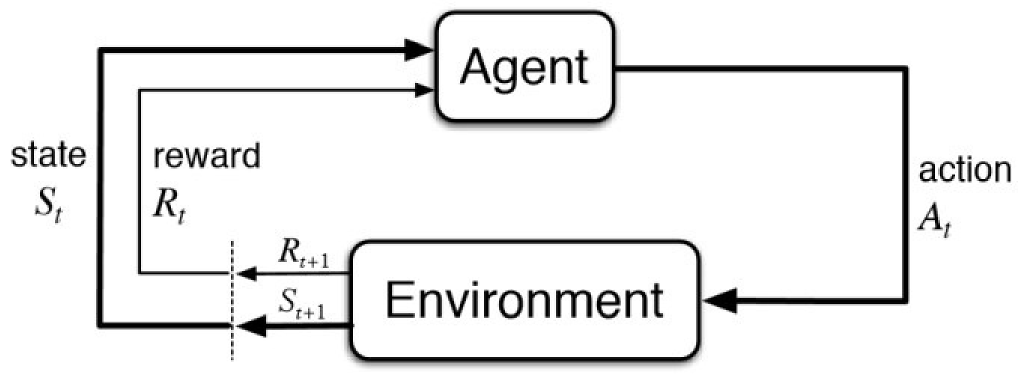
QUANTUM REINFORCEMENT LEARNING

Cristhian Gomez, Miguel Angel Ortiz, Sebastian Medina



REINFORCEMENT LEARNING

Area of Artificial Intelligence focused on building intelligent agents that can take actions that benefit them in a certain environment, maximizing their reward. RL uses the reward to evaluate actions and learn from current or past states to future actions by interacting with the environment initially by trial and error.

QUANTUM REINFORCEMENT LEARNING

QRL has similar components as RL such as a policy, a reward function and an environment. QRL algorithms are different from classical RL algorithms in the way that states and actions are represented (orthogonal bases in the Hilbert space) and how the policy is updated (Grover algorithm).

Algorithm 1: QRL Algorithm

Initialize
$|s^m\rangle = \sum_s^{11\ldots1} C_s |s\rangle$, $f(s) = |a_s^n\rangle = \sum_a^{11\ldots1} C_a |a\rangle$
and $V(s)$ arbitrarily.

Repeat(for each episode)

For all states $|s\rangle$ in $|s^m\rangle = \sum_s^{11\ldots1} C_s |s\rangle$:

1) Observe $f(s) = |a_s^n\rangle$ and get $|a\rangle$
2) Take action $|a\rangle$, observe next state $|s'\rangle$, reward r, then:
   a) Update state value:
   $V(s) \leftarrow V(s) + \alpha(r + \gamma V(s') - V(s))$
   b) Update probability amplitudes:
   Repeat U_{Grov} for L times:
   $U_{Grov} |a_s^n\rangle = U_{a_0^n} U_a |a_s^n\rangle$

Until for all states $|\Delta V(s)| \leq \varepsilon$

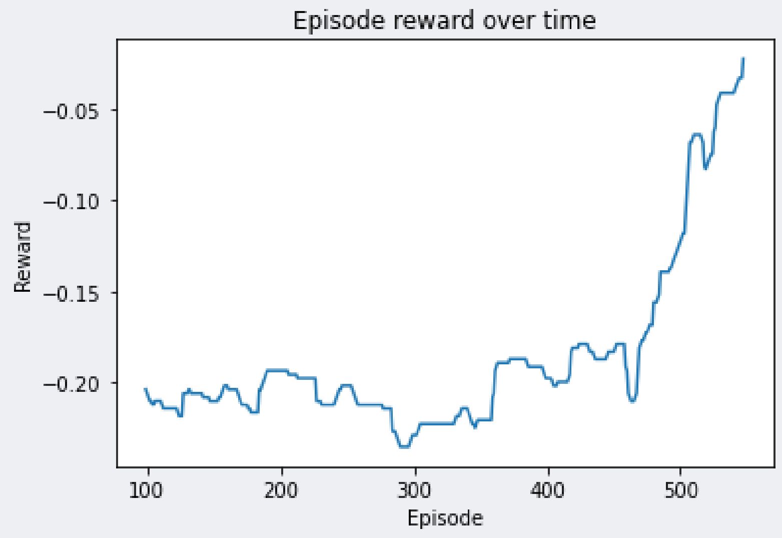
CONCLUSION

Quantum computation can provide powerful tools to improve current algorithms and novel approaches to solving difficult computational tasks. The implementation and results obtained in this work reflect the natural benefit of involving quantum theory in computational problems, specially related to Artificial Intelligence. The results show the feasibility of using quantum computation in a practical way and its superiority when problems lie in highly dimensional spaces that may be sub-optimal to solve in a classic computational approach. Video available at: https://youtu.be/jdp4hEVhB4M

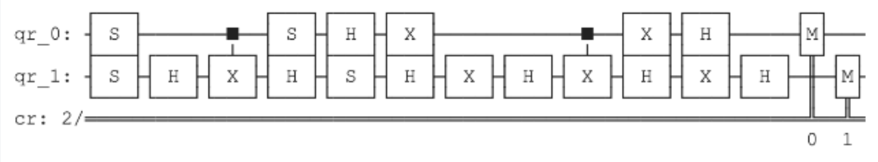
IMPLEMENTATION AND RESULTS

The QRL algorithm was tested with the classical problem of path searching inside a maze, in which it was compared the Quantum algorithm against the classical Q table based on neural networks. The actions on the QRL algorithm were encoded using two bits, and the Grover algorithm was implemented for four different cases.

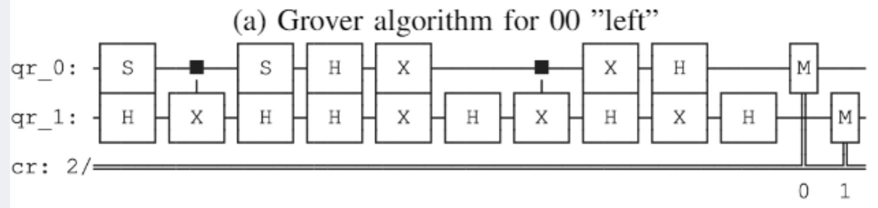
QRL REWARD FUNCTION



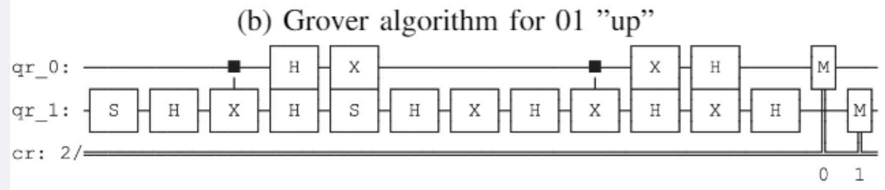
RL REWARD FUNCTION



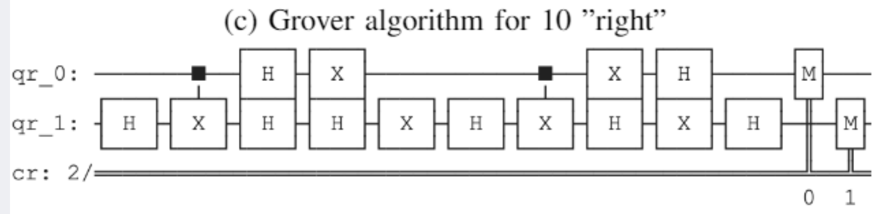
(a) Grover algorithm for 00 "left"



(b) Grover algorithm for 01 "up"



(c) Grover algorithm for 10 "right"



(d) Grover algorithm for 11 "down"

GROVER ITERATIONS