CS-E4800 AI Exercises 5: Symbolic Search, MDPs February 21, 2017

1 Exercise

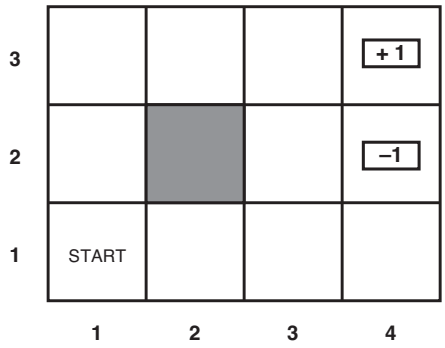
Consider the following system.

State variables: $X = \{b, c\}$
Initial state : $b \vee \neg c$
Transition relation: $(b@0 \leftrightarrow c@1) \wedge (c@0 \leftrightarrow b@1)$

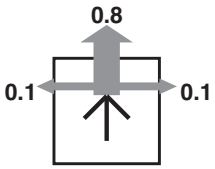
Derive a formula that represents those states that are reachable by two steps with the transition relation, by using the logic-based image operation.

2 Exercise

For the 4×3 world shown in figure below, calculate which squares can be reached from START by the action sequence [Up, Up, Right, Right, Right] and with what probabilities. The intended outcome occurs with probability 0.8, but with probability 0.2 the agent moves at right angles to the intended direction. Collision with a wall (including the walls of the gray cell) results in no movement.



(a)



(b)

3 Exercise

Sometimes Markov decision processes (MDPs) are formulated with a reward function $R(s)$, $R(s, a)$, or $R(s, a, s')$.

- Write (and simplify) Bellman equations for these formulations.
- Show how an MDP with reward function $R(s, a, s')$ can be transformed into a different MDP with reward function $R(s, a)$, such that optimal policies in the new MDP correspond exactly to optimal policies in the original MDP.
  (Extra exercise: Do the same to convert MDPs with $R(s, a)$ into MDPs with $R(s)$.)

4 Exercise

Consider an undiscounted ($\gamma = 1$) MDP having three states $s \in \{1, 2, 3\}$, with rewards $R(s)$ = -1, -2, 0, respectively. State 3 is a terminal state. In states 1 and 2 there are two possible actions: a and b. The transition model is as follows:

- In state 1, action a moves the agent to state 2 with probability 0.8 and makes the agent stay put with probability 0.2.
- In state 2, action a moves the agent to state 1 with probability 0.8 and makes the agent stay put with probability 0.2.
- Action b moves the agent to state 3 with probability 0.1 and makes the agent stay put with probability 0.9.

Answer the following questions:

- What can be determined qualitatively about the optimal policy in states 1 and 2?
- Apply policy iteration, showing each step in full, to determine the optimal policy and the values of states 1 and 2. Assume that the initial policy has action b in both states.