

Обучение агентов для выработки  
неконфликтных действий, направленных на  
достижение конкретных целей, с помощью  
кооперативного взаимодействия

Выполнил: Удод М.Н., гр. 8303  
Руководитель: Филатов Ан. Ю., к.т.н.

# Актуальность

Наблюдается рост решений, использующих машинное обучение и предназначенных для взаимодействия с человеком. При этом:

- Существующие решения в основном считают человека экспертом
- На этапе обучения ИИ делает предположения о поведении напарников. Человек может действовать иначе
- Поведение каждого человека индивидуально
- Использование человека на этапе обучения дорого

Цель: Разработать инструмент для обучения агентов выработки неконфликтных действий в условиях взаимодействия с человеком

# Предложенный алгоритм

1. Получение начального приближения с помощью алгоритмов глубокого многоагентного обучения с подкреплением
2. Запись человеческих траекторий
3. Обучение человекоподобного агента
4. Дообучение в среде с человекоподобным агентом



# Реализация

Язык программирования: Python

1. Для многоагентного обучения используется библиотека MARLlib
2. Траектории сохраняются в формате Minari. Для записи может использоваться CrowdPlay.
3. Для имитационного обучения используется библиотека Imitation

Также реализованы:

- Оптимизация гиперпараметров (Ray Tune и Optuna)
- Настройка через YAML-файл
- Поддержка пользовательских сред и человеческих политик
- Просмотр результатов после шагов 1, 3 и 4

# Заключение

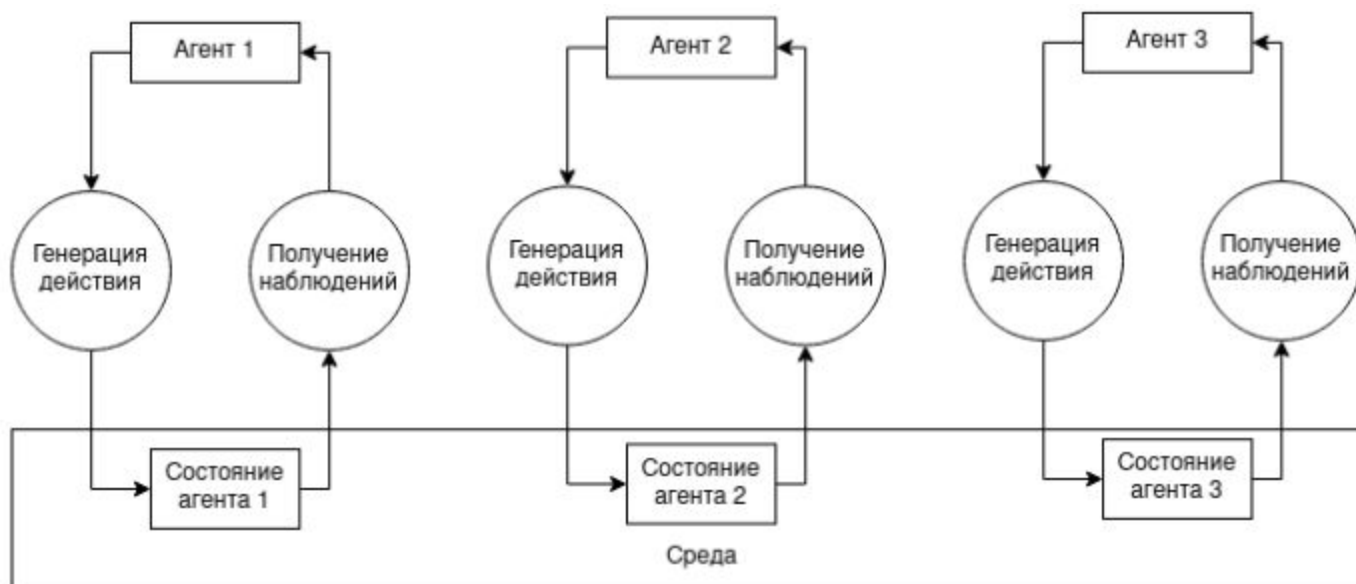
Дальнейшие задачи:

1. Валидация результатов работы инструмента
2. Измерение производительности (изменения вознаграждения за эпизод после дообучения)
3. Публикация в PyPi

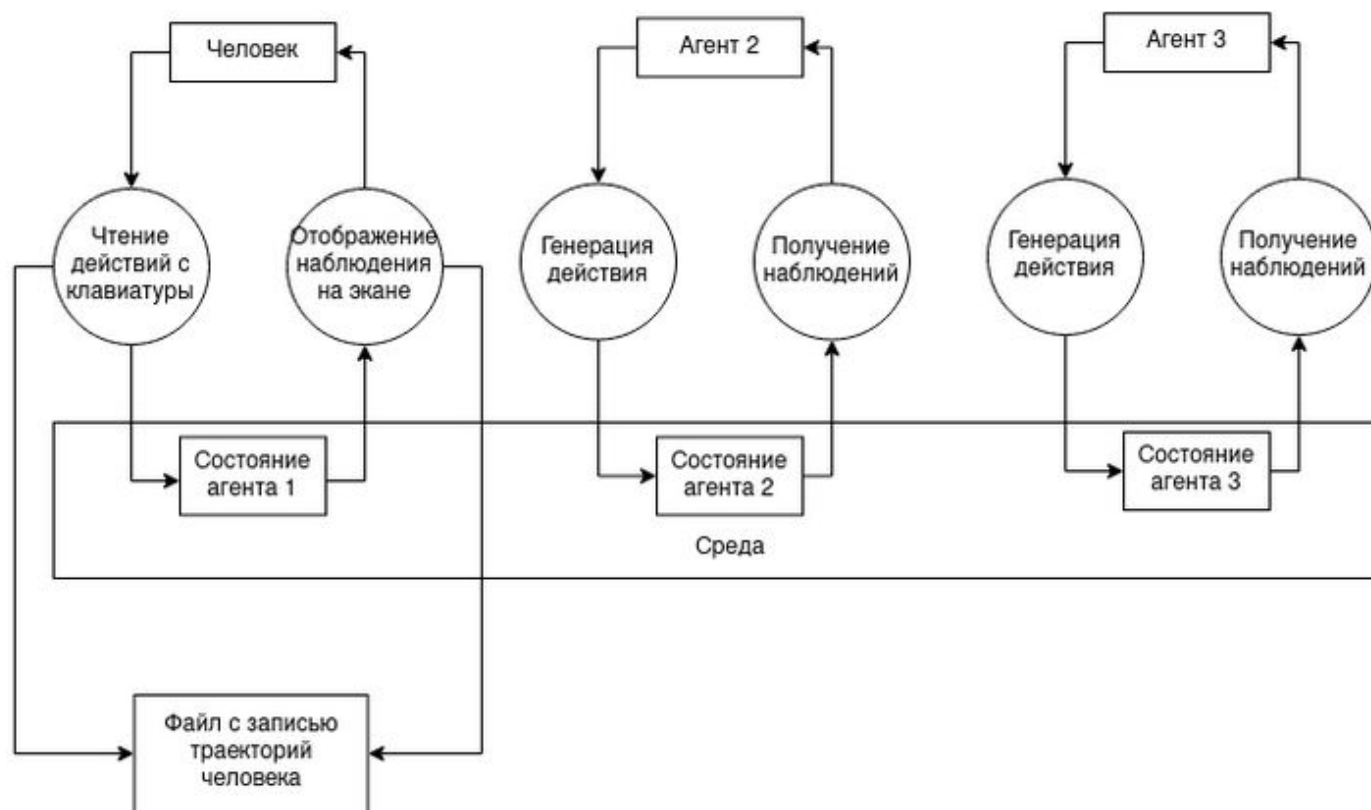


Дополнительные слайды

# Шаг 1



## Шаг 2

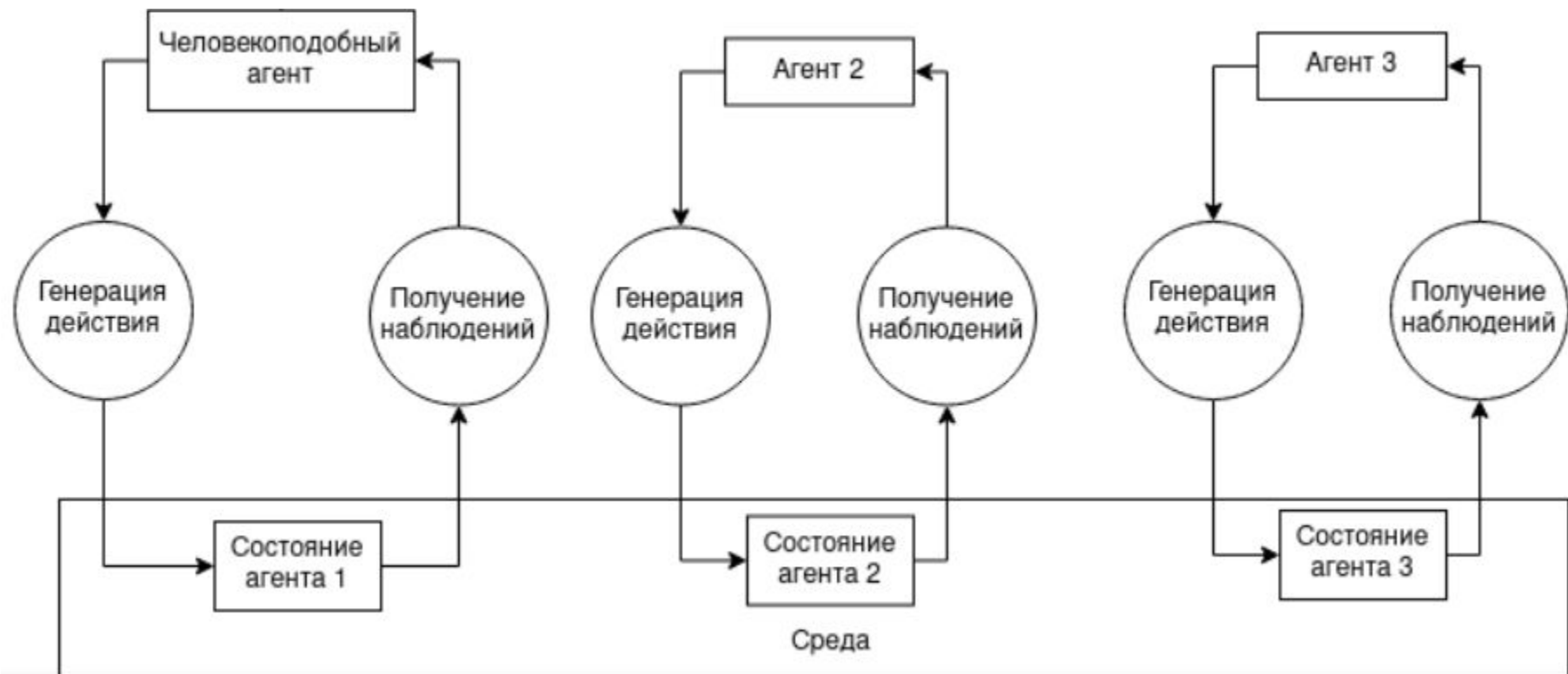




## Шаг 3



## Шаг 4



# Сравнение библиотек для глубокого многоагентного обучения с подкреплением

Библиотека	Число поддерживаемых сред	Число алгоритмов	Разделение параметров	Виды сетей
<a href="#">PyMARL</a>	1 cooperative	5	share	GRU
<a href="#">PyMARL2</a>	2 cooperative	11	share	MLP + GRU
<a href="#">MAPPO Benchmark</a>	4 cooperative	1	share + separate	MLP + GRU
<a href="#">MAlib</a>	4 self-play	10	share + group + separate	MLP + LSTM
<a href="#">EPyMARL</a>	4 cooperative	9	share + separate	GRU
<a href="#">MARLlib</a>	15 no task mode restriction	18	share + group + separate + customizable	MLP + CNN + GRU + LSTM

# Сравнение библиотек глубокого имитационного обучения

Библиотека	Число алгоритмов	Виды сетей
<a href="#">Inverse Reinforcement Learning</a>	3	MLP
<a href="#">Tianshou</a>	3	MLP
<a href="#">Let's do Inverse RL</a>	4	MLP
<a href="#">irl-imitation</a>	3	MLP
<a href="#">DI-engine</a>	8	MLP + CNN + GRU + LSTM
<a href="#">Imitation</a>	8	MLP + CNN