CS533
Intelligent Agents and Decision Making
Homework 3, Winter 2009

1. Prove that for any STRIPS planning problem that the length of the optimal relaxed plan from any state s to the goal g is an admissible heuristic for state s. (NOTE: This does not require a complicated proof.)

2. Consider the Sussman anomaly problem in Figure 11.16 of your book. Write a definition for this problem in STRIPS (as you did in HW1). Using this definition, compute the HSP heuristics h_{max} and h_{add} for the initial state of the problem. Show the steps of each computation.

3. Exercise 17.4 b and first question of part c (NOTES: By terminal state the book means a state with zero reward that can only transition to itself. By undiscounted MDP the book means an infinite horizon MDP where the discount factor is set equal to 1.)

4. Exercise 17.5 in book.

5. **(Policy Evaluation.)** Given a policy π, let V_π be the infinite-horizon, discounted value function (as defined in class), which we know satisfies the following equation at all states s,

$$V_\pi(s) = R(s) + \gamma \sum_{s'} \Pr(s'|s, \pi(s)) \cdot V_\pi(s') \qquad (1)$$

We can compute V_π by solving the above system of linear equations. However, there is also an iterative technique for computing V_π that is often more efficient. Consider the following value-function operator T_π,

$$T_\pi[V](s) = R(s) + \gamma \sum_{s'} \Pr(s'|s, \pi(s)) \cdot V(s')$$

note that $T_\pi[V]$ is simply a value function and $T_\pi[V](s)$ gives the value of state s. As described in your book in the discussion of modified policy iteration, this operator can be used to iteratively compute a sequence of value functions V^k that converge to V_π as follows:

$$\begin{aligned} V^0(s) &= 0, \text{for all } s \\ V^k &= T_\pi[V^{k-1}] \end{aligned}$$

Use the following steps to prove that the sequence does converge to the correct value function.

(a) Show that T_π is a contraction operator with respect to the max-norm. That is show that for any value functions V and V',

$$||T_\pi[V] - T_\pi[V']|| \leq \gamma ||V - V'||$$

(b) Use this fact to prove that $\lim_{k\to\infty} V^k = V_\pi$. You may use equation 1 if desired.

(c) Does the sequence still converge to V_π if we initialize V^0 to random values? Explain.

(d) What value of k is sufficient so that $||V^k - V_\pi|| \leq \epsilon$? Explain.