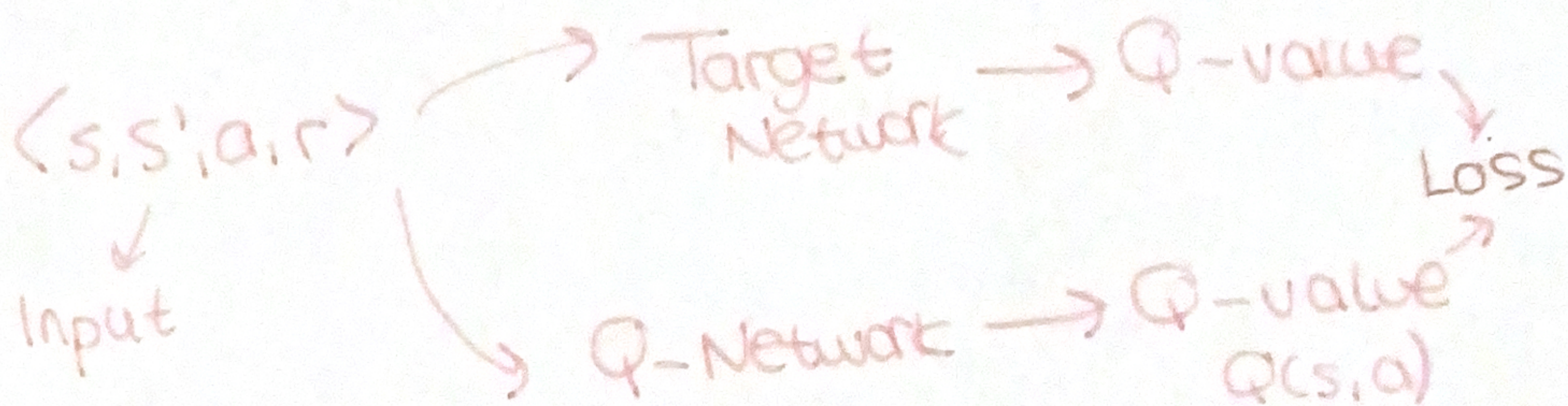
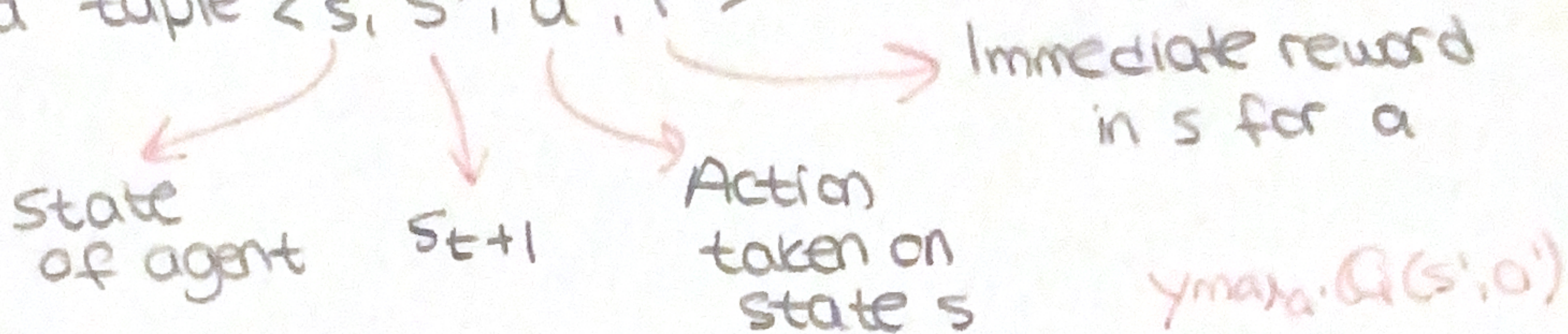


Experience Replay

Neural network training process becomes more stable when training is done on random batch of previous experiences

Experience Replay: Memory that stores experiences

as a tuple $\langle s, s', a, r \rangle$



$$\text{Loss} = (r + \gamma \max_a Q(s', a) - Q(s, a))^2 \quad \leftarrow \text{Backprop}$$

Deep Q-Learning Steps

1. Provide state of environment to agent. Get Q-values from Target Net & Q-Net.
2. Pick action a , based on epsilon value.
3. Perform action a .
4. Observe reward & next state s' . Put them in $\langle s, s', a, r \rangle$
5. Sample random batches from experience replay memory & perform training on Q-Net.
6. Each n^{th} iteration, copy weights from Q to Target.
7. Repeat 2-7 for each episode.

Episode \rightarrow all states between initial & terminal states.

Epoch \rightarrow Forward & backward pass