Homework Assignment 5: Due Thursday October 12

Continuous Time Stopping

Do problem 1.11. of Bertsekas Volume II.

A brief solution suffices. Write down the Bellman operator and identify the optimal policy.

Gittins Index Theorem

Read through this paper to review the proof of the Gittins Index Theorem. Ask yourself the following questions to enforce your comprehension:

- Where did we leverage the independence structure of bandit processes?

- Lemma 2.1 might make you think the Gittins Index is somehow myopic. Why is the process of "reducing" the bandit i^* capturing the problem dynamics, discount factor, etc?

- Why did this proof require a semi-Markov formulation?

For those of you who are interested in reinforcement learning, you may want to read this landmark paper which introduced the "Options framework" for Hierarchical RL. There are natural connections to the ideas we've just studied (Semi-Markov processes, 'reducing' a bandit arm etc.).
`http://www-anw.cs.umass.edu/barto/courses/cs687/Sutton-Precup-Singh-AIJ99.pdf`