Follow your Nose: Using General Value Functions for Directed Exploration in Reinforcement Learning

Durgesh Kalwar[1], Omkar Shelke[1], Somjit Nath[1], Hardik Meisheri[1], Harshad Khadilkar[1,2]

[1]TCS Research, Mumbai

[2]IIT Bombay

scan me

Contribution

- We extend upon temporally extended version of the ϵ-greedy exploration strategy by using auxiliary task learning with the help of General Value Functions (GVF) to perform directed exploration thereby further improving state space coverage during exploration.
- This is generalized formulation to include domain knowledge about the environment by providing GVF cumulant which also improves latent representation.

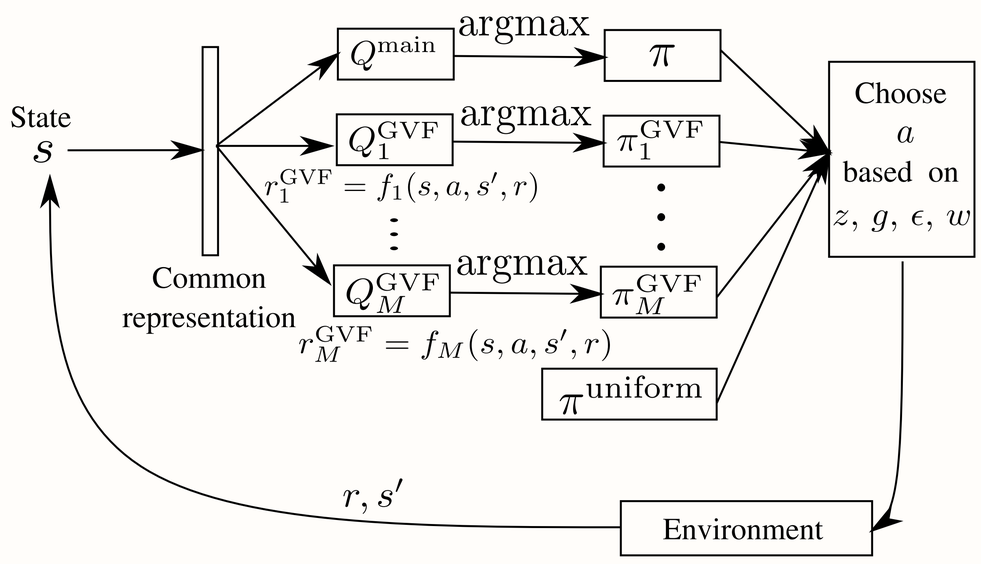


Figure 1: High-level architecture of the DEZ-greedy strategy.

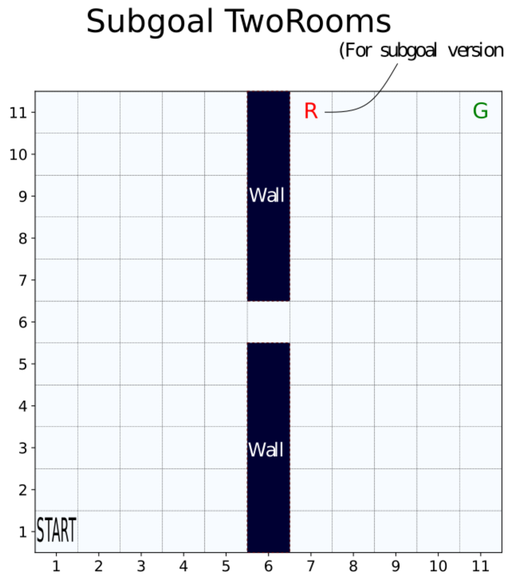
Pseudo-code

Function *DEZGreedy*(ϵ, Z_{max})**:**

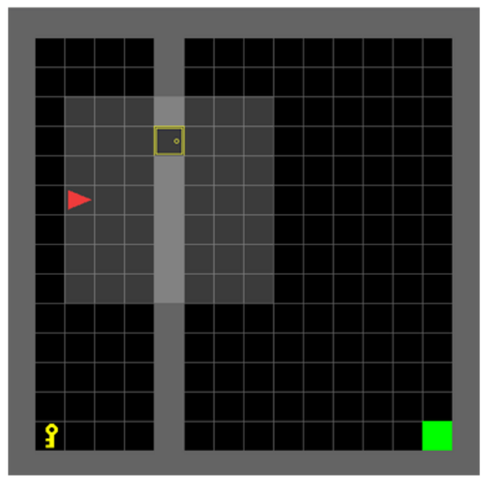
Countdown timer $z \leftarrow 0$
Uniformly sampled random action $w \leftarrow -1$
Selected GVF index $g \leftarrow 0$
while *True* **do**
  Observe State s
  if $Z == 0$ **then**
    if $\text{random}() < \epsilon$ **then**
      Sample Duration: $z \sim [1, Z_{max}]$ // Explore
      Sample GVF: $g \sim [0, M]$
      if $g == 0$ **then**
        Sample action: $w \leftarrow U(A)$
        $a \leftarrow w$
      $a \leftarrow \text{argmax}(Q_g^{GVF})$
    else
      $a \leftarrow \text{argmax}(Q^{Main})$ // Exploit
  else
    if $g == 0$ **then**
      $a \leftarrow w$
    else
      $a \leftarrow \text{argmax}(Q_g^{GVF})$
    $z \leftarrow z - 1$
  Take action a

Environments

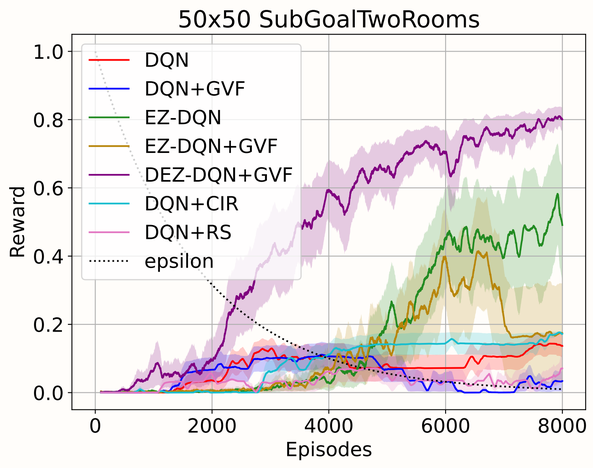
Subgoal TwoRooms



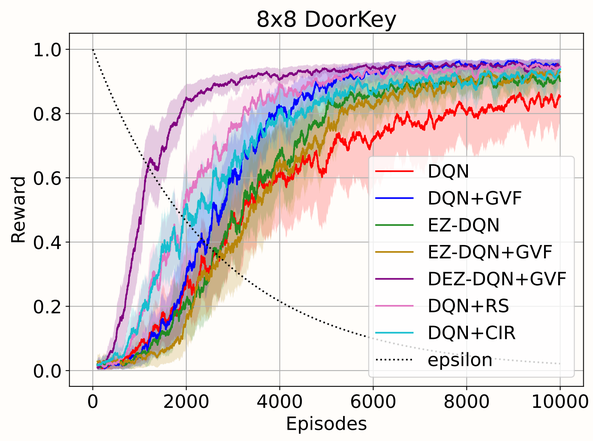
MiniGrid DoorKey



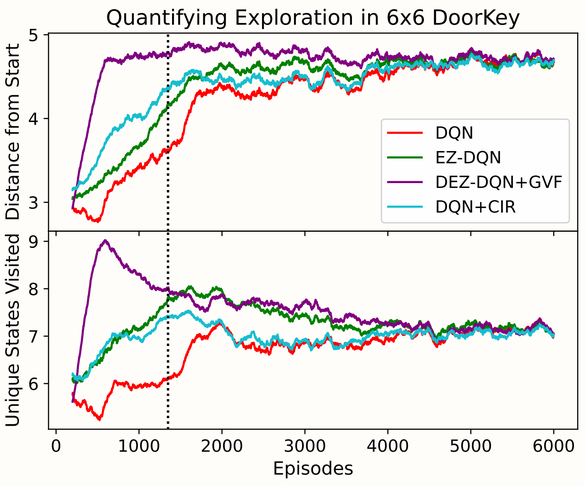
Results & Discussion



- There is a marked performance difference between algorithms that uses persistence and normal ϵ-greedy.



- DEZ- greedy policy (Purple) converges early compared to other baselines.



- Directional exploration helps DEZ to not only explore farther from the initial state but also reach unique states quickly.

References

[1] Will Dabney et al, Temporally-Extended ϵ-Greedy Exploration, ICLR 2021, *https://openreview.net/forum?id=ONBPHFZ7zG4*

AAMAS, 29 May – 02 June 2023, London