

May 4

①

2	-0.04	+1
1	-0.04	-1
	1	2

$$p(1,2 | 1,1, u_p) = .8$$

$$p(2,1 | 1,1, u_p) = .1$$

$$p(1,1 | 1,1, u_p) = .1$$

$$Y = .9$$

u	u'	u	u'
0	0	0	0
0	-0.04	0	1
-0.04	.709	1	.81
0	0	0	0
-0.04	-0.04	0	-1
-0.04	-0.162	-1	-0.3736

$$-0.04 + .9(.8 \times 0 + .1 \times 0 + .1 \times 0) = -0.04$$

$$-0.04 + .9(.8 \times 1 + .1 \times -0.04 + .1 \times -0.04) = .6728$$

$$1 + .9(.8 \times 1 + .1 \times 1 + .1 \times -0.04) = 1.8064$$

$$-1 + .9(.8 \times 1 + .1 \times -0.04 + .1 \times -1) = -0.3736$$

$$-0.04 + .9(.8 \times -0.04 + .1 \times -0.04 + .1 \times -1) = -0.1624$$

(1)

Policy-Iteration ( $S, A, P, R, V$ )

For all  $s \in S$   $\pi(s) =$  random legal action,  $U(s) = 0$

Repeat

$U \leftarrow$  Policy Eval ( $S, A, P, R, V, \pi, U$ )

For each  $s \in S$

if  $\max_{a \in A} \sum_{s' \in S} P(s'|s, a) U(s') > \sum_{s' \in S} P(s'|s, \pi(s)) U(s')$   
where  $a \leftarrow \pi(s)$

then  $\pi(s) \leftarrow \operatorname{argmax}_{a \in A} \sum_{s' \in S} P(s'|s, a) U(s')$

Until  $\pi$  doesn't change

Partially Observable MDPs (POMDPs)

Still enumerate states

Don't observe states, get information on possible

states you may be in

Maintain belief state - prob. distribution over states,

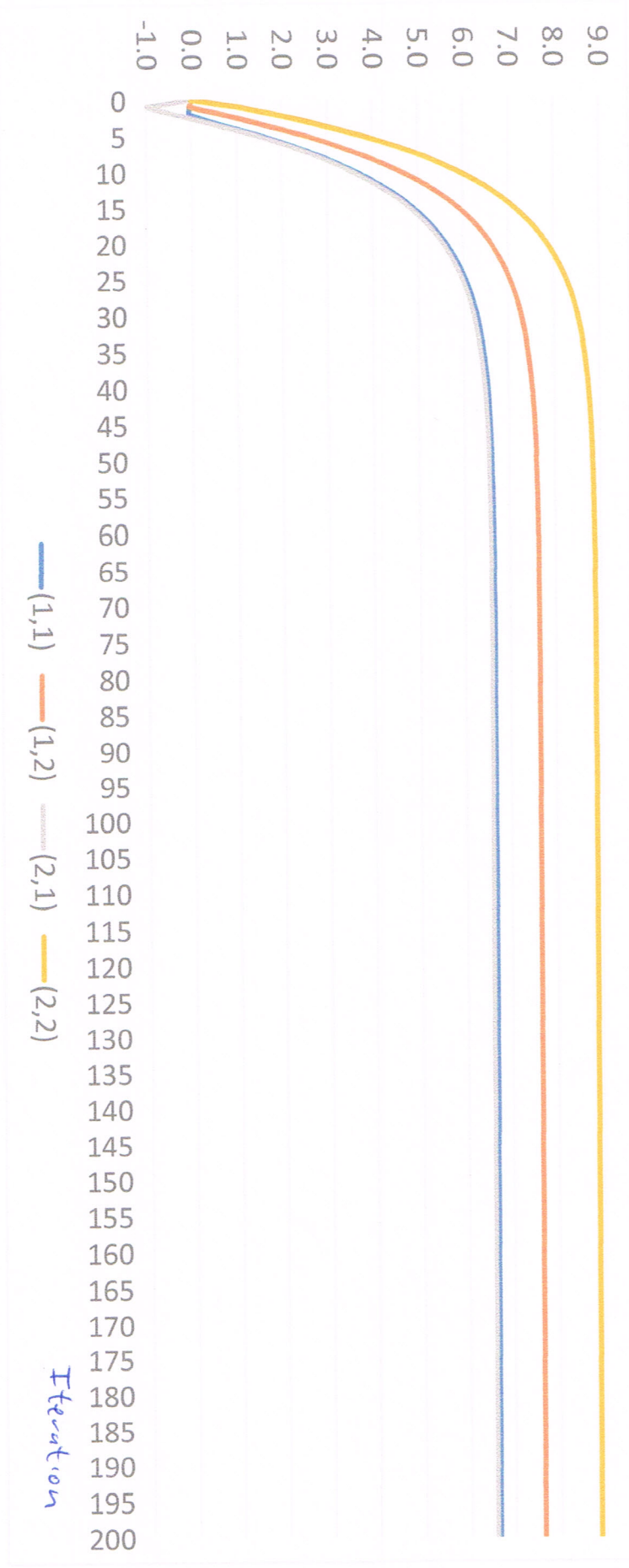
Treat these as states, solve MDP over these states

Iteration #

Trace of Value Iteration

Iteration	$u(1,1)$	$u(1,2)$	$u(2,1)$	$u(2,2)$	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration	Iteration
1	0.000	0.000	0.000	0.000	33	6.385	7.374	6.319	8.614	67	6.647	7.635	6.580	8.876						
2	-0.040	-0.040	-1.000	1.000	34	6.412	7.400	6.345	8.641	68	6.648	7.636	6.581	8.877						
3	0.407	0.673	-0.374	1.806	35	6.437	7.425	6.370	8.665	69	6.649	7.637	6.582	8.877						
4	0.968	1.317	0.263	2.524	36	6.458	7.447	6.392	8.687	70	6.649	7.637	6.582	8.878						
5	1.517	2.498	1.443	3.163	37	6.478	7.466	6.411	8.707	71	6.650	7.638	6.583	8.879						
6	2.025	3.011	1.956	4.251	38	6.496	7.484	6.429	8.724	72	6.650	7.638	6.584	8.879						
7	2.486	3.474	2.419	4.714	39	6.512	7.500	6.445	8.740	73	6.651	7.639	6.584	8.879						
8	2.903	3.891	2.836	5.131	40	6.526	7.514	6.459	8.755	74	6.651	7.639	6.584	8.880						
9	3.278	4.266	3.211	5.506	41	6.539	7.527	6.472	8.768	75	6.651	7.640	6.585	8.880						
10	3.615	4.603	3.549	5.844	42	6.550	7.539	6.484	8.779	76	6.652	7.640	6.585	8.881						
11	3.919	4.907	3.852	6.148	43	6.561	7.549	6.494	8.790	77	6.652	7.640	6.585	8.881						
12	4.193	5.181	4.126	6.422	44	6.570	7.558	6.503	8.799	78	6.652	7.641	6.586	8.881						
13	4.439	5.427	4.372	6.668	45	6.579	7.567	6.512	8.807	79	6.653	7.641	6.586	8.881						
14	4.661	5.649	4.594	6.889	46	6.586	7.574	6.519	8.815	80	6.653	7.641	6.586	8.882						
15	4.860	5.848	4.793	7.089	47	6.593	7.581	6.526	8.822	81	6.653	7.641	6.586	8.882						
16	5.039	6.028	4.973	7.268	48	6.599	7.587	6.532	8.828	82	6.653	7.641	6.586	8.882						
17	5.201	6.189	5.134	7.430	49	6.605	7.593	6.538	8.834	83	6.653	7.641	6.587	8.882						
18	5.346	6.335	5.280	7.575	50	6.610	7.598	6.543	8.839	84	6.653	7.642	6.587	8.882						
19	5.477	6.465	5.410	7.706	51	6.614	7.602	6.548	8.843	85	6.654	7.642	6.587	8.882						
20	5.595	6.583	5.528	7.824	52	6.618	7.606	6.552	8.847	86	6.654	7.642	6.587	8.882						
21	5.701	6.689	5.634	7.930	53	6.622	7.610	6.555	8.851	87	6.654	7.642	6.587	8.883						
22	5.796	6.784	5.730	8.025	54	6.625	7.613	6.558	8.854	88	6.654	7.642	6.587	8.883						
23	5.882	6.870	5.815	8.111	55	6.628	7.616	6.561	8.857	89	6.654	7.642	6.587	8.883						
24	5.959	6.948	5.893	8.188	56	6.631	7.619	6.564	8.860	90	6.654	7.642	6.587	8.883						
25	6.029	7.017	5.962	8.258	57	6.633	7.621	6.566	8.862	91	6.654	7.642	6.587	8.883						
26	6.092	7.080	6.025	8.320	58	6.635	7.624	6.569	8.864	92	6.654	7.642	6.587	8.883						
27	6.148	7.136	6.081	8.377	59	6.637	7.625	6.571	8.866	93	6.654	7.642	6.587	8.883						
28	6.199	7.187	6.132	8.427	60	6.639	7.627	6.572	8.868	94	6.654	7.642	6.588	8.883						
29	6.244	7.232	6.177	8.473	61	6.641	7.629	6.574	8.869	95	6.654	7.642	6.588	8.883						
30	6.285	7.273	6.218	8.514	62	6.642	7.630	6.575	8.871	96	6.654	7.643	6.588	8.883						
31	6.322	7.310	6.255	8.551	63	6.643	7.631	6.577	8.872	97	6.654	7.643	6.588	8.883						
32	6.355	7.344	6.289	8.584	64	6.644	7.633	6.578	8.873	98	6.654	7.643	6.588	8.883						
					65	6.645	7.634	6.579	8.874	99	6.654	7.643	6.588	8.883						
					66	6.646	7.635	6.580	8.875	100	6.654	7.643	6.588	8.883						

$u$



(1,1) (1,2) (2,1) (2,2)

Iteration