# **REINFORCEMENT LEARNING LIBRARY**

## Этапы обучения с подкреплением

Реализация обучения с подкреплением выполняется в несколько этапов:

1. инициализация система, которую надо обучить
2. применяется алгоритм обучения с подкреплением
3. анализ полученных результатов
4. оптимизация алгоритма обучения с подкреплением (при необходимости)

Данные этапы выполняются не зависимо от того, какой алгоритм обучения с подкреплением будет применён. Но при этом способ реализации каждого из этапов может быть выбран индивидуально, в зависимости от условий и сложности решаемой задачи.

В библиотеке предлагается реализация этапов 1 и 2, написанная на языке программирования Java. Реализация этапа 1 представляет собой набор объектов и методов, который позволяет инициализировать обучаемую систему. На этапе 2 есть возможность выбора алгоритмов обучения SARSA или Q-Learning. При этом, доступны методы, реализующие основные шаги указанных алгоритмов. Таким образом пользователь получает возможность самостоятельно модифицировать алгоритм обучения в пределах TD-обучения. Это даёт возможность повышать качество обучения, а также не зацикливаться на самостоятельной реализации шагов, которые будут однотипными для алгоритмов TD-обучения. Также некоторые методы пользователь может переопределять. Это предаёт гибкость коду и удобство использования, так как чётко установленных правил инициализации системы обучения или реализации алгоритмов обучения нету.

Этапы 3,4 должны быть реализованы непосредственно пользователем путём анализа полученных в результате этапа 2 данных, а также изменений параметров обучения в целях улучшения качества обучения. Результатами для анализа могут быть, к примеру, скорость обучения, эпизод, на котором был достигнут успех в обучении и другие параметры, которые пользователь выберет или определит самостоятельно.

Стоит отметить, что этапы 1 и 2 зависят от поставленной задачи обучения с подкреплением. Поэтому пользователь должен чётко определить цель обучения, а также что конкретно и чему надо обучить. Только после этого можно переходить к этапу 1. Таким образом можно выделить начальный, или подготовительный, этап для пользователя – постановка задачи обучения. Без корректной постановки задачи переход к реализации обучения с подкреплением является бессмысленным.

Далее будет описана более детально непосредственно структура этапов 1, 2 и её реализация в RL библиотеке. Также будут приведены некоторые примеры реализации этапов 3,4.

## 1 Этап инициализации обучаемой системы

### **Структура этапа инициализации обучаемой системы**

Этап определения обучаемой системы состоит из следующих шагов:

1. инициализация агента обучения, то есть того, кого надо обучить
2. инициализация состояний, в которых может пребывать агент в процессе обучения
3. инициализация действий, которые может выполнять агент в процессе обучения

Результатом данного этапа будет чёткое определение того, что мы хотим обучить и чему.

4) выбрать алгоритм, в соответствии с которым будет проводится обучение

5) определить все необходимые параметры в зависимости от выбранного алгоритма

6) описать процесс обучения

7) проанализировать полученные результаты

2) выбирается алгоритм для реализации обучения с подкреплением

3) определяются параметры обучения для применения выбранного алгоритма обучения

4) применение выбранного алгоритма с определёнными ранее параметрами

Agent:

State: