CS 295A/395D: Artificial Intelligence

Probabilistic state over time: Markov Chains

Prof. Emma Tosch

20 April 2022

The University of Vermont

Logistics

- Reminder: worksheet due in class on Friday
- Post questions to Teams
- Be sure to get started on your programming assignment
  - Reminder: you may work in pairs!

Today: Markov Chains
(i.e., probabilistic state transitions)

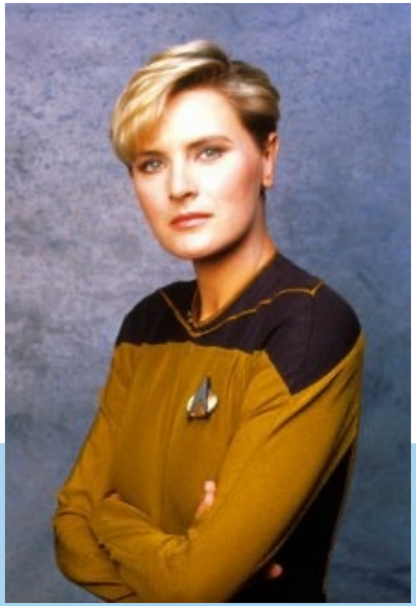
Recall: LTL Example

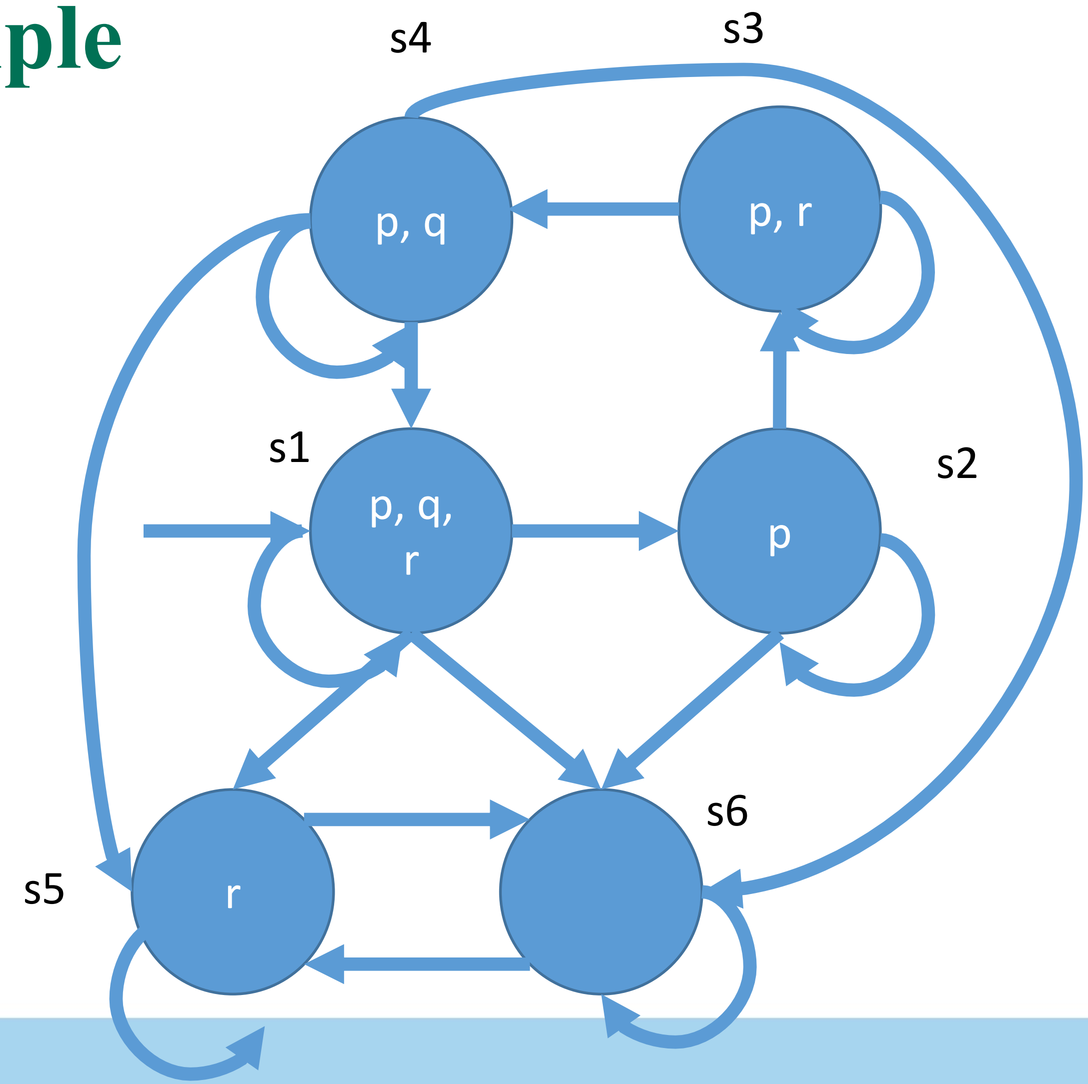
Example:

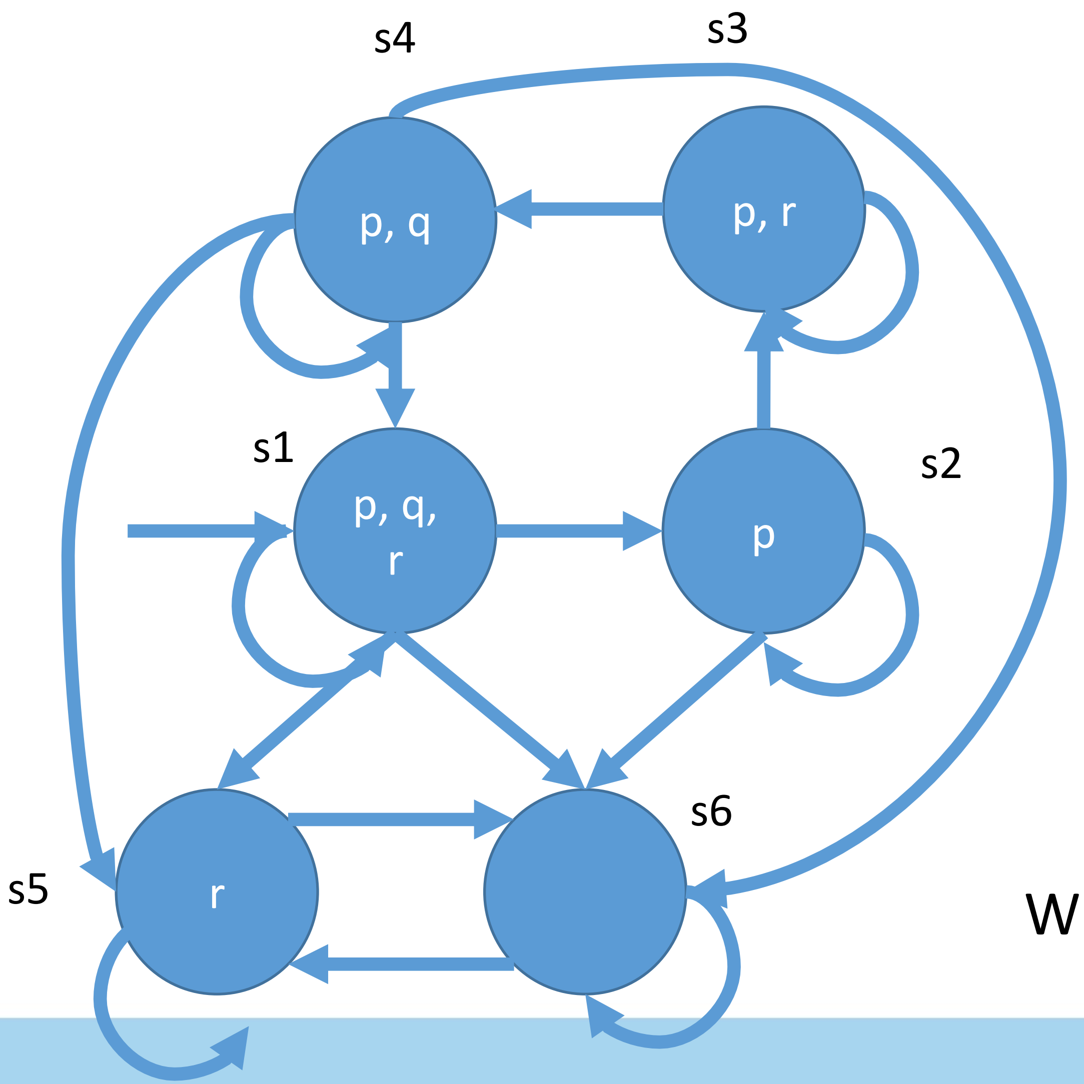
p = Yar alive

q = Yar on our ship

r = transporter ready







Previously:

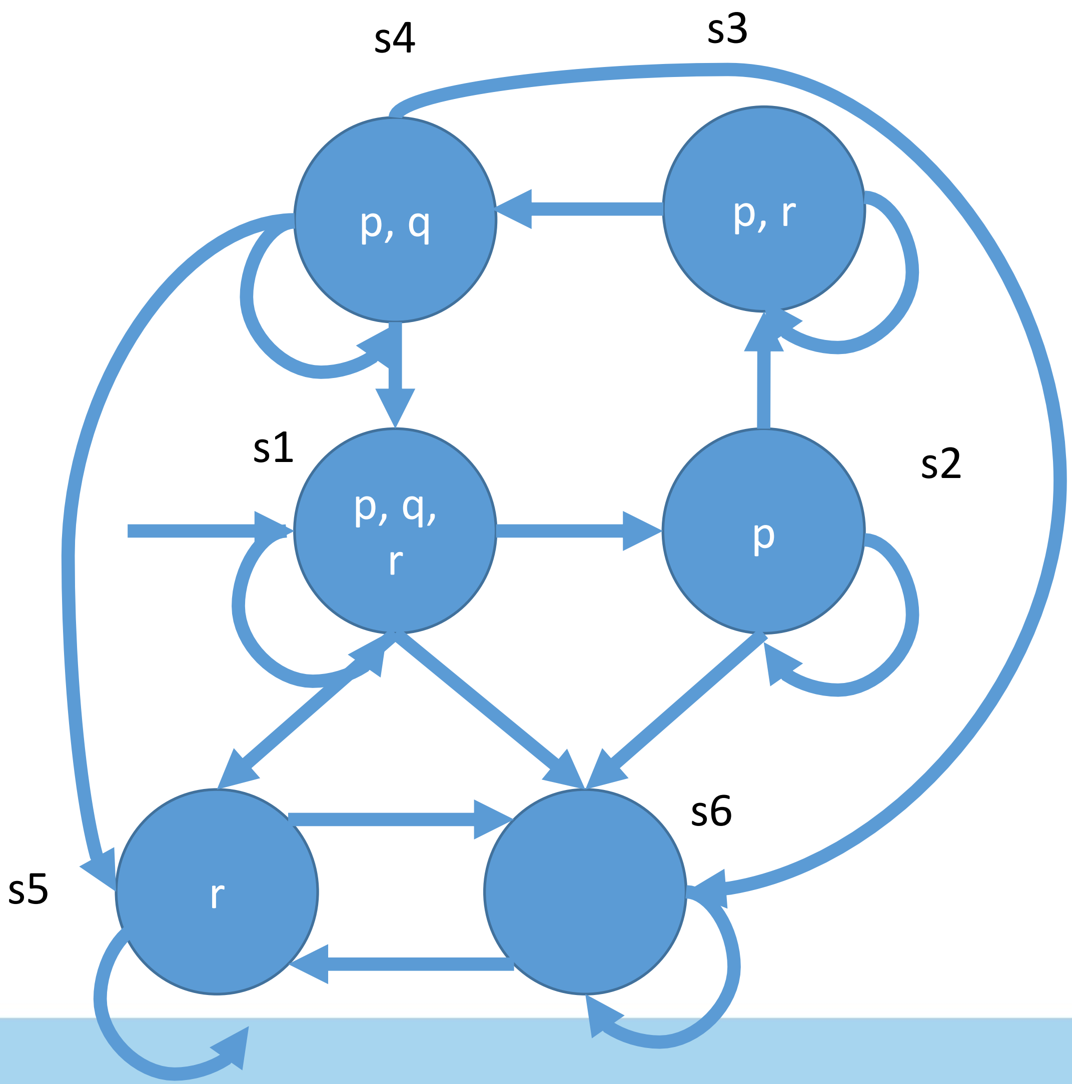
- Each state contains atoms
- Represented discrete time steps via path, e.g.

$$\pi = s_1 s_1 s_2 s_2 s_2 s_3 s_3 \ldots$$

Such that:

$$\pi^3 = s_2 s_2 s_2 s_3 s_3 \ldots$$

What if paths are probabilistic?



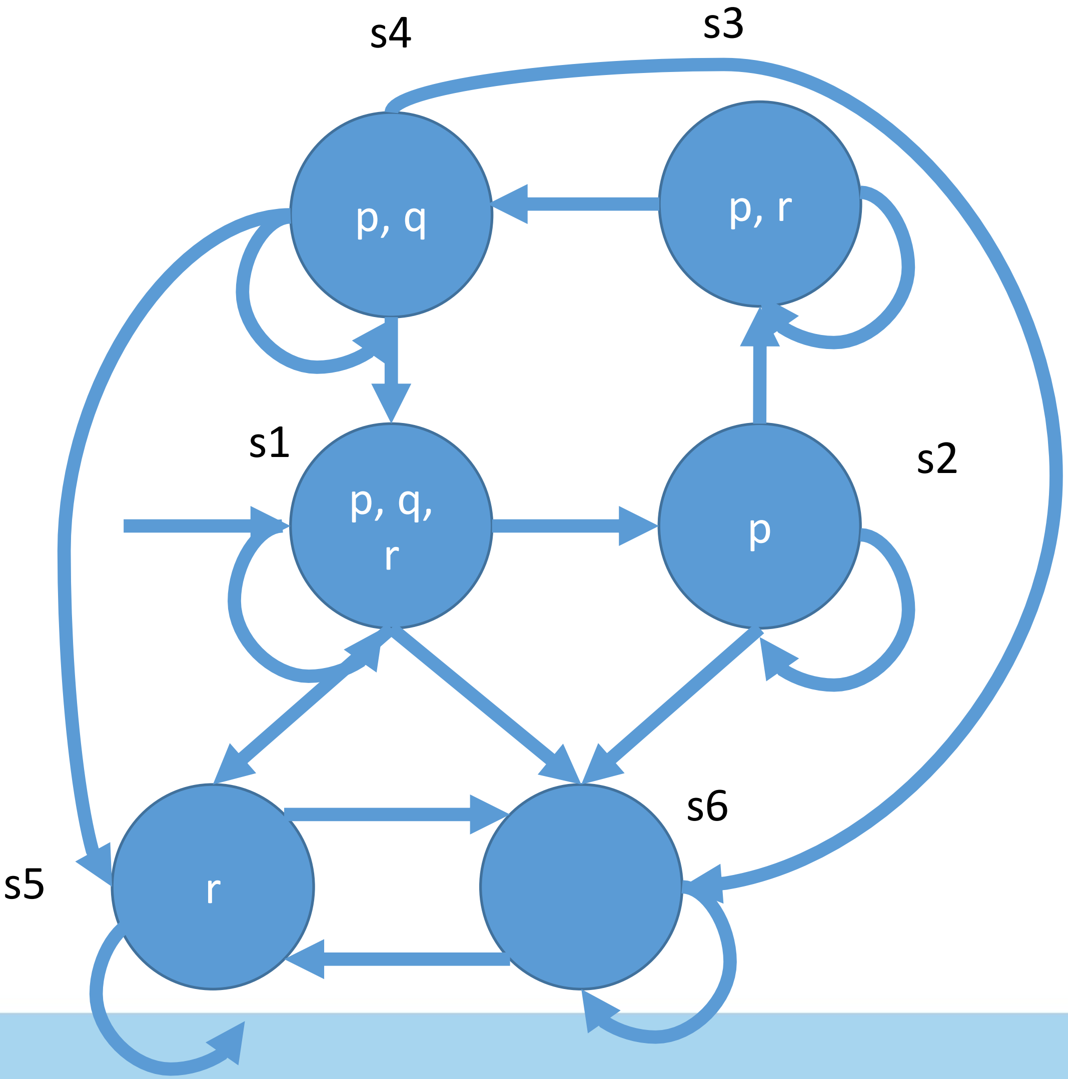
Idea: instead of asking

$$\pi \vDash \phi$$

Ask a probabilistic query:

- $P(\pi)$
- $P(\pi \vDash \phi)$
- $P(\pi \mid \pi \vDash \phi)$

How would we compute these?



What is the event space?

- Attempt 1: All possible combinations of state

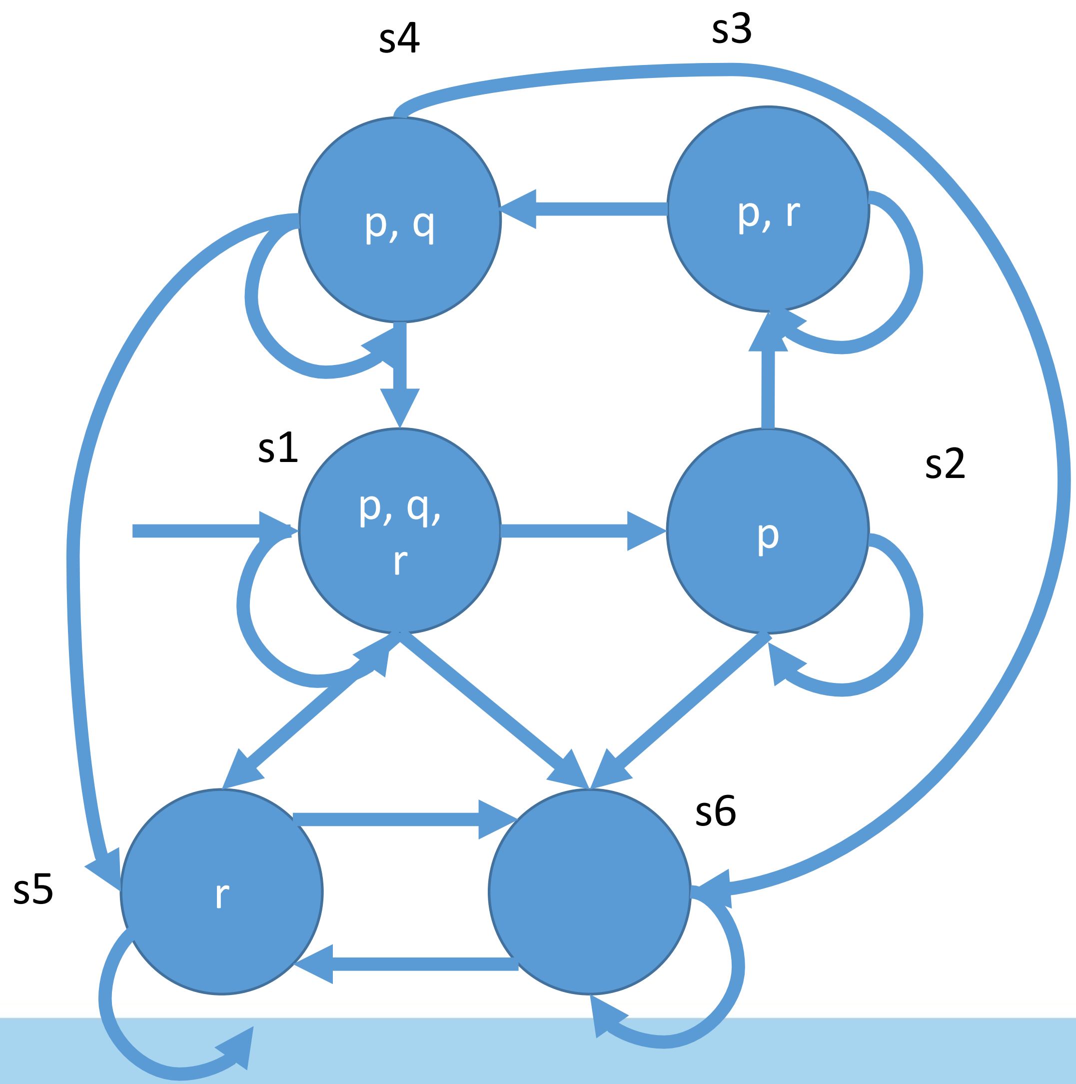
$$|\Omega| = 6! \cdot 6! \cdots$$

Devise an enumeration scheme:

$s_1 s_1 s_1 \ldots$

$\boxed{s_2 s_2 \ldots}$

Not a valid path according to our model!



What is the event space?

- Attempt 2: All valid paths

$$|\Omega| = 1 \cdot 4 \cdot ?$$

(Still infinite, though)

Devise an enumeration scheme:

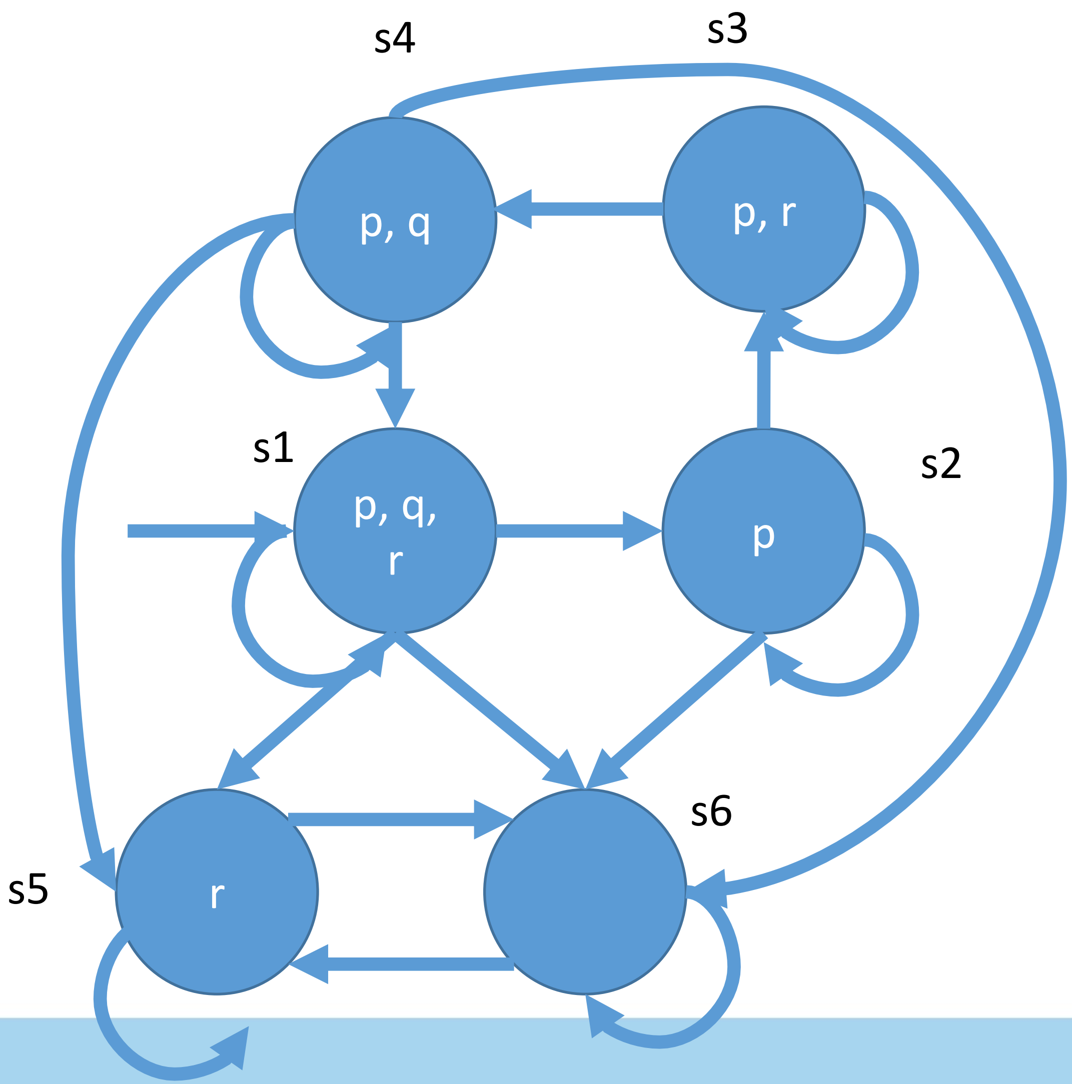
Now what?

$s_1 s_1 s_1$ …

$s_1 s_2 s_2$ …

$s_1 s_2 s_3 s_3$ …

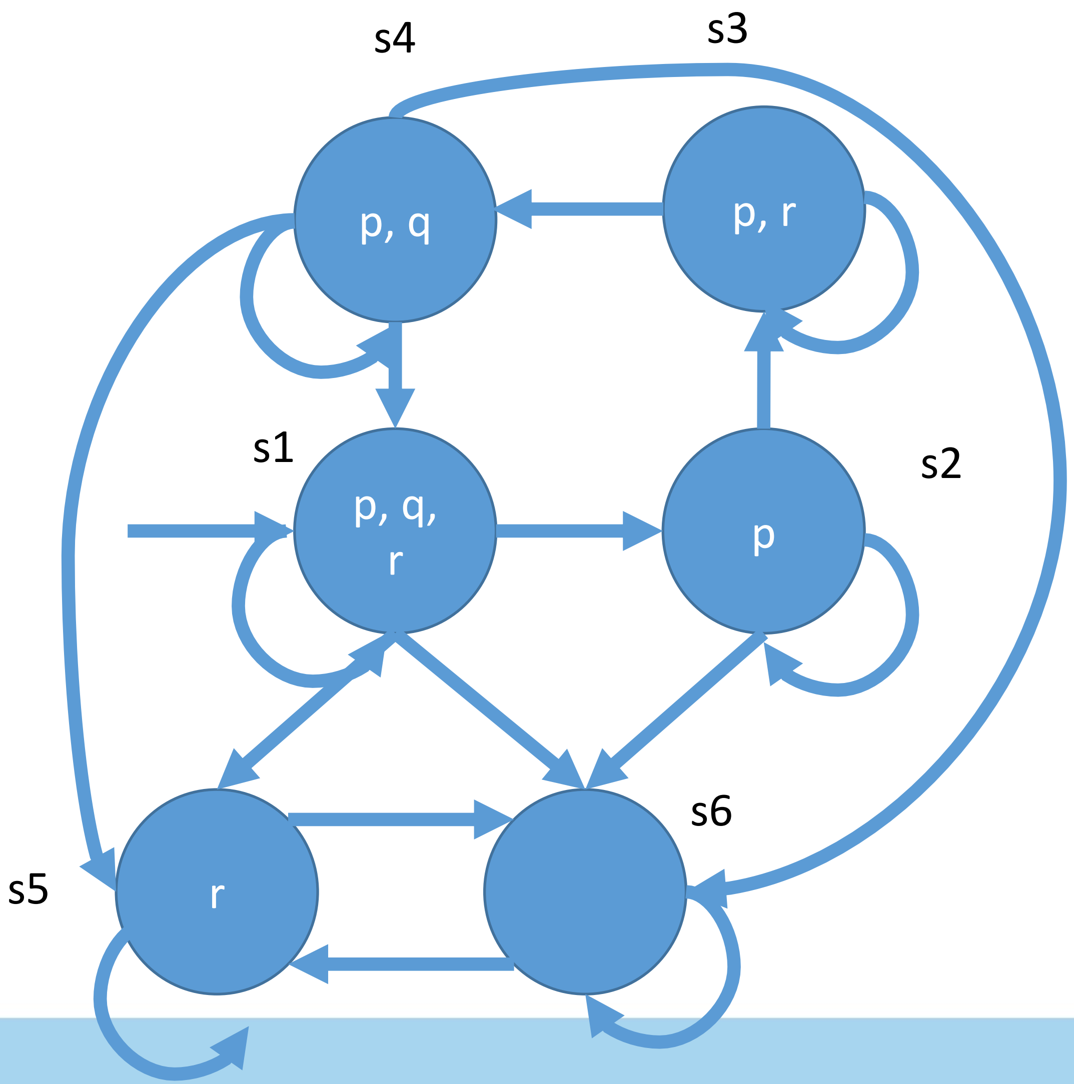
$s_1 s_2 s_6 s_6$ …



What is the event space?

- Attempt 2: All valid paths
  - What is $P(\pi)$?
  - Have Ω and a π
  - Then $P(\pi) = \frac{1}{\Omega}$

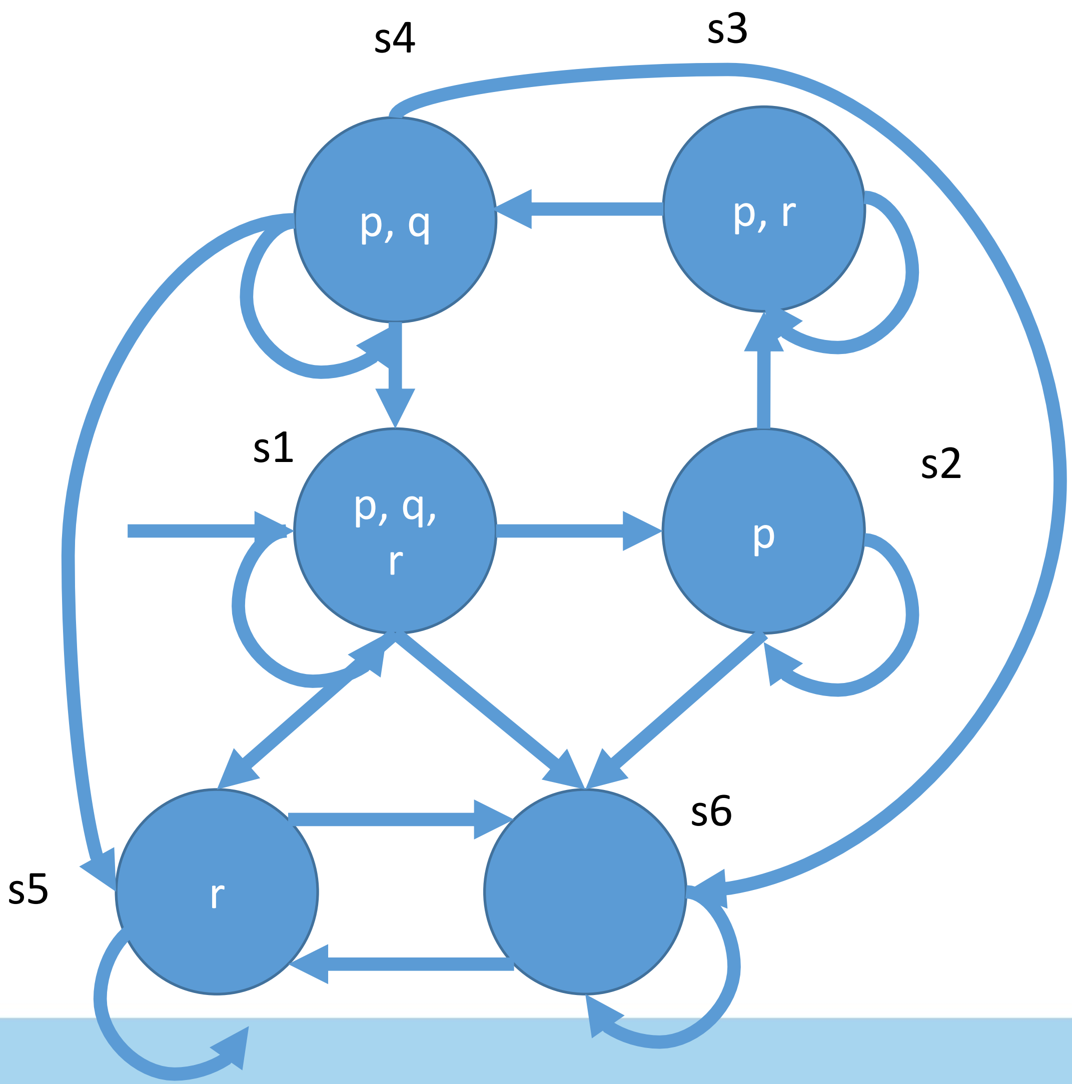
    Is this a useful quantity?



What is the event space?

- Attempt 2: All valid paths
  - What is $P(\pi \vDash \phi)$?
  - Then $P(\pi) = 1$ if $(\pi \vDash \phi)$ and 0 otherwise

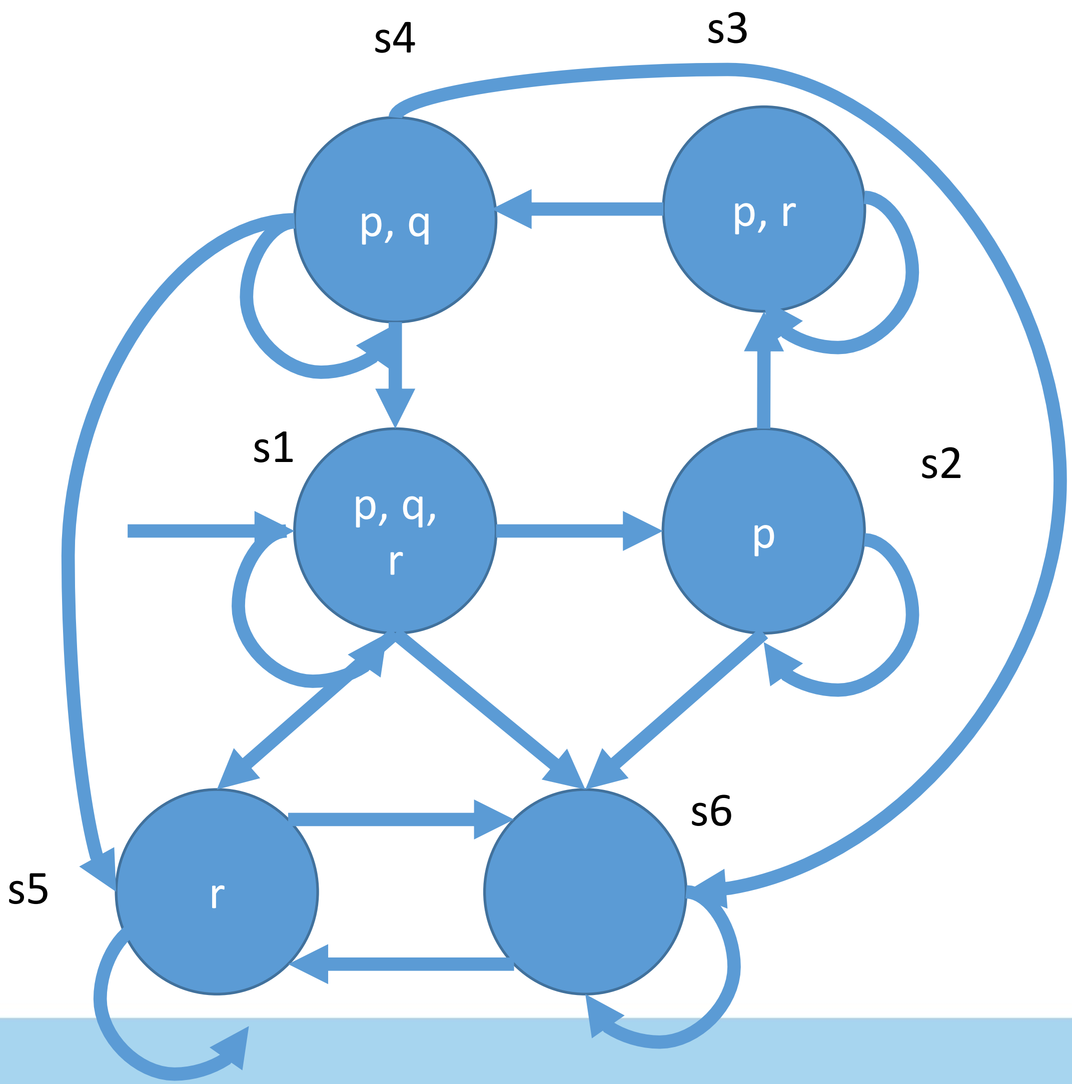
    Is this a useful quantity?



What is the event space?

- Attempt 2: All valid paths
  - What is $P(\pi \mid \pi \vDash \phi)$?
  - This is a conditional probability
    - Need to restrict the event space:
      - $\Omega' = \{\pi' \mid \pi' \vDash \phi\}$
      - Then $P(\pi \mid \pi \vDash \phi) = \frac{1}{\Omega'}$

Is this a useful quantity?



Problem: Assumes all paths equally likely!

$$P(s_1 s_1 \dots) = P(s_1 s_2 \dots)?$$

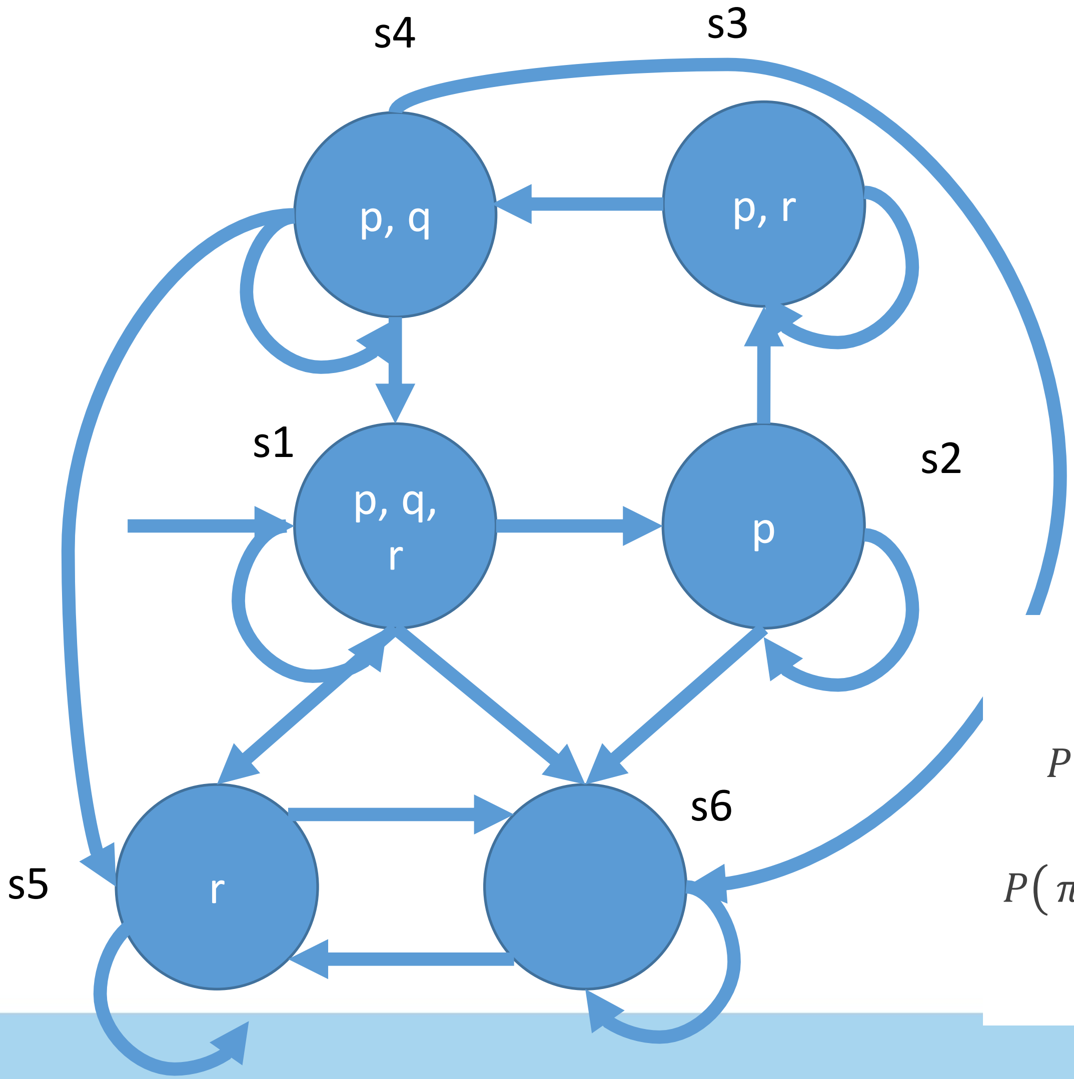
Scenario:

- Assume the transporter is ready.
- The longer she is on the enemy ship, the greater the probability she returns:

Let π^t represent a random path such that

π^t = s means the first state of the path is s

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}\right) >$$
$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2}\right)$$



Scenario:

- Assume the transporter is ready.
- The longer she is on the enemy ship, the greater the probability she returns:

Let π^t represent a random path such that

π^t = s means the first state of the path is s

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}\right) > P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2}\right)$$

However, there may be diminishing returns:

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}\right) - P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2}\right) = \epsilon$$

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \dots, \boldsymbol{\pi^{t-4}} = \boldsymbol{s_3}\right) - P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-4}} = \boldsymbol{s_2}\right) \ll \epsilon$$

Recall: Conditional probability tables

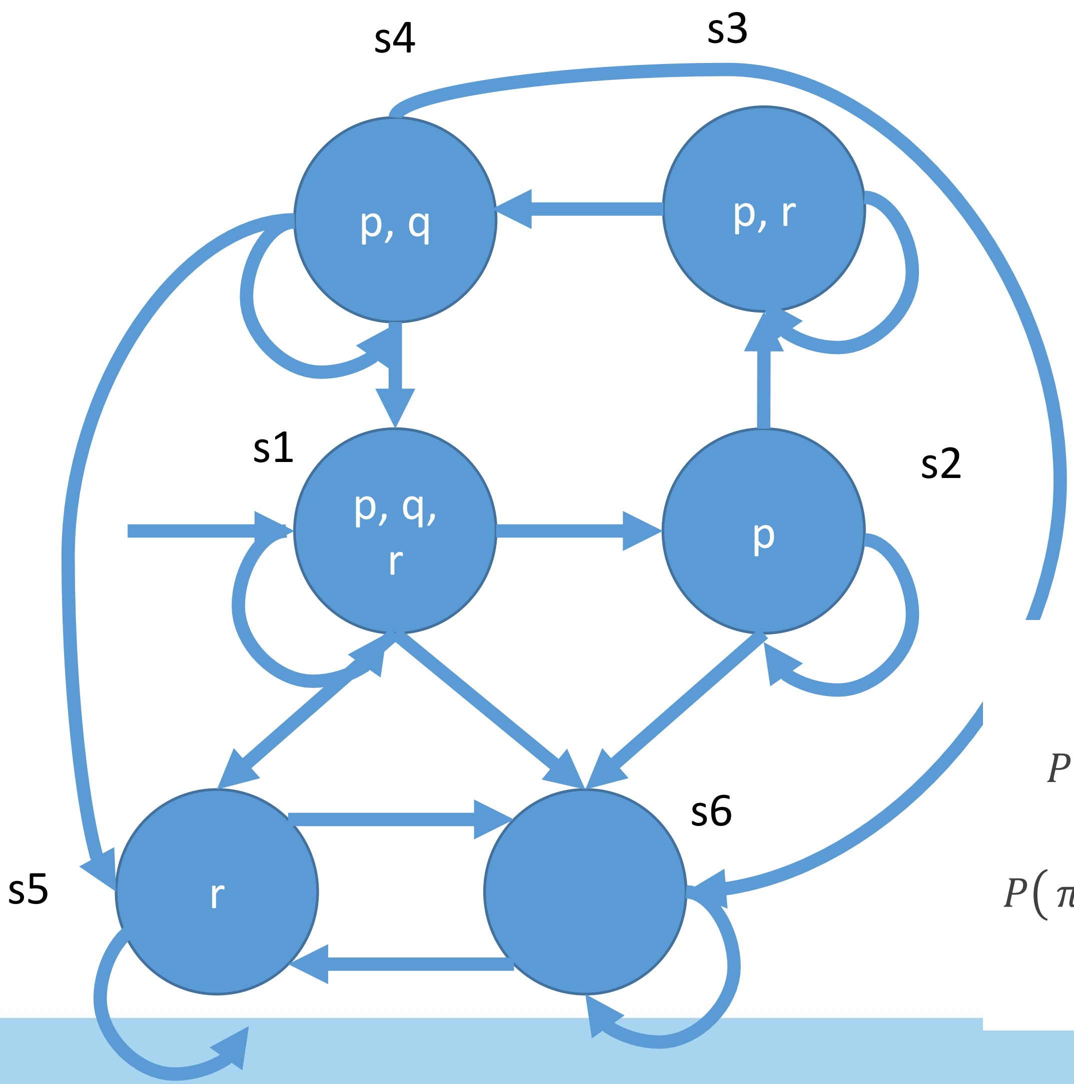
Assume wlog that all variables are binary

$$P(X \mid Y_1, Y_2, \ldots, Y_n)$$

Then the *minimum* number of parameters will be 2^n

(General case: $(|\mathcal{X}| - 1) \cdot (|\mathcal{Y}_1| + |\mathcal{Y}_2| + \cdots + |\mathcal{Y}_n|)$)

Returning to our scenario…



Scenario:

- Assume the transporter is ready.
- The longer she is on the enemy ship, the greater the probability she returns:

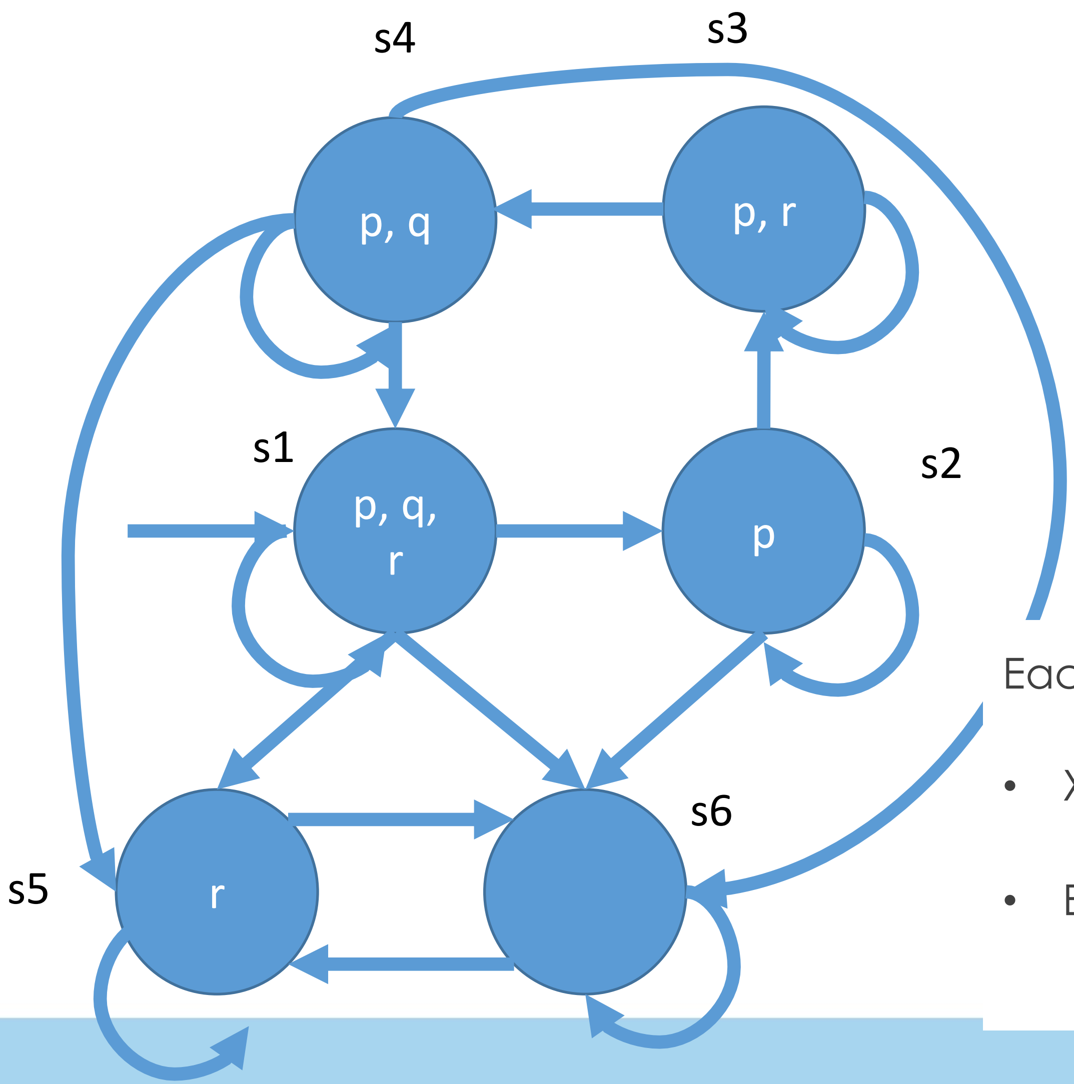
Let π^t represent a random path such that π^t = s means the first state of the path is s

$$P(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}) > P(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2})$$

However, there may be diminishing returns:

$$P(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}) - P(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2}) = \epsilon$$

$$P(\pi^t = s_4 \mid \pi^{t-1} = s_3, \dots, \boldsymbol{\pi^{t-4}} = \boldsymbol{s_3}) - P(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-4}} = \boldsymbol{s_2}) \ll \epsilon$$



Scenario:

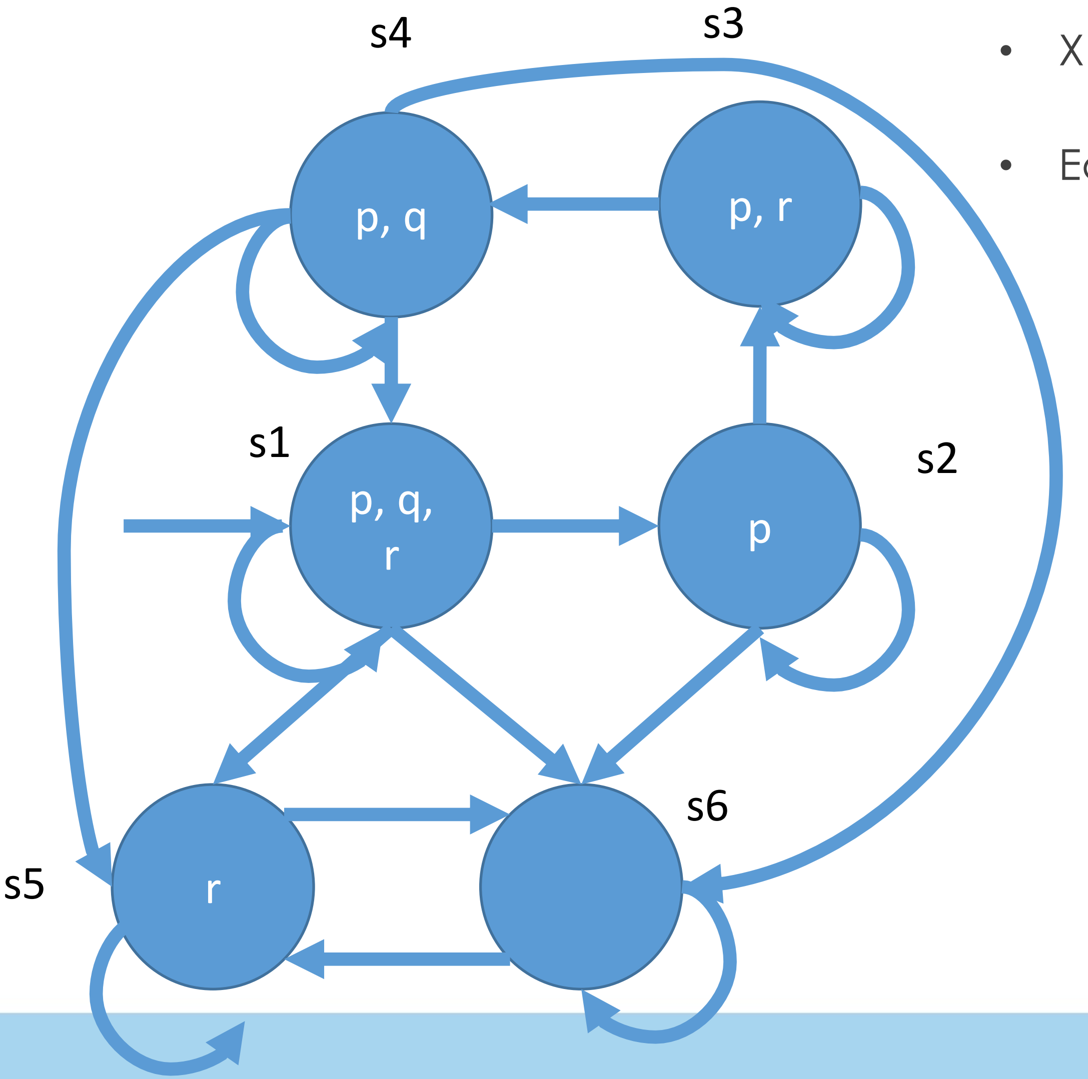
- Assume the transporter is ready.
- The longer she is on the enemy ship, the greater the probability she returns:

Let π^t represent a random path such that π^t = s means the first state of the path is s

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}\right) > P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2}\right)$$

Each variable in $P(X \mid Y_1, Y_2, \ldots, Y_n)$ is actually a state in our model

- X is the current state at time t
- Each Y_i is the state at $t - i$

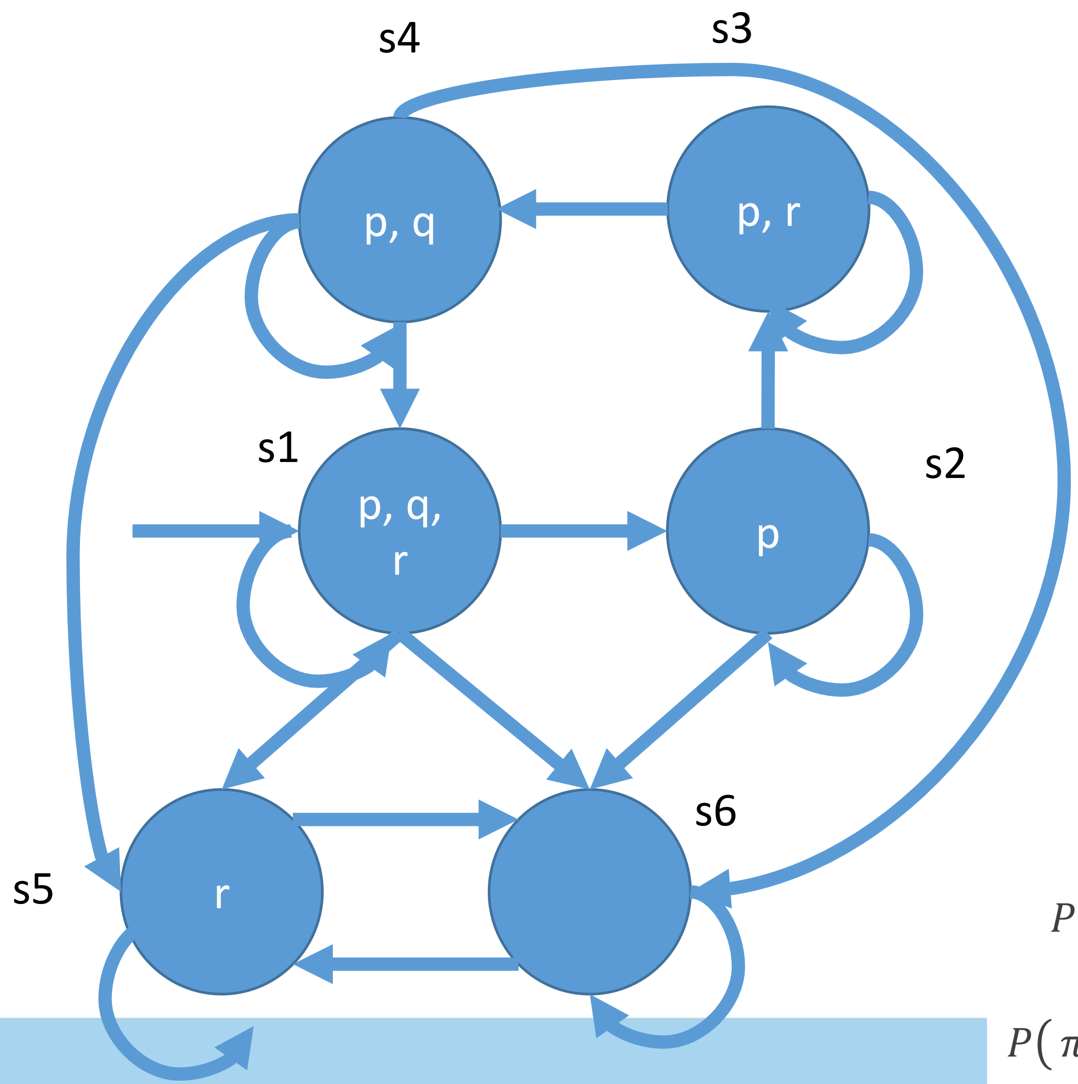


Each variable in $P(X \mid Y_1, Y_2, \dots, Y_n)$ is actually a state in our model

- X is the current state at time t
- Each Y_i is the state at $t - i$

Rather than $P(\pi)$, consider $P(\pi^t \mid \pi^{t-1})$

- Reframing as a conditional probability
- Reframing as a finite subsequence
- Now expressing in terms of a transition



Rather than $P(\pi)$, consider $P(\pi^t \mid \pi^{t-1})$

- Reframing as a conditional probability
- Reframing as a finite subsequence
- Now expressing in terms of a transition

New problem:

- number of parameters can grow with the size of the path (i.e., with *t*)

Recall:

However, there may be diminishing returns:

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_3}\right) - P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-2}} = \boldsymbol{s_2}\right) = \epsilon$$

$$P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \ldots, \boldsymbol{\pi^{t-4}} = \boldsymbol{s_3}\right) - P\left(\pi^t = s_4 \mid \pi^{t-1} = s_3, \boldsymbol{\pi^{t-4}} = \boldsymbol{s_2}\right) \ll \epsilon$$

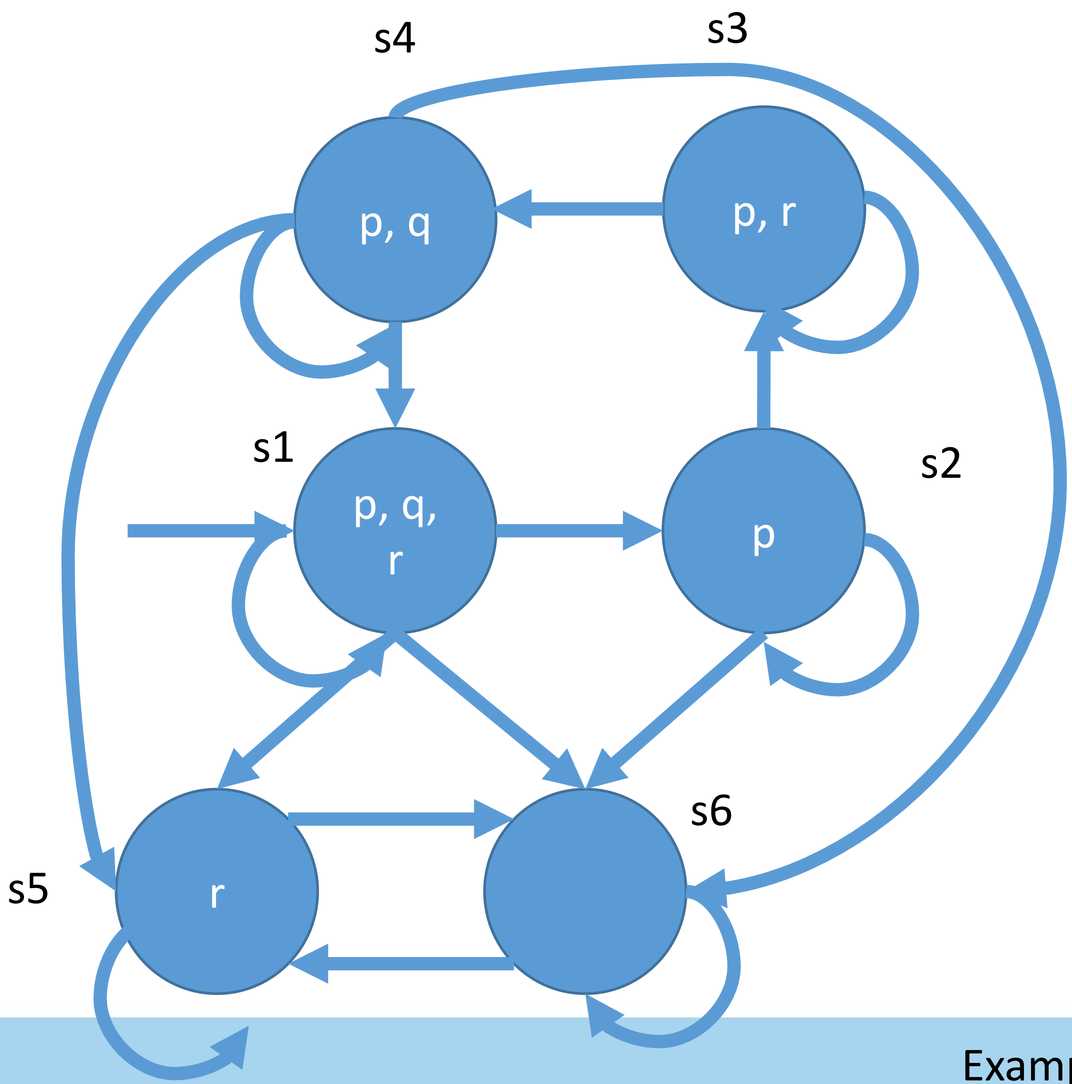
Idea: Simplifying assumption!

Assume there exists some finite window *k* such that only the *k* previous states matter.

- What do we mean by *matter*?
- Independence!

$$P(\pi^t \mid \pi^{t-1}, \ldots, \pi^1) = P\left(\pi^t \mid \pi^{t-1}, \ldots, \pi^{t-k}\right), \qquad k \geq 1$$

Most common setting is k=1



Now: can label transitions with probabilities

Why?

This example: assume all *transitions* are equally likely

Note: previously we assumed all paths equally likely

Example queries on the board

Markov chains

This is an example of a discrete-time Markov chain:

- A sequence of random variables indexed by “time”
- “Markovian” or “Markov property” denotes locality
    - Specifically that $P(X \mid Y_1, \ldots, Y_n) = P(X \mid Y_1)$
- While a random variable, may not be a ”variable” as we normally think of them
    - Compare: Bayes nets

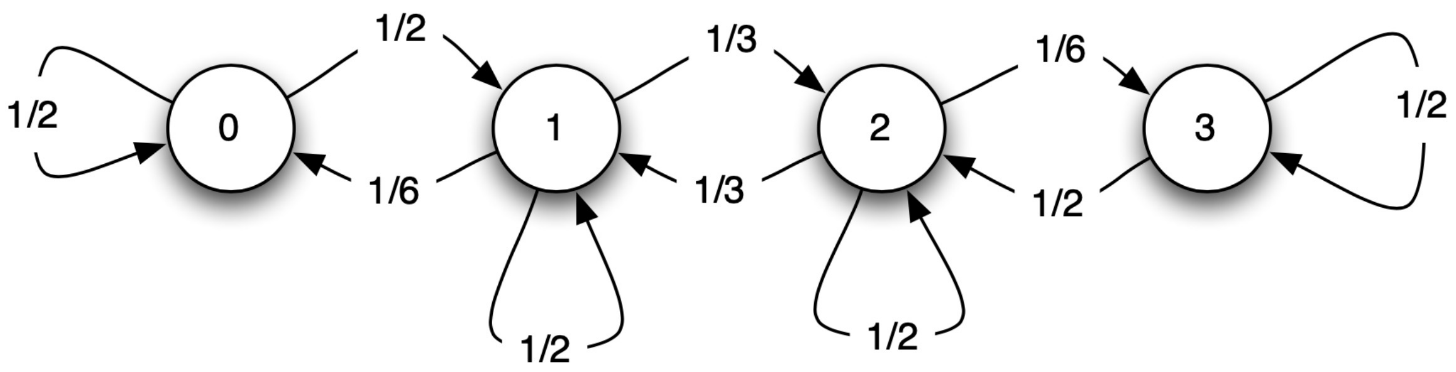
Markov Chains, formally

Another example

- Waiting in line
- Every minute, someone joins…
  - With probability 1 if the line has length 0
  - With probability 2/3 if the line has length 1
  - With probability 1/3 if the line has length 2
  - With probability 0 if the line has length 3
- Every minute the server serves someone with probability ½

Suppose one person is in line at noon. How many people do we expect in line at 12:10?

Another example



- Waiting in line
- Every minute, someone joins…
  - With probability 1 if the line has length 0
  - With probability 2/3 if the line has length 1
  - With probability 1/3 if the line has length 2
  - With probability 0 if the line has length 3
- Every minute the server serves someone with probability ½

Suppose one person is in line at noon. How many people do we expect in line at 12:10?