Robot Learning

Assignment 6

Due Tuesday, June 27st, before class*.*

This problem is known as the *tiger problem* and is due to Cassandra, Littman and Kaelbling (1994).

A person faces two doors. Behind one is a tiger, behind the other a reward of +400. Initially, the person has no information behind which door the tiger or the reward are.
The person can either listen or open one of the doors.
When opening the door with the tiger, the person will be eaten, which has an associated cost of –1000.
Listening costs –100. When listening, the person will hear a roaring noise that indicates the presence of a tiger, but only with 0.75 probability will the person be able to localize the noise correctly. With 0.25 probability, the noise will appear as if it came from the door hiding the reward.

6. Provide a formal model of the POMDP, in which you define the state, action, and measurement spaces, the cost function, and the associated probability functions.

    5 points

6.1) What is the expected cumulative reward/cost of the open-loop action sequence: "Listen, listen, open door 1"? Explain your calculation!

    5 points

6.2) What is the expected cumulative reward/cost of the action sequence: "Listen, then open the door for which you did not hear a noise"? Again, explain your calculation!

    5 points

6.3) Manually perform the one-step backup operation of the POMDP. Plot the resulting piecewise linear value functions in a diagram, just like the ones presented in the lecture. Provide diagrams for all intermediate steps.

    5 points