Sebastien Gaspoz
Marin Thiercelin

Assignment 3 :Report

Part I : Generalisation

To compute the state representation, we extracted several features from the screen array.
We first used an iteration on the screen that tried to locate the pacman in the screen .
Once the pacman was located, we iterated on the pixels around it (a 60*60 pixel square) to look for
the closest ghost and we represented the state as a tuple (distance,direction). The distance being the
manhattan distance of the closest ghost and direction being one of the North, East, West, South.
This gave a state-action space of size 60*4*4 = 960. This representation, even if it isn't really
powerfull, gave a good generalisation. The agent was able to learn that it was not rewarding to go in
the direction of the ghosts, but, as the walls were ignored, was not able to avoid it really often.

The second approach was to remove the background from the screen , and to take into account the
objects (coins, ghosts, cherries) in a perimeter around the pacman giving 12 features per screen and
per action. We considered this a stronger representation and that could lead to better results, but it
lead to a bigger state-action space and made it harder to generalize. As before, we ignored the walls
and that limited the improvement of the agent since it tended to get stuck in a corner.

Part II : Exploration

To enforce exploration we used a function that
given a state , select the least frequently used
action in that state, with a probability 1/T. Where
T is a function of the frequency of the state : T =
10 if state frequency n < 10 times, T = 20 if n <
50, T=50 if n < 100, T = n if n > 100.
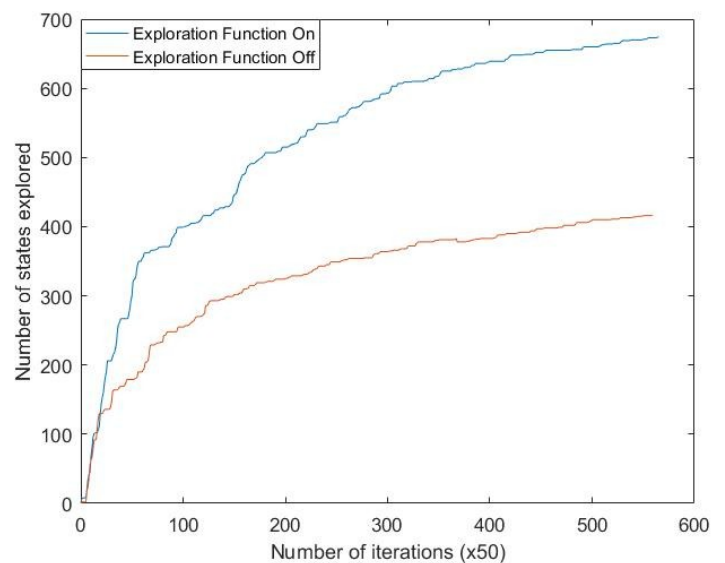
Part III : Performance

We let our agent learn on more than 2000 games and
we noticed a slight improvement in the score.
The agent learned to go away from the ghosts which
in turn lead to a better score, but didn't take into
account the walls, and usually ended up being trapped
in a corner. Also the agent didn't learn anything about
the situation where the ghosts are far away and this
limited the improvement.