

# Mario

Zidar Miha (63060317)

January 20, 2013

## 1 Introduction

In this report we will present the results of Q learning for Super Mario from 12th rl competition 2009.

## 2 Methods

For learning we used a simple Q matrix, where states were represented as strings surrounding Mario. To make monsters more important, we clipped the string to just have Mario and the nearest monster. For each step we propagated the reward back with the function:

$$Q[s][a] = Q[s][a] + \frac{reward}{1 + offsetfromcurrentaction/10} \quad (1)$$

This proved to be much better than the original formula:

$$Q[s][a] = Q[s][a] + \alpha * (reward + \gamma * Q[s'][a'] - Q[s][a]) \quad (2)$$

## 3 Results

Here we have the result of the previously mentioned algorithm, and control runs.

- random - selects all actions randomly.
- random forward - selects all actions randomly, but without the option to go back.
- simple Q learning - Q learning with equation 2
- inc Q learning 1 - same as above but with equation 1 and with all movements set to run, thus reducing the action space from 12 actions to possible 6 actions.
- inc Q learning 2 - same as Q learning 1, but with full movement set, and added option of exploration once a good method has been found.

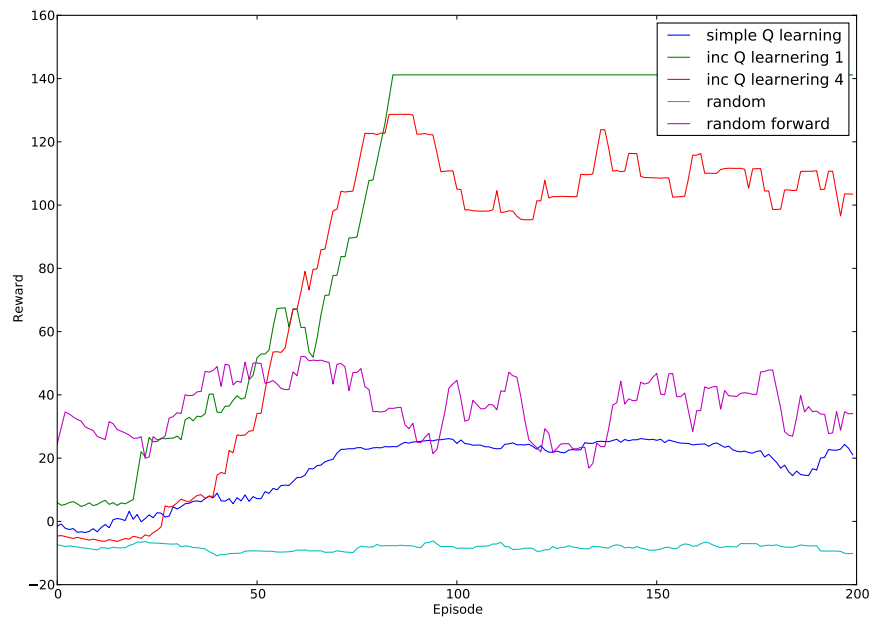


Figure 1: Result of learning with 200 trials