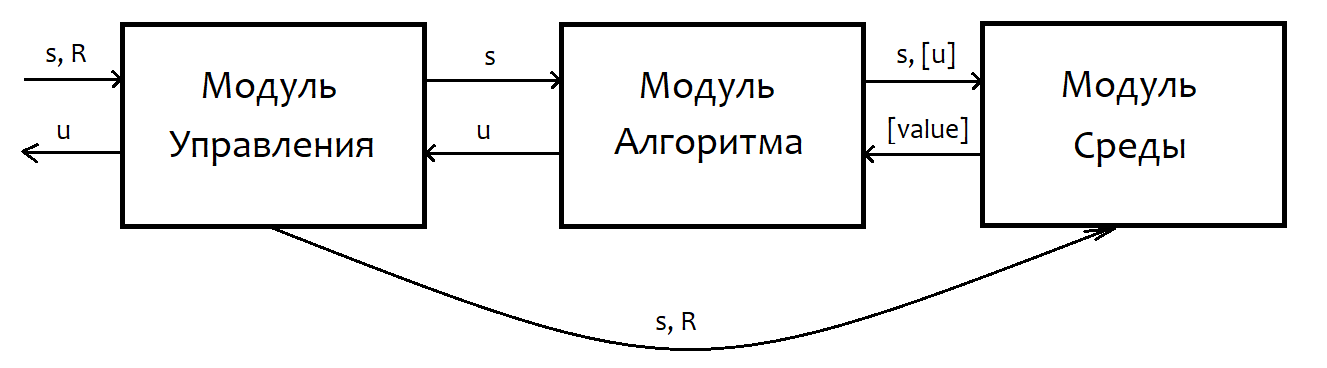
**Описание HLD собственной модели**

В рамках данного проекта предложена следующая структура архитектуры:



**Блок «Модуль управления»** принимает на вход текущее состояние (s) и сохраняет его в текущую траекторию. С последним состоянием траектории блок принимает итоговую «ценность» этой траектории (R) и отправляет эту траекторию в «Модуль среды». В задачи блока также входит отправить в «Модуль алгоритма» текущее состояние и получить от него управление (u), которое он вернёт назад клиенту вызова.

**Блок «Модуль алгоритма»** принимает на вход текущее состояние (s) и запрашивает у блока «Модуль среды» предсказания по текущему состоянию с каждым из возможных действий (u). Из полученных ответов выбирается наиболее выгодный в плане действия, и это действие возвращается блоку «Модуль управления».

**Блок «Модуль среды»** содержит в себе методы предсказания на базе RDF (Random Decision Forest) и на базе таблицы. RDF изначально обучается на таблице случайных запусков. В ходе работы алгоритма блок «Модуль управления» посылает набор траекторий (s, R), по мере накопления определённого количества которых данные добавляются в таблицу, и RDF переобучается. Метод предсказания на базе RDF по набору состояний (s) и действий (u) возвращает блоку «Модуль алгоритма» набор предсказанных значений награды (value). Метод предсказания по таблице по набору состояний (s) ищет в таблице ближайшие по Евклидову расстоянию состояния и выдаёт действие (u) при максимальном значении награды.