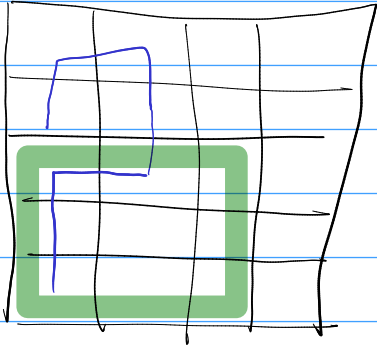


15 March
2023

Lecture 02 → Exploration

$$Q(s, a) = r + \gamma \max_{a'} Q(s', a')$$



Key difficulties

① Sparse goals (rewards)

② Lack of models

③ Uneven exploration of state space

Possible fix: Directed exploration — frontier states — curiosity
— novelty — options —

■ $Q(s, a, \beta) = r_t^e + \beta r_t^i$