6.231 Dynamic Programming and Optimal Control
Homework Due Nov. 15, 2015
Prof. Dimitri Bertsekas

Problem 1

2.3 (Another Version of Value Iteration)

Consider n-state discounted MDP of Section 2.1 and the version of the VI method that starts with an arbitrary function $J : S \mapsto \Re$ and generates recursively $FJ, F^2J, \ldots$, where F is the mapping given by

$$(FJ)(i) = \min_{u \in U(i)} \frac{g(i,u) + \alpha \sum_{j \neq i} p_{ij}(u) J(j)}{1 - \alpha p_{ii}(u)}.$$

Show that $(F^k J)(i) \to J^*(i)$ as $k \to \infty$ and provide a rate of convergence estimate that is at least as favorable as the one for the ordinary method. Show that F is the DP mapping for an equivalent DP problem where there is 0 probability of self-transition at every state.

Problem 2

3.2

A gambler engages in a game of successive coin flipping over an infinite horizon. He wins one dollar each time heads comes up, and loses $m > 0$ dollars each time two successive tails come up (so the sequence TTTT loses $3m$ dollars). The gambler at each time period either flips a fair coin or else cheats by flipping a two-headed coin. In the latter case, however, he gets caught with probability $p > 0$ before he flips the coin, the game terminates, and the gambler keeps his earnings thus far. The gambler wishes to maximize his expected earnings.

(a) View this as an SSP problem and identify all proper and all improper policies.

(b) Identify a critical value $\overline{m}$ such that if $m > \overline{m}$, then all improper policies give an infinite cost for some initial state.

(c) Assume that $m > \overline{m}$, and show that it is then optimal to try to cheat if the last flip was tails and to play fair otherwise.

(d) Show that if $m < \overline{m}$ it is optimal to always play fair.

6.231 Dynamic Programming and Stochastic Control
Fall 2015