life



Solving for Total Reward

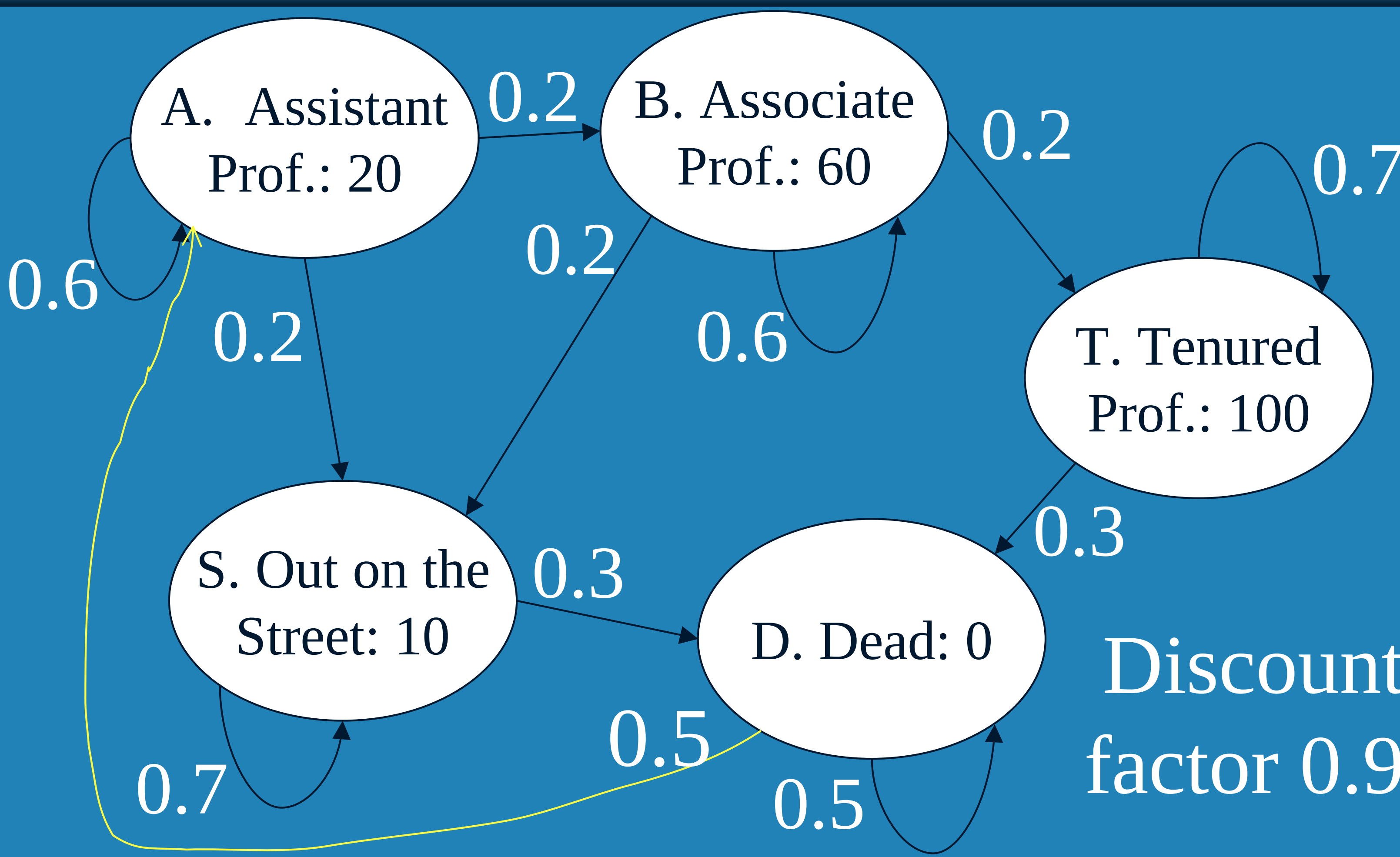
- L(i) is expected total reward received starting in state i.
- How could we compute L(A)?
- Would it help to compute L(B), L(T), L(S), and L(D) also?

Working Backwards



Discount factor 0.9

Reincarnation?



Discount factor 0.9

System of Equations

L(A) = 20 + .9(.6 L(A) + .2 L(B) + .2 L(S))

L(B) = 60 + .9(.6 L(B) + .2 L(S) + .2 L(T))

L(S) = 10 + .9(.7 L(S) + .3 L(D))

L(T) = 100 + .9(.7 L(T) + .3 L(D))

L(D) = 0 + .9 (.5 L(D) + .5 L(A))

Transition Matrix

Let P be the matrix of probs: Pij = Pr(next = j | current = i) To

From

	A	B	S	T	D
A	0.6	0.2	0.2		
B		0.6	0.2	0.2	
S			0.7		0.3
T				0.7	0.3
D	0.5				0.5

Matrix Equation

L(A)=	20+.9	.6	.2	.2			L(A)
L(B)=	60+.9		.6	.2	.2		L(B)
L(S)=	10+.9			.7		.3	L(S)
L(T)=	100+.9				.7	.3	L(T)
L(D)=	0+.9	.5				.5	L(D)

$L = R + \gamma \, P \, L$

Solving the Equation

$L = R + \gamma P L$

$L - \gamma P L = R$

$I L - \gamma P L = R$ (introduce identity)

$(I - \gamma P) L = R$

$(I - \gamma P)^{-1} (I - \gamma P) L = (I - \gamma P)^{-1} R$

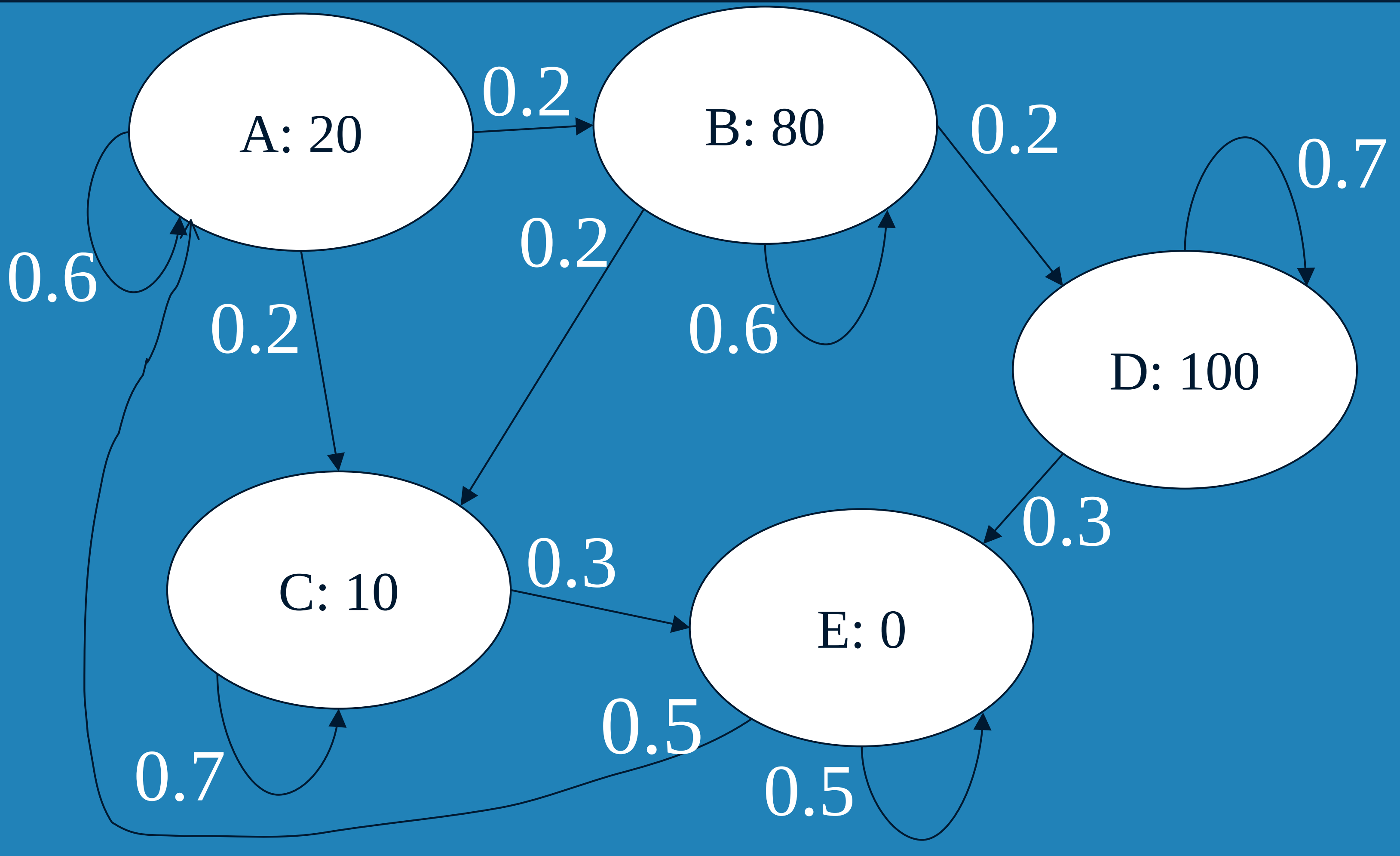
$L = (I - \gamma P)^{-1} R$

Matrix inversion, matrix-vec mult.

Markov Chain

- Set of states, transitions from state to state.
- Transitions only depend on current state, not the history: *Markov property*.

What Does a MC Do?



MC Problems

- Probability of going from s to s' in t steps.
- Probability of going from s to s' in t or fewer steps.
- Averaged over t steps (in the limit), how often in state s' starting at s?
- How many steps from s to s', on average?
- Given reward values, expected discounted reward starting at s.

Examples

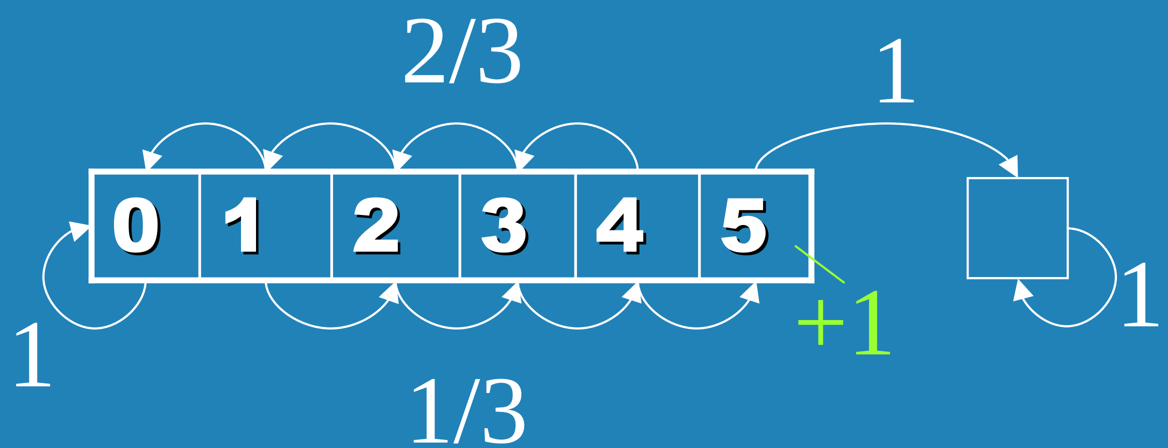
- Queuing system: Expected queue length, time until queue fills up.
- Chutes & Ladders: Avg game time.
- Genetic Algs: Time to find opt.
- Statistics: Time to “mix”.
- Blackjack: Expected winnings.
- Robot: Prob. of crash given controller.
- Gamblers ruin: Time until money gone.

Gambler's Ruin

Gambler has 3 dollars.
Win a dollar with prob. 1/3.
Lose a dollar with prob. 2/3.
Fail: no dollars.
Succeed: Have 5 dollars.

Probability of success?
Average time until success or failure?
Set up Markov chains.

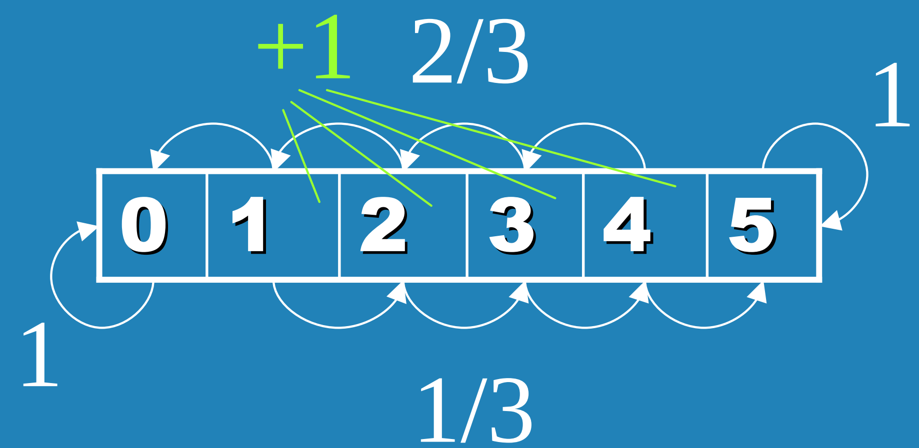
Ruin Chain



Solve for L(3) using γ=1.

Gives probability of success.

Gambling Time Chain

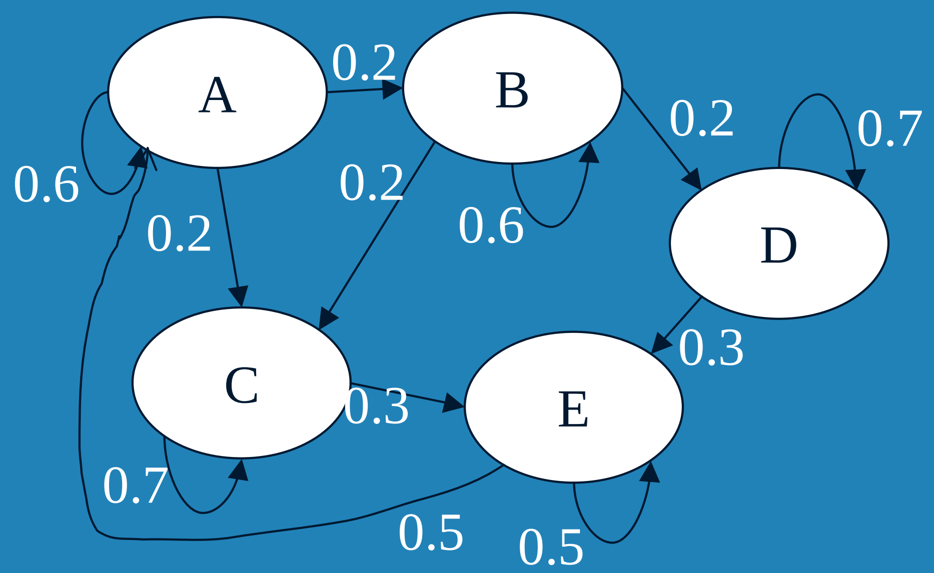


Solve for L(3) using γ =1.

L(3) will be the number of steps in 1, 2, 3, 4 states before finishing.

Stationary Distribution

What fraction of the time is spent in D?



L(D) =

0.7 L(D) + 0.2 L(B)

Also: L(A)+L(B)+L(C)+L(D)+L(E) = 1

Other Uses: Uncertainty

- Can add a notion of “actions” to create a *Markov decision process*. Useful for AI planning.
- Can add a notion of “observation” to create *hidden Markov model* (we’ll see these later).
- Add both to get *partially observable Markov decision process* (POMDP).

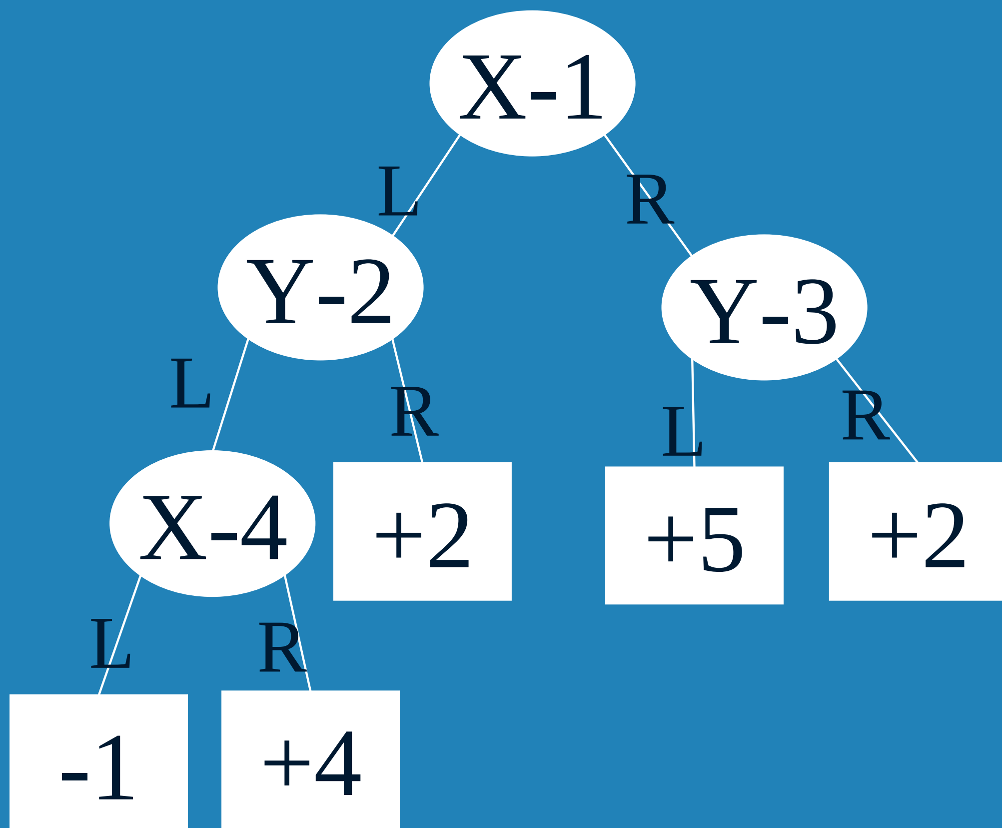
What to Learn

- Solving for the value of Markov processes using matrix equations.
- Setting up problems as Markov chains.

Homework 5 (due 11/7)

1. The value iteration algorithm from the *Games of Chance* lecture can be applied to deterministic games with loops. Argue that it produces the same answer as the “Loopy” algorithm from the *Game Tree* lecture.
2. Write the matrix form of the game tree below.

Game Tree



Continued

3. How many times (on average) do you need to flip a coin before you flip 3 heads in a row? (a) Set this up as a Markov chain, and (b) solve it.
4. More soon...