

Вопросы к докладу «Multi objective reinforcement learning survey».

Климкин Андрей

14 марта 2017 г.

1. В чем состоит главное отличие задачи MORL от задачи RL. Почему мы не можем обойтись только методами для обычного RL (привести пример, где MORL может быть полезнее чем обычный RL)?
2. Что такое функция скаляризации (с примером), как она может помочь для решения задачи MORL? Какие основные недостатки использования функции скаляризации для решения MORL?
3. Какие два основных подхода есть для решения задач MORL? Приведите примеры алгоритмов каждого из подходов.
4. Что такое Consistently tracking policy? Как данная проблема решается в Pareto Q-learning, и почему она не возникает в задачах обычного RL (в случае, когда награда является скаляром)?
5. Что такое Hypervolume indicator (с формулой)? Для чего он нужен, каким важным свойством он обладает?
6. В чем смысл Stochastic mixture policy подхода? Какую проблему помогает решить данный подход? Каким главным недостатком он обладает?