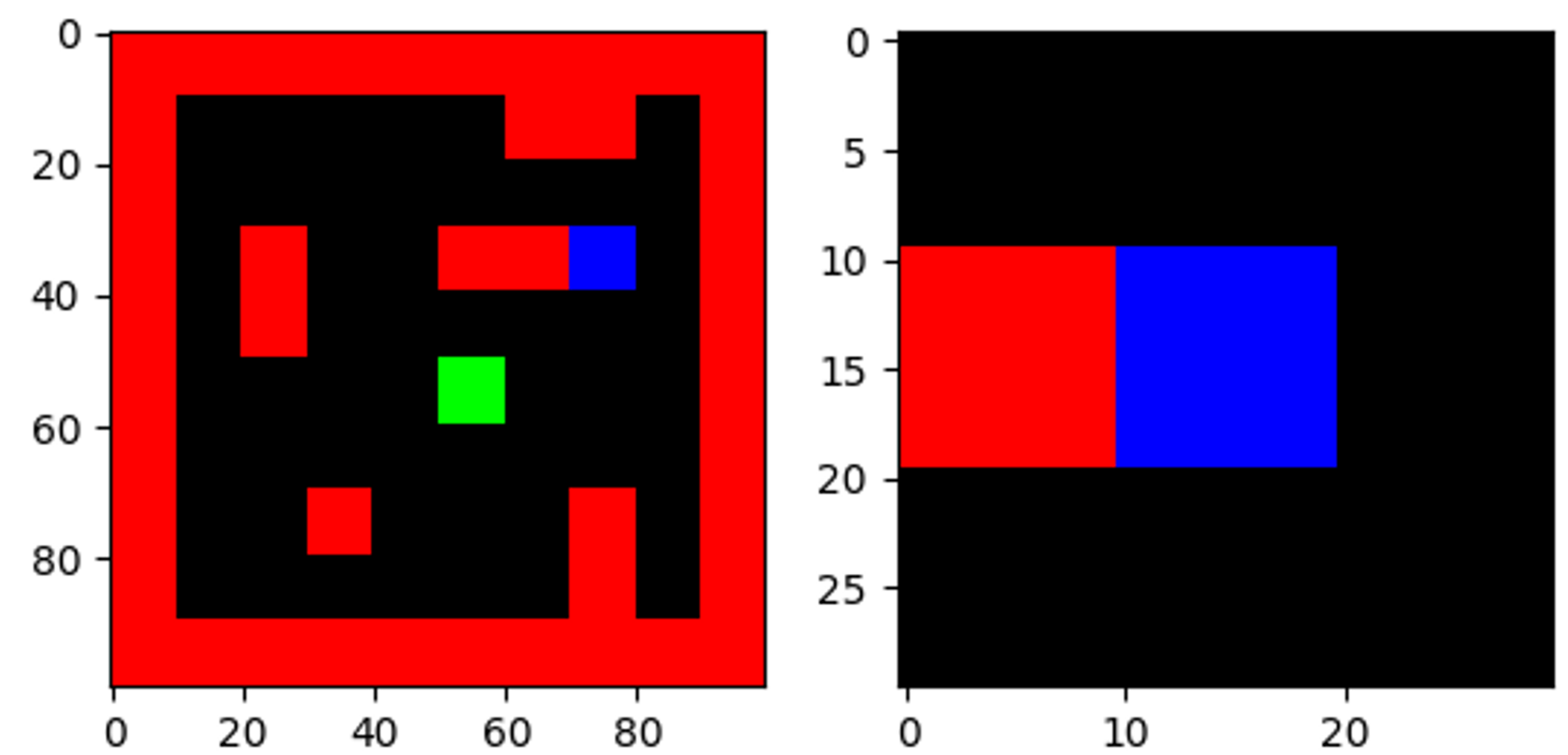


# Deep Q-learning variants

Archit Basal, Ilia Dobrusin

## Deep Q-learning (DQN)

- Use Deep Neural Network to approximate Q-function
- We show three **extensions to regular Q-learning**
- **Visual Planning task**: find route to goal in a grid maze map with small local view
- Train for 100.000 steps with random sampling from previous experience



## Target Network

- Simple extensions of regular DQN
- **Use second (target) network** to calculate target Q-values and next action (for loss calculation)
- **Keep target network static**, only update source network weights
- Synchronize source and target network periodically

## Double DQN

- Regular DQN overestimates Q-values
- Similar to target network
- **Use source network for action prediction**
- Calculate **target Q-value with target network**

$$Q_{target} = r + \gamma Q(s', \argmax(Q(s', a, \theta), \theta'))$$

## Duelling Networks

- Q-values indicate how good each action is for given state
- Split Q-value calculation into two:
  - **Value function**  $V(s)$ : indicates quality of current state
  - **Advantage function**  $A(a)$ : Comparison of actions compared to each other

$$Q(s, a) = V(s) + A(a)$$

- **Duelling approach**: Let network compute separate values for V and A by splitting network internally after convolutions

