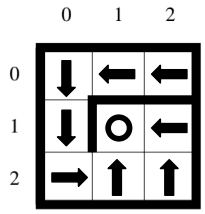
1. Friendlier Pacman

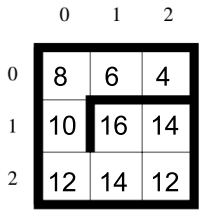
1. Assume finite horizon of 8 (so Pacman takes exactly 8 actions) and no discounting ($\gamma = 1$). Now, for each state...

...fill an the optimal policy:



(available actions: ↑, ↓, →, ←, ∘)

...and fill in the corresponding V*(s) value (utility):



Hint: Pacman gets a reward for taking any action while in (1,1), and is not rewarded for the action to enter that state.

1.1.

1.2. Q values:

1.2.1. 4

1.2.2. 14

1.3. No Actions to stay in place:

1.3.1. False

1.3.2. 8

1.4. (1/(1-0.5)) * 2 = 4

2. Markov Decision Processes

2.1.

State	A	B	C	D	E
K	4	3	2	1	0

2.2.

State	A	B	C	D	E
K	∞	∞	∞	1	0

(a) Mark **all** of the statements that must be true for any MDP.

☐ For no state s and for all policies π, $V^*(s) \geq V^\pi(s)$

☒ For some state s and some policy π, $V^*(s) \geq V^\pi(s)$

☒ For all states s and all policies π, $V^*(s) \geq V^\pi(s)$

○ None of the above

(b) Mark **all** of the statements that are true for value iteration.

☐ Each iteration of value iteration produces a value function that has higher value than the prior value functions **for all states**.

☒ Value iteration can produce a value function that has higher value than the earlier value functions **for some state**.

☐ Each iteration of value iteration produces a value function that has lower value than the prior value functions **for all states**.

☒ Each iteration of value iteration produces a value function that has value at least as high as the prior value functions **for all states**.

○ None of the above

2.3.

(a) ○ $\frac{1}{2}$ ○ $\frac{1}{3}$ ○ 2 ○ -1 ○ 3 ○ γ ✖ None

(b) ✖ $\max_a$ ○ $\min_a$ ○ None

(c) ○ $T(s', a, s'')$ ✖ $T(s, a, s')$ ○ None

(d) ✖ $R(s, a, s')+$ ○ $R(s, a, s')-$ ○ None

(e) ○ $\max_a$ ○ $\min_a$ ○ $\max_{a'}$ ✖ $\min_{a'}$ ○ $\gamma \max_a$ ○ $\gamma \min_a$ ○ $\gamma \max_{a'}$ ○ $\gamma \min_{a'}$ ○ None

(f) ○ $\sum_{s'}$ ✖ $\sum_{s'} T(s, a, s')$ ○ $\sum_{s''}$ ○ $\sum_{s''} T(s', a', s'')$ ○ None

(g) ○ $R(s, a, s')+$ ○ $R(s, a, s')-$
○ $R(s', a', s'')+$ ○ $R(s', a', s'')-$ ✖ None

(h) ○ $\frac{1}{2}V^*(s')$ ✖ $\gamma V^*(s')$ ○ $V^*(s')$ ○ $\frac{1}{2}V^*(s'')$
○ $\gamma V^*(s'')$ ○ $V^*(s'')$ ○ None

2.4.