1. Why doesn't the bot avoid the radiation in the default map? What would have to be different for the bot to avoid as much of it as possible?
2. Under the default reward, the bot runs away from the enemy. What is the smallest value for enemyDead that would make it so that the bot is willing to kill the enemy if they cross paths. Explain why. What is the smallest value for enemyDead that would induce the bot to seek out the enemy and kill it. Explain why.
3. What effect does switching enemyMode from 1 (follow the influence map) to 2 during training have on the behavior of the bot, if any? How does more or less training episodes help or hurt? Hint: experiment with play = 2.

1. The bot does not avoid the radiation for several reasons. The first reason is the penalty for being in the radiation is relatively low. The bot only get penalizes by being in the radiation, rather then being penalizes each turn after it has been touched by the radiation. However, the bot would get 20 points each round, as soon as it rescues a human. Thus, the penalty for being inside radiation is almost nothing comparing to the reward it would get after successfully rescues a human. Furthermore, the bot would get 10 points by staying in the base station. Thus once the bot reach the base station, the bot would not make any further movement if humans were already rescues. To have the bot avoid the radiation as much as possible, the bot should get penalize each turn after it get touches by the radiation, and this penalty should be slightly higher than the reward it would receive each turn from rescuing a human.
2. If the enemy is in the area where a human is in, the bot will choose to smash the agent regardless of the value of the enemyDead. This is because the robot would gain a huge reward every step after rescuing a human, thus making the penalty of enemyDead insignificant. For the bot to seek out the enemy and kill it, enemyDead must be a value that is positive. As long as enemyDead contains a negative penalty, there is no benefit for the bot to kill the agent under situation that the enemy is not blocking the agent to get a higher reward, such as rescuing a human. Thus the smallest value for enemyDead that would induce the bot to seek out the agent and kill it would be 1. The smallest value for enemyDead that would make the bot to kill the enemy when cross path would be 0. Under such situation, the agent doesn't gain reward from killing the bot, but it doesn't receive a penalty neither.
3. When the bot is in random mode, it attends a better reward score at a faster speed. Perhaps by having the enemy moving randomly, the optimal solution appears sooner. While using the influence map, the enemy would take unnecessary steps when training. Furthermore, there would be space that enemy would not have go proceeding some steps if it use influence map. Having more training episodes would not result in a better score, since the max reward score have been coverage early on around 100 iterations. However, having less training episodes before the reward score converges to the max reward score would have result in a poorly trained bot.