

MAST30001 2013, Recommended Problems, Chapter 4
Lecturer: Nathan Ross

Instructions: Answer the following questions. Justify all work and give clear, concise explanations, using prose when appropriate. Clarity, neatness and style count.

1. (Discounted Reward—see the introduction of Section 4.2 of Borovkov) Recall the setting and notation of our discussion of Markov decision theory. Many times a later reward is less valuable in some sense (e.g., the earlier you sell, the longer you could put the money in a savings account and accumulate interest). One such discounted return model sets the total reward for the process as (again recall the notation from lecture or from Chapter 4 of Borovkov)

$$\sum_{i=1}^n z^i R(X_i, a_i)$$

for some $0 < z \leq 1$ (the case $z = 1$ is the usual case with no discounting). For the example from lecture where three buyers make consecutive iid offers that you must take or decline immediately before seeing the next offer; the distribution of the offers Z_i , $i = 1, 2, 3$ is (in thousands)

$$\mathbb{P}(Z_i = 500) = 0.3, \quad \mathbb{P}(Z_i = 510) = 0.5, \quad \mathbb{P}(Z_i = 520) = .2;$$

find the strategy that optimizes the expected discounted reward and compute this expected discounted reward (these answers will depend on z). Comment on how the value of z affects the optimal strategy and its expected reward.