Ferran Cantariño i Iglesias - 173705
Marc Rabat Pla - 172808

Informe - Pràctica 2

Estat de la pràctica

- **Què funciona?**
  All the questions required to solve the asked questions are done. We have not implemented the optional question Q8.

- **Què no funciona?**

- **Quins problemes hem trobat?**
  ValueIterationAgent and QLearning Agent have been the more complex parts if we speak in terms of reaching the solution. Also, the parameters in questions 2, 3 & 6 are difficult to understand precisely how they affect to the agent directly, so we reach the values after test-fail iteration for different values.

Respostes

Is there an epsilon and a learning rate for which it is highly likely (greater than 99%) that the optimal policy will be learned after 50 iterations? question6() in analysis.py should return EITHER a 2-item tuple of (epsilon, learning rate) OR the string 'NOT POSSIBLE' if there is none. Epsilon is controlled by -e, learning rate by -l.

We have a probability of (2/4)^10 of finding the 10 ending states of the bridge, this probability done in 50 iterations augments to a probability of 0.04882% of finding the 10 states after 50 iterations. This probability is too low to depend on the epsilon or learning rate parameters, so the problem is NOT POSSIBLE.