Assignment 7

1. 1)states:
   x_1={door1 is reward door, door2 is tiger door}, x_2={door1 is tiger door, door2 is reward door};
   2)actions:
   u_1={open door1}, u_2={open door2}, u_3={listen};
   3)cost(reward) function:
   $V_1(b) = \max\{200p_1 - 1000(1 - p_1), -1000p_1 + 200(1 - p_1), -50\}$
   4)measurement space:
   $\{z_1, z_2\} = 0.8, 0.2$={prob correctly localize noize, prob to make mistake}

2. "Listen listen open door1"
   $V_1(b) = -50 - 50 + 200p_1 - 1000(1 - p_1)$ -we choose u_1 anyway, independently of result of u_3, thus we just add doubled expected reward of doing u_3 and expected reward after doing u_1

3. "Listen, then open the door for which you did not hear a noise" $V_1(b) = -50 + 200 * 0.8 - 1000 * 0.2 = -90$ after committing u_3 we act according to the results of measurements, thus we will open the door with best measurement with probability 0.8.