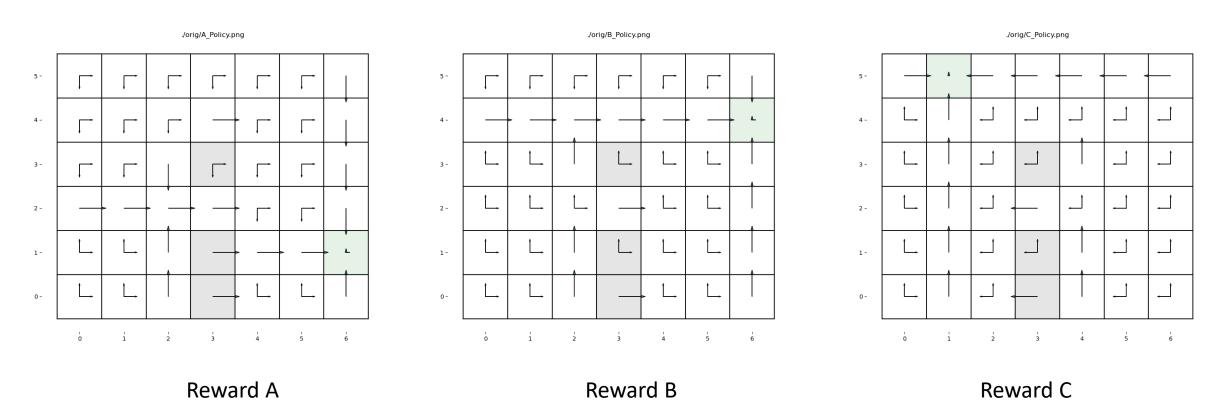
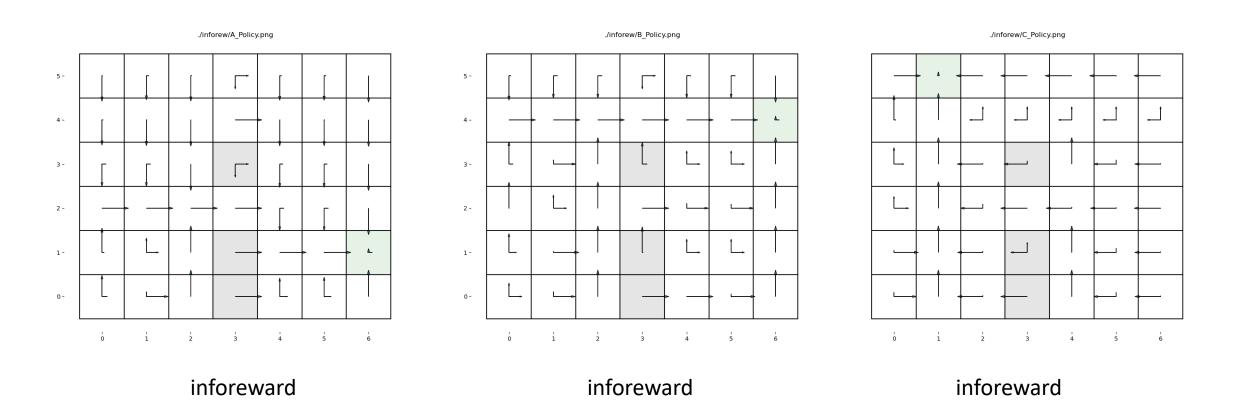
Ashish Kumar Singh (UID:105479019)

### Original Policy Visualization



Ashish Kumar Singh (UID:105479019)

### New Policy with InfoReward Visualization



Ashish Kumar Singh (UID:105479019)

### **Changes in new policy wrt original**

#### Goal A

We see that states nearby Goal B and Goal C shows major changes in policy. Specifically, states like (4,2),(4,3),(5,2),(5,3) could transition to both right and down in original policy. But in new signaling policy, down is highly preferred for these states as downward direction shows that we are moving towards Goal A and helps the observer to differentiate between goal pursued.

#### Goal B

We see that states nearby Goal A and Goal C shows major changes in policy. Specifically, state (1,3) prefers to go right in new signalling policy as going up could confuse the observer with goal C. For state (5,2) new policy prefers to go right so get maximum reward as well as not confusing with goal A as for goal A policy would have preferred going down.

#### Goal C

We see that states nearby Goal A and Goal B shows major changes in policy. Specifically, state (5,3) and (5,2) prefers to go left in new signalling policy as it signals that goal are not A or B.

Overall, we see that new policy reflects signaling as many states move to a direction in which it becomes very clear what the goal is.

# Ashish Kumar Singh (UID:105479019)

### Original Value Visualization

./orig/A_Value.png							/orig/B_Value.png							/orig/C_Value.png									
5 -	28.006	32.229	36.921	42.134	47.927	54.363	61.515	5 -	42.134	47.927	54.363	61.515	69.461	78.29	88.1	5 -	88.1	99.0	88.1	78.29	69.461	61.515	54.363
4 -	32.229	36.921	42.134	47.927	54.363	61.515	69.461	4 -	47.927	54.363	61.515	69.461	78.29	88.1	99.0	4-	78.29	88.1	78.29	69.461	61.515	54.363	47.927
3 -	36.921	42.134	47.927	-44.637	61.515	69.461	78.29	3-	42.134	47.927	54.363	-37.485	69.461	78.29	88.1	3-	69.461	78.29	69.461	-37.485	54.363	47.927	42.134
2 -	42.134	47.927	54.363	61.515	69.461	78.29	88.1	2 -	36.921	42.134	47.927	54.363	61.515	69.461	78.29	2 -	61.515	69.461	61.515	54.363	47.927	42.134	36.921
1-	36.921	42.134	47.927	-29.539	78.29	88.1	99.0	1-	32.229	36.921	42.134	-51.073	54.363	61.515	69.461	1-	54.363	61.515	54.363	-51.073	42.134	36.921	32.229
0 -	32.229	36.921	42.134	-37.485	69.461	78.29	88.1	0 -	28.006	32.229	36.921	-56.866	47.927	54.363	61.515	0-	47.927	54.363	47.927	-56.866	36.921	32.229	28.006
	Ö	1	2	3	4	5	6	•	ó	1	2	3	4	5	6		0	i	2	3	4	5	6

Reward A Reward B Reward C

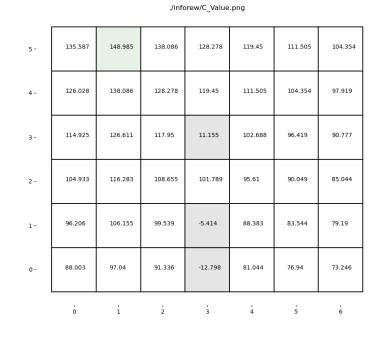
# Ashish Kumar Singh (UID:105479019)

./inforew/B Value.png

#### New Value with InfoReward Visualization

	/inforew/A_Value.png													
5 -	68.908	74.423	79.876	85.507	93.342	100.761	109.005							
4 -	74.899	81.026	87.085	93.342	102.046	110.29	119.45							
3 -	78.776	85.585	92.317	0.796	108.94	118.1	128.278							
2 -	83.085	90.65	98.129	106.44	116.6	126.778	138.087							
1 -	75.443	81.885	89.317	15.934	125.111	136.42	148.985							
0 -	68.566	74.747	81.385	4.065	112.85	123.445	135.087							
	0	1	2	3	4	5	6							

5 -	81.31	88.677	96.03	104.2	114.111	124.29	135.6
4 -	88.677	96.863	105.033	114.111	125.123	136.433	149.0
3 -	81.309	88.677	96.863	5.2	113.361	124.29	136.433
2 -	74.678	80.476	88.677	95.28	104.2	114.111	125.123
1-	67.877	74.228	80.809	-11.748	94.447	104.2	114.944
0 -	61.755	67.855	73.728	-18.791	87.455	95.505	104.45
	6	1	2	3	4	5	6



Reward A + inforeward

Reward B + inforeward

Reward C + inforeward