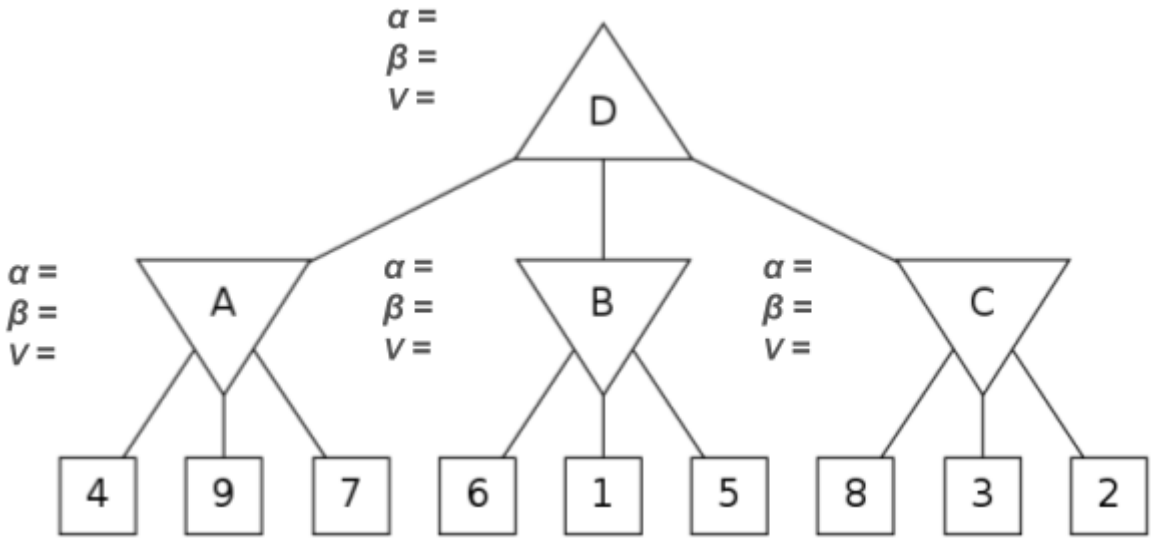
Artificial Intelligence Homework 3

1. Problem Formulation (30 pts)

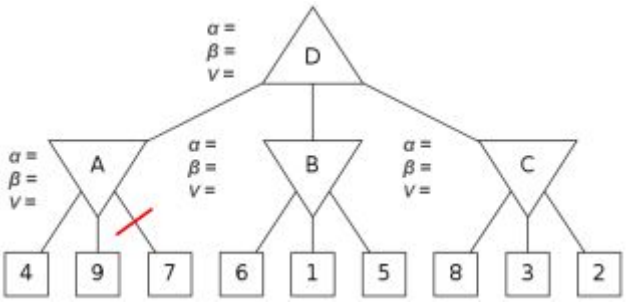
Please define the Go game's (19*19) states, actions, branching factor and transition model.
hint: follow the definition of 'brancing factor' on the textbook.

2. Alpha-Beta Pruning (30 pts)

Consider the game tree shown below. Triangles that point up, Δ, represent choices for the maximizing player; triangles that point down, ▽, represent choices for the minimizing player.



Assuming both players act optimally, use alpha-beta pruning to fill the α, β ,V on the above graph. And represents the pruning by a stroke. (see example below, the red stroke means leaf node 7 is pruned.)



3. MDP (40 pts)

Consider the mini-grid world in Figure 1, in which A is the start state and double-rectangle states are exit states. From an exit state, the only action available is *Exit*, which results in the listed reward and ends the game (by moving into a terminal state X, not shown). From non-exit states, the agent can choose either *Left* or *Right* actions, which move the agent in the corresponding direction. There are no living rewards; the only non-zero rewards come from exiting the grid.

Let's assume that value iteration begins with initial values $V_0(s) = 0$ for all states s. For now, the discount is $\gamma = 1$ and legal movement actions will always succeed (i.e. the state transition function is deterministic).

Figure 1: A Mini Grid World.

(a) (8 pts) What is the optimal value $V^*(A)$? Justfiy your answer briefly.

(b) (4 pts) What is the first iteration k for which $V_k(A)$ will be non-zero?

(c) (4 pts) What will $V_k(A)$ be when it is first non-zero?

(d) (4 pts) After how many iterations k will we have $V_k(A) = V^*(A)$?

(e) (8 pts) If $\gamma = 0.5$, what is the optimal value $V^*(A)$? Justfiy your answer briefly.

(f) (4 pts) For what range of values γ of the discount will it be optimal to go *Right* from A?

(g) (8 pts) Let's assume that the *Left* and *Right* movement actions are now stochastic and fail with probability f. When an action fails, the agent stays in place. The *Exit* action does not fail. If the failure probability is $f = 0.5$ and the discount $\gamma = 1$, what is the optimal value $V^*(A)$? Justfiy your answer briefly.

Deadline:

May 15, 2018 Hand in your assignment during the lecture.

Submission Method

Write your answer on an **A4** paper either in Chinese or English. Remember to write your department, student ID, and your name on top of the paper.