RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN
COMPUTER SCIENCE VI AUTONOMOUS INTELLIGENT SYSTEMS
Prof. Dr. Sven Behnke
Endenicher Allee 19a

Robot Learning

Assignment 2

Due Tuesday, April 23rd, before class.

2.1) Consider a student taking an exam, which consists of k tasks.
For simplicity, we assume that the tasks $i = 1, \dots, k$ can either be solved, which results in the full number r_i of points, or not be solved, resulting in zero points ($r_i = 0$).
After working on a task, the student knows whether the task has been solved or not.
The student may attempt to solve each task a second time, but only when it has not been solved before.
The probability p_i^2 of solving the task in the second attempt is only half of the probability p_i^1 in the first attempt. The first-attempt probabilities p_i^1 are as follows:

Task i	Points r_i	First attempt solution probability p_i^1
1	12	0.25
2	4	0.4
3	10	0.35
4	5	0.6
5	7	0.45
6	3	0.5
7	50	0.15

Formulate this problem as a Markov Decision Process! 4 points

2.2) Assume that the student can attempt only $N = 10$ tasks in the exam.
For passing the exam, the student needs to get at least 50% of the available points.
Model the probability of passing the exam.

4 points

2.3) The student considers two policies for choosing the tasks:
π_A: work on the tasks in sequential order, according to index i.
π_B: work on the tasks in the order of increasing difficulty
(decreasing solution probability)
In both cases, the first non-solved task will be attempted again.
Compare the expected return of both policies! 4 points

2.4) Suggest an improved policy π_C that has a higher expected return than both of the above policies.

4 points

2.5) Give an example for a process model where the Markov assumption is not justified.
How can the state be augmented to make the assumption valid again?

4 points