

Correction Notes - Assignment 2

Group: Manzil, Cüneyt and Dhananjay

Task 1

The state and action space is not defined on a per-task basis! It must encode the status of all tasks, and in the case of the actions, adapt to the available tasks. You encode the outcome of your actions into separate actions, but this approach is flawed (−1 point). The student can only attempt a task, whether this attempt fails or not must be encoded in the resulting transition from current state to the next!

Also, please hand in your assignments as a single PDF!

Points: 3/4

Task 2

The question is not on the expected return of any policies, but on the chance of passing the exam. Therefore, your approach here is, unfortunately, off-topic.

Points: 0/4

Task 3

The expected points are calculated correctly, and your approach is quite intuitive. I like this solution!

Points: 4/4

Task 4

Your π_C is clearly optimal, and you show that it is better than π_A and π_B using your script. Well done!

Points: 4/4

Task 5

Your example is of course entirely correct, I have nothing to complain about here.

Points: 4/4

Total Points: 15/20 Points