

Обучение с подкреплением

13 ноября 2016 г.

1. Сформулируйте задачу о многоруком бандите.
2. В чем заключается метод UCSB?
3. Сформулируйте общую задачу обучения с подкреплением.
4. Дайте определение ценности состояния и действия в состоянии (формулы).
5. Опишите, как происходит шаг в алгоритме Q-learning.
6. Как происходит обновление $V(s)$ в $TD(\lambda)$?
7. Какова основная проблема при применении стандартных алгоритмов в реальной жизни?