# Q Learning

Q learning is an unsupervised learning. Normally it is a typical use of reinforcement learning. The algorithm is based on Markov Decision Process, MDP. An event consists of several states and actions. Most significantly, the next state is only relevant to the current state, and has no relation to the historical states.

## Q table

The key idea for Q learning is Q table, which rows and columns represent states and actions. The value Q(S, A) of Q table is a evaluation of current state S takes the action A.

## Bellman Equation

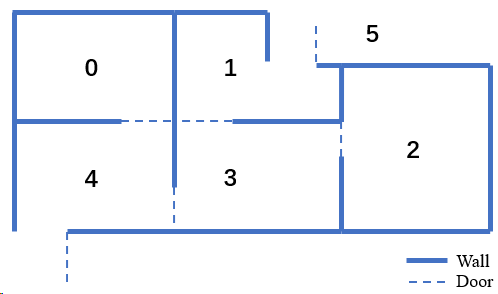
During the training process, we use Bellman Equation to update the Q table.

Q(S, A) = R(S, A) + γ\*max(Q(S’, A’))

In the upper equation, R(S, A) is the reward to state S with action A. γ is wreck rate. Q(S’, A’) is the Q value of next state S’ with all related action A’.

## Implementation

This script will implement on the route search by a robot. Let’s imagine a big house, which has five rooms, number from 0 to 4. And number 5 is outside area. What’s more, from room 1 and room 4 can go outside directly. The schematic is shown as below.



Now the target is to make a robot find a shortest route to go outside from a random room.