

Научные сотрудники из лаборатории моделирования, Samsung Research, Samsung Electronics, задались целью предложить новую структуру обучения с подкреплением AP-NPQL (Non-Parametric Q Learning with Action Primitives) и оценить ее эффективность, применительно к манипуляторам, получающих информацию об окружающем мире через камеры.

Задачи манипулирования, на которых проводилось тестирование алгоритма: толкание тарелки, укладка коробок, переворачивание чашки, захват и размещение тарелки – актуальные задачи, например, для робота, занимающегося работой по дому.

Обучение с подкреплением актуально в данном случае, потому что на основе визуальных данных робот должен определить, какой объект нужно взять, где его схватить, где и как поставить.

Однако у RL есть недостатки, например, низкая выборка и ошибочные «оптимальные» стратегии (последние вызваны попаданием в локальные оптимумы при использовании метода градиентного спуска). Все это усложняется высокой размерностью визуального ввода.

В качестве усовершенствования алгоритма обучения с подкреплением предлагается упрощение параметрической части алгоритма и введения набора примитивов действий.