Notes while reading Reinforcement learning an introduction (Sutton/Barto)

Willem

January 2013

Contents

Chapter 1

Introduction

Chapter 2

Mutli-armed bandits

Chapter 3

Finite Markov Decision Process

Chapter 4

Dynamic Programming

4.1 Exercise 4.8

The reward is only obtained when the capital is above 99. When the capital is at 50, there is a 50% chance you can win the game. So this obviously is the optimal policy. When you reach 51: it would be rather odd to bet the entire capital, as you don't need to risk it all to reach 100. Bigger downside, but same upside. So the best course of action is to bet with 1, see if you can grow this above 50. If you lose it, you still have a 50% chance to win by betting it all.

Chapter 5

Mutli-armed bandits