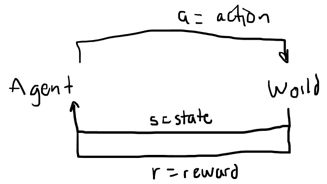
Week 3

Saturday, February 19, 2022 10:40 AM



<u>World</u>

Graph: nodes connected by lines/arrows.
↓
Circles

The world contains a set of States
↳ describes the configuration of the world.

→ states will be the nodes in the graph: $s_1, s_2, \dots s_n$
↳ the n states are visualized through the graph

→ Set of states : $S = \{s_1, \dots, s_n\}$ ie: the $\#^{\text{e}}_{\text{x}}$ state in the game
↓
Capitalized S
each State: s_1, s_n.

→ if we can pass with an ACTION from a state to another, we connect the nodes with a line

→ To define the lines, we thus need to define the ACTIONS

$A = \{a_1, \dots, a_m\}$ → not the same as n (states)

→ We transit from a state to another through an action (a).

$$s_1 \xrightarrow{a} s_2$$

Write an arrow for each action you need to transition from a node to another

The set of States (S) and set of actions (A) are not enough to describe World, we need:

Transition Model : from which node to which node and with what action

1. Deterministic : defined by a two variable function

$$s_f = f_{MT}(s, a)$$

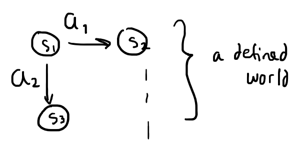
↳ final State ↓ current state ↳ action to do

This function can be seen in a matrix

$$\begin{array}{cc} & \xrightarrow{a} \\ s \downarrow & \begin{bmatrix} \boxed{s_f} \end{bmatrix} \end{array}$$

$s_i a_j = s_f$
↳ the state to which we transition

Example: we have a world:



Thus

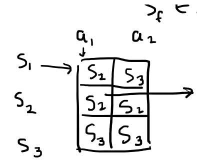
$s \in S = \{s_1, s_2, s_3\}$ → the domain of s is S

$a \in A = \{a_1, a_2\}$ → the domain of a is A

$s_f = f_{MT}(s, a)$
↳ Transition Model Function
↳ another variable whose domain is S also
$s_f \in S$

$$= \begin{array}{ccc} & a_1 & a_2 \\ s_1 \rightarrow & \boxed{s_2} & \boxed{s_3} \\ s & \dots & \end{array}$$ since there's no $s_1 \xleftarrow{a} s_2$,

$s_f \in S$



	a_1	a_2
s_1	s_2	s_3
s_2	s_2	s_2
s_3	s_3	s_3

since there's no $s_1 \xleftarrow{a} s_2$, we stay in the current state

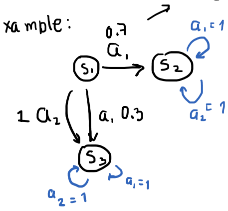
= Transition model function

2. Non-deterministic: it is modeled through a probability distribution that is conditioned.

$$s_f \sim P_{MT}(s_f \mid s, a)$$

you need a probab. distribution per pair of actions possible x

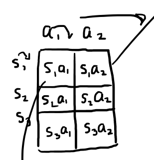
Example:



the a_1 takes us from s_1 to s_2 with a probability of 0.7 (probability of transition)

$s_f = P_{MT}(s_f \mid s, a)$ =

given s,a



$P_{MT}(s_f \mid s_1, a_2)$

if s_1 and a_1 happens we have a probability Distribution of transitioning to all states

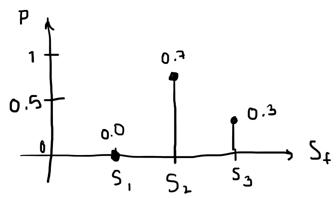
$P_{MT}(s_f \mid s_2, a_1)$

in total, we have 6 probability distribution function to define the world

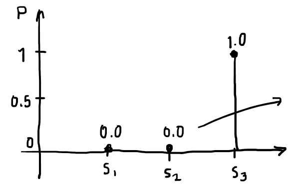
Let's see one of the six prob distrib func:

$P_{MT}(s_f \mid s_1, a_1)$

1)



2)



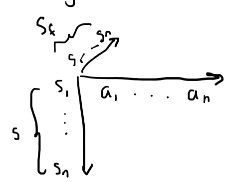
if there's no transition, p = 0.0

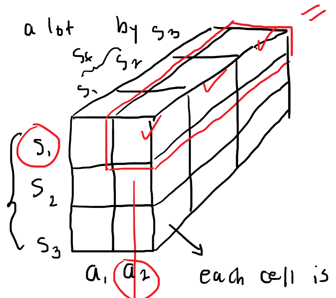
these 3 cubes are this plot

$P_{MT}(s_f \mid s_1, a_2)$

p = 1.0

We can skip plotting a lot by drawing a 3D box:





each cell is a probability value

we can draw this 3D box like:

	a_1	a_2	
s_1	0	✓	= ∅
s_2	0	0	
s_3	0	0	

$s_f = s_1$

	a_1	a_2	
s_1	0.7	✓	= ∅
s_2	1	1	
s_3	0	0	

$s_f = s_2$

	a_1	a_2	
s_1	0.3	✓	= 1
s_2	0	0	
s_3	1	1	

$s_f = s_3$

- Plot 1)
- Plot 2)
- Implicit probability

we say 1 since there's no

transition to S_f with a_n,
thus we stay in current state s
with probability $= 1$.

→ there's a probability of transition since in a NON-DETERMINISTIC world there is a randomness assumed

$s_1 \xrightarrow{a=0.7} s_2$ → th is transition may not happen even though action a happens.

After defining S, A, f_{MT}, we need to define a prize and punish.

normally $\left\langle \begin{array}{l} (+) \text{ prices} \\ (-) \text{ punishes} \end{array} \right\}$ the program aims to accomplish the highest price possible

→ Price: 3 variable function f_R (can also be deterministic / non deterministic)

deterministic ↓ the common

non deterministic ↓ not so useful and makes it very complex

$r_f = f_R(s, a, s_f)$

- s: current state
- a: action done
- s_f: arrival state after action a

We can draw a 4D plot by grouping 3D boxes

We now have defined The World, but how to define the AGENT?
↳ we define how the agent performs the action.

ACTIONS:

NON DET: helps model uncertainty

1. Deterministic $a = f_\pi(s)$

π = Politic, since given a state s it tells which action a to do.

↓ one variable defined as a vector

$$\begin{array}{c} s_1 \\ \vdots \\ s_n \end{array} \begin{bmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_n \end{bmatrix}$$

→ if s_1, we perform $\tilde{a}_1$

2. NON Deterministic:

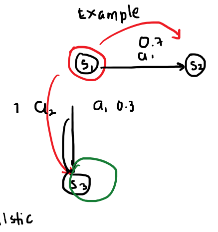
$a \sim P_\pi(a \mid s)$ ↳ given (or conditioned by) s state

$$s \left\{ \begin{array}{c} s_1 \\ \vdots \\ s_n \end{array} \right. \begin{array}{c} \begin{array}{ccc} a_1 & \cdots & a_m \end{array} \\ \begin{bmatrix} P_{11} & \cdots & P_{1m} \\ & \ddots & \\ P_{n1} & \cdots & P_{nm} \end{bmatrix} \end{array} \begin{array}{l} = 1 \\ \\ = 1 \end{array}$$

→ given s_1, we have m probabilities since we have m possible actions

} each row sums 1

Example



Deterministic action

Transition model: $P_{MT}(s_f \mid s, a)$

	a_1	a_2
s_1		
s_2		
s_3	0	

$S_f = s_1$

	a_1	a_2
s_1		
s_2		
s_3	0	

$S_f = s_2$

	a_1	a_2
s_1		
s_2		
s_3	1	

$S_f = s_3$

deterministic (its result is nondet.)

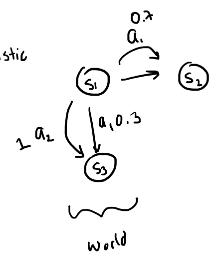
30% s_3
70% s_2

$$a = f_\pi(s) = \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} \begin{bmatrix} a_1 \\ a_2 \\ a_1 \end{bmatrix}$$

we defined arbitrarily what actions are done given a state

we stay in s_3 if we perform a_1 after seeing the tables

Non deterministic action



Same $P_{MT}(s_f \mid s, a)$

$a \sim P_\pi(a \mid s)$ = Probability conditioned to s

	a_1	a_2
s_1	0.6	0.4
s_2	0.5	0.5
s_3	1	0

$P_\pi(a \mid s_1) = 1$

$= 1$

$P_\pi(a \mid s_3) = 1$

These are the probabilities that define the stochastic action a

These are ...

world

$P_\pi(a|S_3) = 1$

action a

These are the probabilities of the agent (inside the world)

→ How does the agent acts inside the world?

↳ Suppose we first position agent in S_1, thus the probability of the agent taking action a_1 is 60% and a_2 is 40% (decided by a random generator)

$\underbrace{\text{action } a_1 \text{ is } 60\%}_{\text{you can go either to state } S_2 \text{ or } S_3}$

$\underbrace{a_2 \text{ is } 40\%}_{\text{directly goes to } S_3}$

→ The probability of arriving to S_2 from S_1 is $= \underset{\text{probability of taking action } a_1}{(0.6)} \underset{\text{prob given } a_1 \text{ of arriving to } S_2}{(0.7)} = \boxed{0.42}$