Algorithm 1 Deep Q-Learning for H2O-Cloud With Experience Replay

Algorithm 2 Deep Q-Learning for H2O-Cloud With Experience Replay

Initialize environment, including clusters/servers/VMs
Initialize the dqnAgent, including replay buffer/dqn network
Reading datasets to get jobID/dcpu/dmem to create tasks
for $episode = 1 \ldots, E$ **do**
    Run $env.reset()$ to initialize the cloud platform
    for $task = 1 \ldots, T$ **do**
        Get current state s_t from environment
        Choose action a_t(cluster) based on s_t
        Run $env.step(task, action)$ to get new state s_{t+1}
        reward r_t, reject signal and options(valid actions)
        while $reject = 1$ and $options > 1$ **do**
            $round\,robin()$ to get new action a_{t+1}
            if new action != action **then**
                action = new action
                Run $env.step(task, action)$ to get new state s_{t+1}
                reward r_t, reject signal and options(valid actions)
            end if
        end while
        Store transition (s_{t+1}, a_t, r_t, s_t) in memory
        Store decision a_t and reject signal
        Update state to new state $s_{t+1} = s_t$
        Sampling mini-batch of transitions to learn by
        replaying experience by dqnAgent
        Every Y steps decrease ϵ
        Every ζ steps, copy Q to $\hat{Q}$
    end for
end for
return All actions and reject signals