# Homework 3. Partially observable Markov decision problems



Figure 1: A "Doom" player attacked by 4 Imps and a Shotgun guy.

Consider once again the game "Doom", already featured in Homework 2 (see Fig. 1 for an example screenshot). In this homework, you will again model an extremely simplified instance of a "Doom" game as a partially observable Markov decision problem. In Homework 2, the focus of the agent was to get the necessary keys to reach the exit. In this homework, in contrast, the focus of the agent is to *avoid an enemy* to safely reach the exit.

◇

Suppose that the agent moves in the environment depicted in Fig. 2. Moving in the same environment there is an *Imp*—a monster that, if it stands in the same cell as the agent, will inflict a large amount of damage to the agent. The Imp moves between cells 4 and 5. At each step, there is a 0.3 probability of moving to the next cell, and a 0.7 probability of remaining in the same cell.

At each step, the agent may move in any of the four directions—up, down, left, and right—or, in alternative, *listen* attentively for the Imp's grunting. Movement actions (up, down, etc.) across a *grey* cell division succeed with a 0.8 probability and fail with a 0.2 probability (in
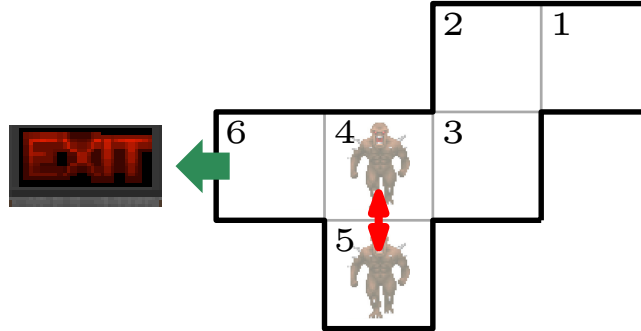
Figure 2: Environment where the agent must reach the exit (located in cell 6).

which case, the agent remains in the same cell). Assume that the action "left" always succeeds in taking the agent through the exit. The action "Listen", in contrast, keeps the position of the agent unchanged.

The agent is able to see the Imp with probability 1 if it stands in the same cell. If it stands in a cell adjacent to the Imp after executing a movement action, it is able to hear the Imp's grunting with a probability 0.3, and with a probability 0.7 it hears nothing. On the other hand, if the agent stands in in a cell adjacent to the Imp after executing a listening action, it is able to hear the Imp's grunting with a probability 0.7, but with a probability 0.3 it still hears nothing.

## Exercise 1.

(a) Identify the state space, $\mathcal{X}$, and the action space, $\mathcal{A}$, and the observation space $\mathcal{Z}$ for the POMDP. As in Homework 2, you should treat "Exit" as an absorbing state. Moreover, take into consideration that the agent *knows its own position*.

(b) Write down the transition probability matrix for the action "right".

(c) Write down the observation probability matrix for the action "listen".