
Algorithm 1 Multiagent learning algorithm

Require: env, algorithm, episodes, output, steps, eval, ▷ hyper parameters
load agents and environments
for $agent \in agent\ set$ **do**
 load the algorithm ▷ all the agents use the same algorithm
 train the agent and evaluate every 100 iterations
end for

Algorithm 2 training steps

Require: episodes, output, steps, eval, ▷ basically kepted the default value.
Initialize a neural network has the value estimator of $Q(s,a)$. ▷ The neural network structure are 50
neuros and 2 hidden layers for all the algorithms to bes fare
for 10000 *episode* **do**
 for 100 *episode* **do**
 while step < step limit **do**
 $s(t), a(t) \leftarrow DNN$
 $a(t+1) \leftarrow DNN$
 $env(s(t), a(t+1))$
 record(s, a, r), (s, a, r), ...
 end while
 end for
 train the DNN with record
 evalute the DNN:
 for *steps* **do**
 $s(t), a(t) \leftarrow env(s(0), a(0))$
 $a(t+1) \leftarrow DNN(s(t), a(t))$
 end for
end for

pesudocode

Zhaoyang Chen

May 2023

1 Introduction