CSCI373 : Artificial Intelligence (Fall 2016)

HW3

0. **Due on 10/26/16 at 8:00am**. Electronic submission only, as described below.
1. Submit a zip file containing a pdf **solutions file** and **source code** via "*turnin –c 373 hw3sub.zip*".
2. This homework will **not** be peer-assessed.

[MDP] You are a cute adorable bunny rabbit (in B1) on the north bank of a river, eyeing some yummy carrots in the south bank (in B4) (see Fig 1a). There are three terminal states (indicated by the boxed final reward). Though you are a fine swimmer (coming from a long, proud line of president-attacking swamp rabbits), the river is treacherous, and whenever you are in the river and try to move to a neighboring square (even onto grass), there is a 25% chance you will be swept one square west towards the waterfall, instead of your intended move (E.g., at B2, you have 25% chance of moving to A2.). Falling off the waterfall will kill you! Note that moving on grass (even into water) is always reliable (you can get to the target cell 100% of the time).

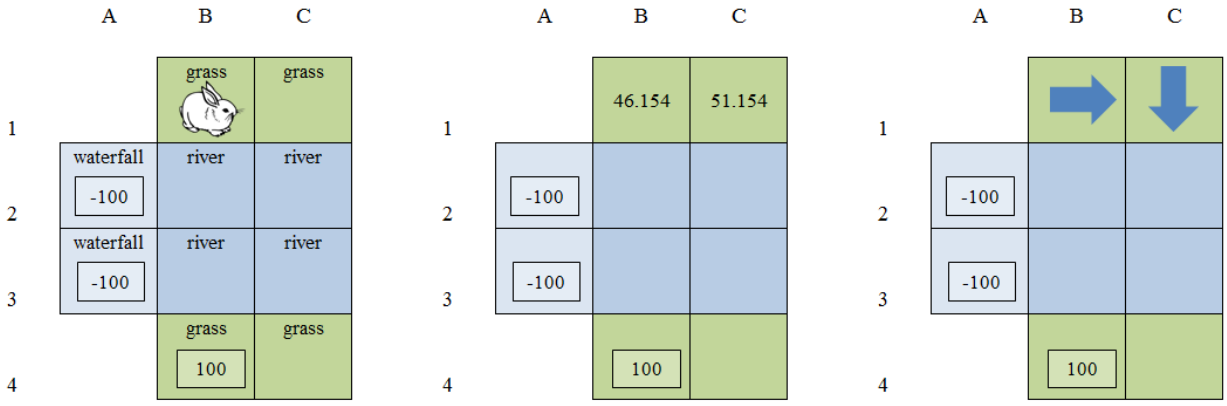

Figure 1: From left to right, a) The world, b) Partial utilities for R=-5 and c) Partial optimal policy for R=-5

Let R indicate the reward for non-terminal states, that is, R(B1) = R(B2) = R(B3) = R(C1) = R(C2) = R(C3) = R(C4) = R. By decreasing R we decrease your patience for remaining in the world. Actions available to you in each state may include North, South, East, and West (corresponding with an attempt to move in that direction). Movements that attempt to move off the board are excluded (e.g., attempting to move north is not a valid action in B1).

For each of the following subproblems, you are to provide a) a utility graph analogous to the book's Figure 17.3, b) an optimal policy graph analogous to the book's Figure 17.2, and c) a brief interpretation (1 or more sentences) of the obtained policy (To compute the utility function, the suggestion is to use the value iteration algorithm of Figure 17.4. Assume your utility is based on simple additive rewards, with the discount factor γ equal to 1.)

1. Suppose you are slightly peckish, so your R(s) = −5. Just to get you started on what your answers should look like and what sorts of values you should get, Fig 1b and 1c show an incomplete answer for this question with only the top two states' utilities and policies filled in.
2. Suppose you are really hungry, so your R(s) = −60.
3. Suppose you are totally starving, so your R(s) = −150.