

Использование обучения с подкреплением в задаче автоматического тестирования мобильных приложений

Фомин Сергей Александрович

Научный руководитель:

Турдаков Денис Юрьевич

Научный консультант:

Сорокин Константин Сергеевич

ИСП РАН

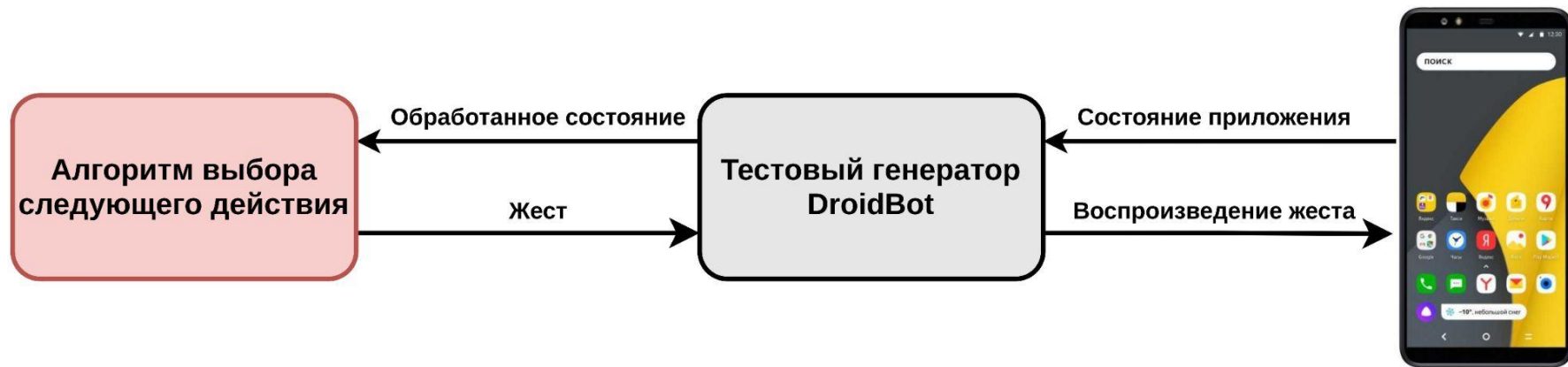
28.04.2021

Тестирование мобильных приложений

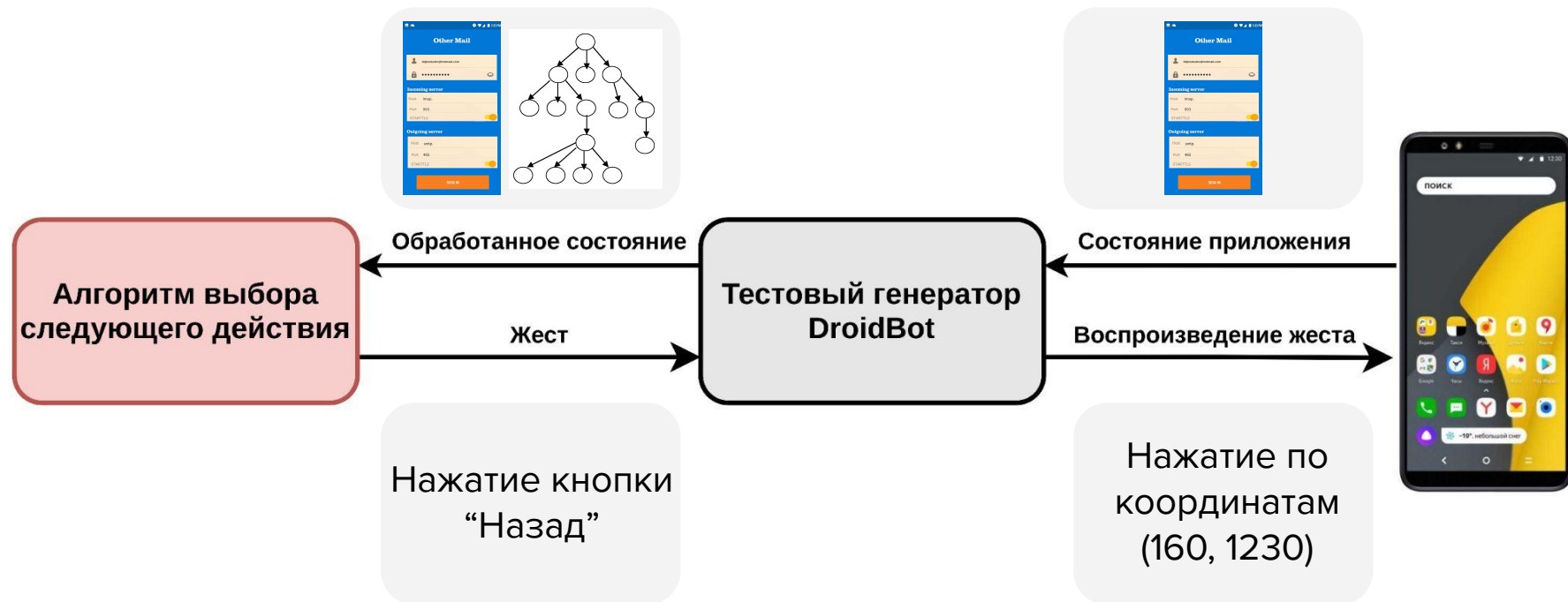
- Цель: проверка готового приложения на наличие ошибок и сбоев
- Взаимодействие с пользовательским интерфейсом UI жестами: нажатие, долгое нажатие, вертикальная и горизонтальная прокрутка



Автоматическое тестирование



Автоматическое тестирование: Пример



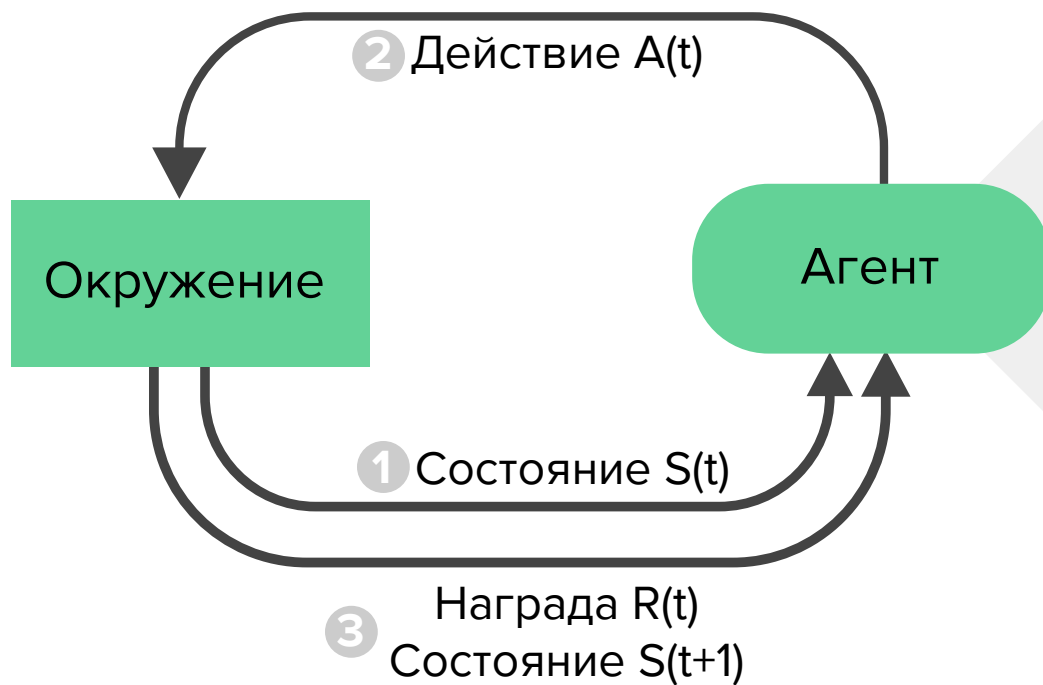
Актуальность

- Востребованность мобильной разработки
- Промышленная востребованность: ручное тестирование - дорого, долго и ненадежно
- Отсутствие оптимального инструмента:
 - Dynodroid: An Input Generation System for Android Apps (2013)
 - Automated Test Input Generation for Android (2015)
 - QLearning-Based Exploration of Android Applications (2018)
 - A Deep Learning based Approach to Automated Android App Testing (Humanoid 2019)
 - Deep Reinforcement Learning for Functional Software Testing (2020)

Постановка задачи

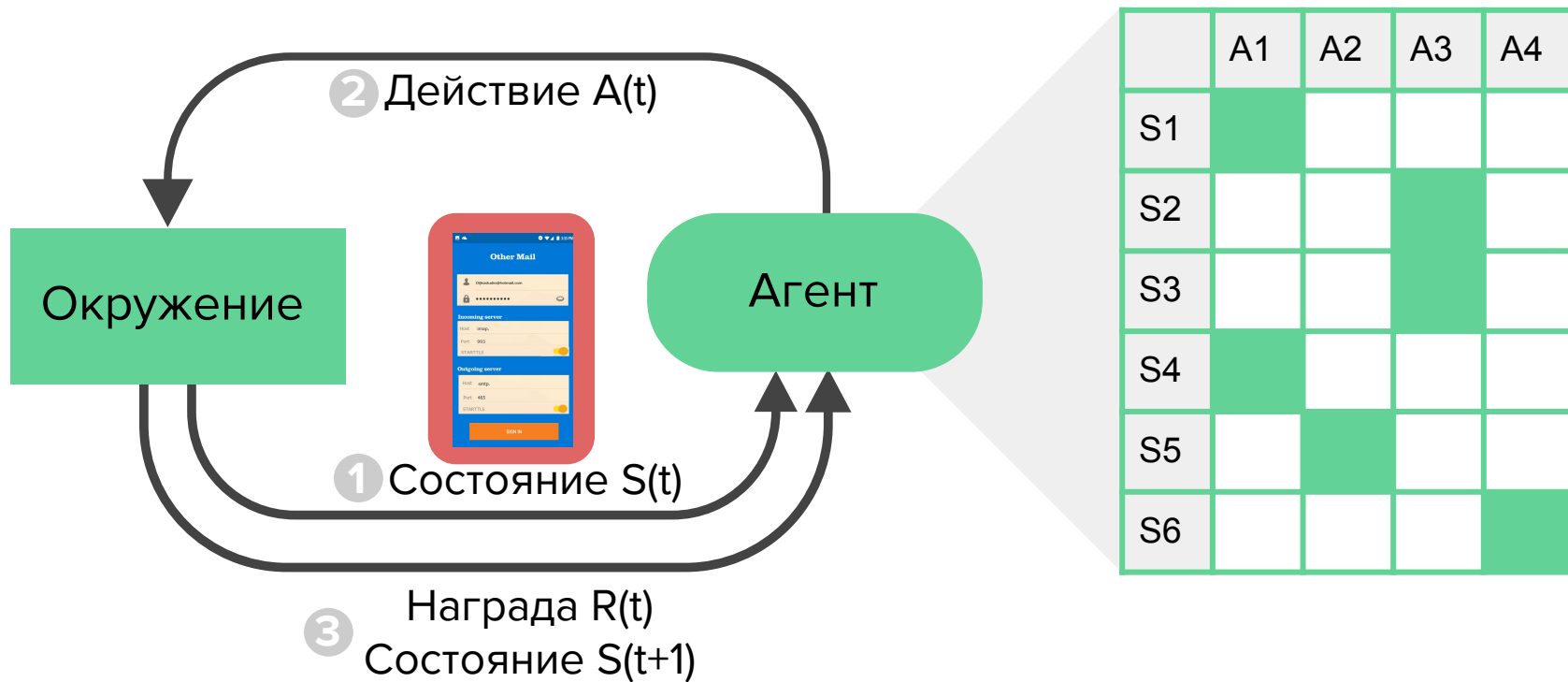
- Исследовать применимость методов обучения с подкреплением в задаче тестирования мобильных приложений через взаимодействие с графическим интерфейсом
- Реализовать и внедрить алгоритмы обучения с подкреплением в систему тестирования DroidBot + Humanoid
- Сравнить разные стратегии и функции наград алгоритмов обучения с подкреплением используя метрику уникальных состояний
- Сравнить лучшую стратегию обучения с подкреплением с инструментом Humanoid

Обучение с подкреплением

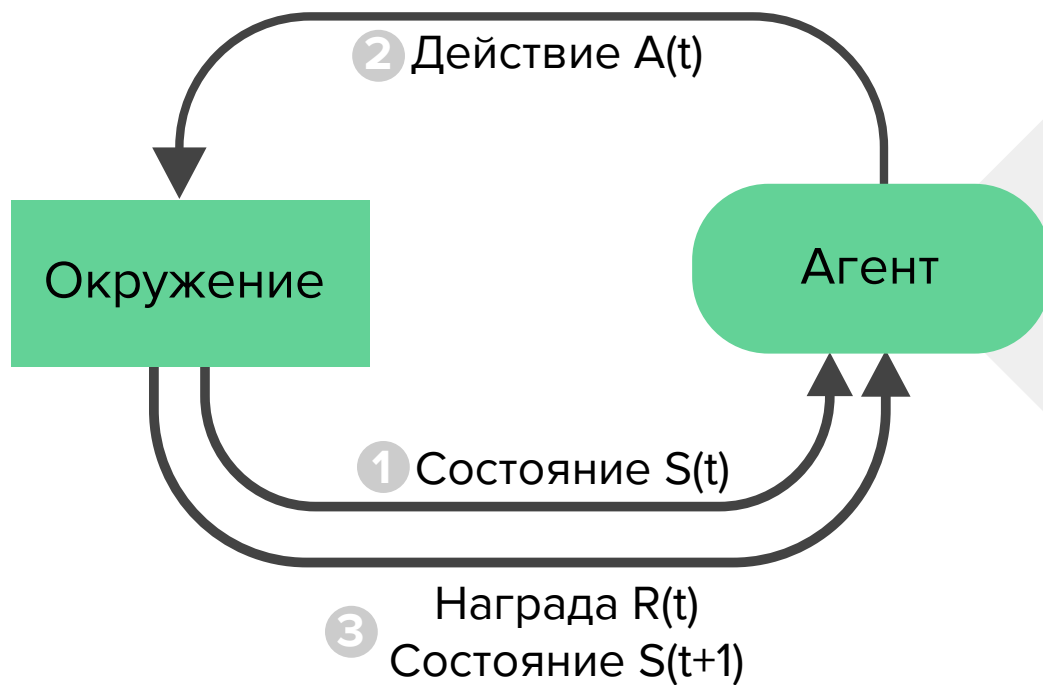


	A1	A2	A3	A4
S1				
S2				
S3				
S4				
S5				
S6				

Обучение с подкреплением: Пример

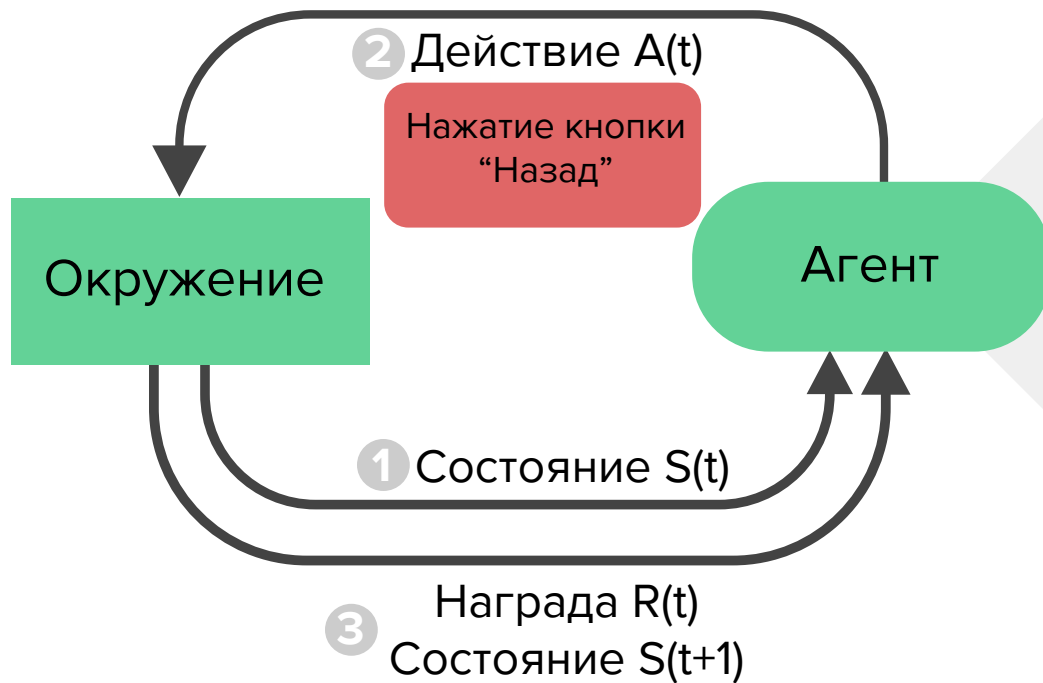


Обучение с подкреплением: Пример



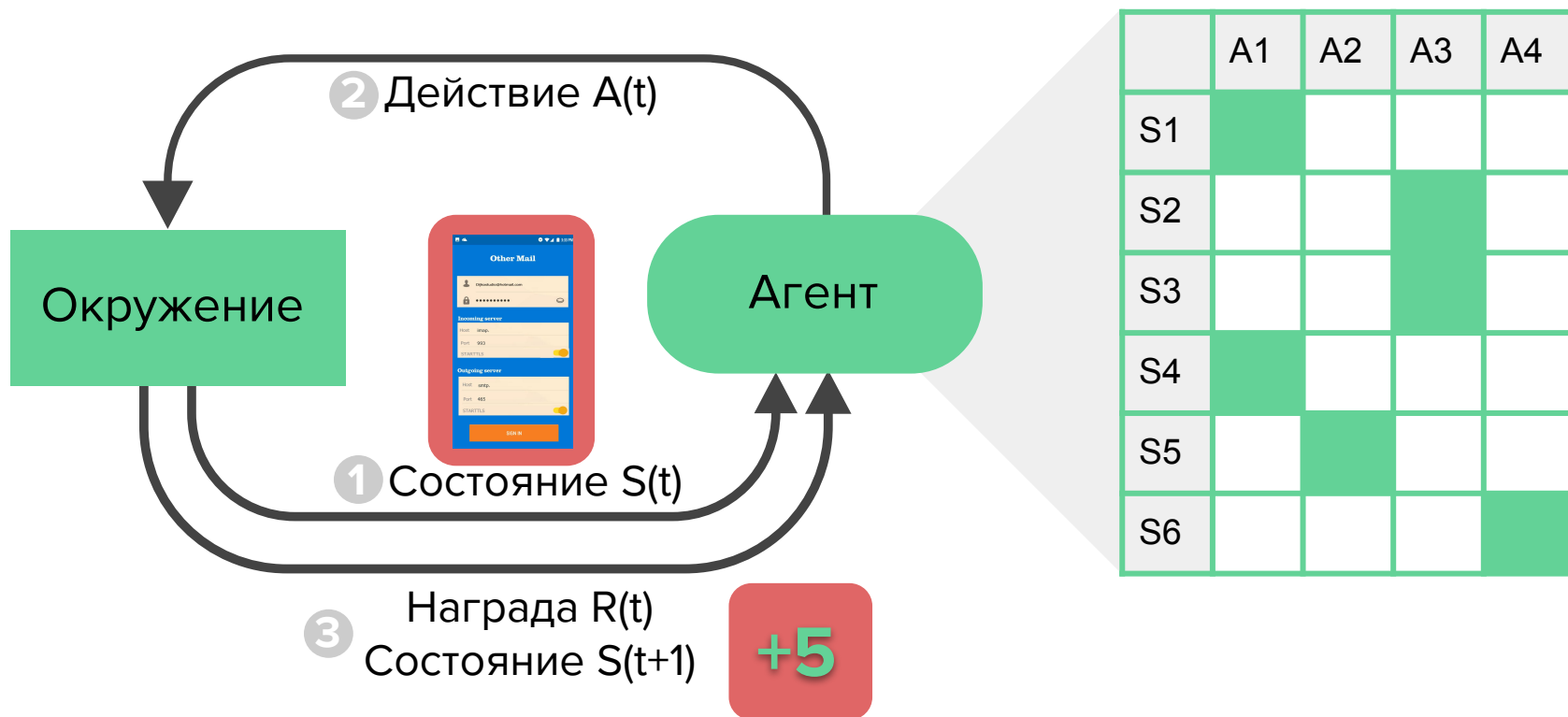
	A1	A2	A3	A4
S1				
S2				
S3				
S4				
S5				
S6				

Обучение с подкреплением: Пример



	A1	A2	A3	A4
S1				
S2				
S3				
S4				
S5				
S6				

Обучение с подкреплением: Пример



Независимые от приложения стратегии

- QBE: QLearning-Based Exploration of Android Applications (2018)
 - Абстрактное состояние - количество интерактивных элементов на экране
 - Абстрактное действие
- Сверточная нейронная сеть
 - Состояние - изображение
 - Действие - бинарная маска места нажатия на экране

Зависимые от приложения стратегии

- Базовая версия алгоритма
 - Состояние - список интерактивных элементов в текущий момент времени
 - Действие - взаимодействие с устройством, доступное в текущем состоянии
 - Награда: Обратная частота нажатий

$$\frac{1}{count(state[t], event[t])}$$

Зависимые от приложения стратегии

- Награда: Количество интерактивных элементов в новом состоянии

$$\frac{|events|}{count(state[t], event[t])}$$

- Награда: Обратное число к количеству интерактивных элементов

$$\frac{1}{|events| count(state[t], event[t])}$$

- Награда: Расстояние между деревьями состояний (Tree Edit Distance)

$$\frac{dist(state[t], state[t + 1])}{count(state[t], event[t])}$$

Улучшение стратегий

- Эпсилон жадная стратегия: Добавление случайности в выбор действия на первых этапах обучения
- Предобучение Q-таблицы: Обучение таблицы в течение 4 эпизодов взаимодействия по 12 минут + тестирование в течении одного эпизода
- Предобучение Q-таблицы с эпсилон жадной стратегией: В этапы обучения добавляется случайность, которая уменьшается к последнему эпизоду

Стратегия	applebees	booking	ebay	faceapp	nytimes	wsj
Абстрактные состояния	30	74	41	36	34	36
Сверточная нейронная сеть	27	51	24	37	27	53
Обратная частота нажатий	32	116	67	54	77	96
Количество интерактивных элементов	30	85	45	48	45	117
Обратное количество интерактивных элементов	34	85	40	48	49	85
Расстояние между состояниями	32	86	46	39	74	82
Эпсилон Жадная	35	94	60	63	68	84
Предобучение	68	107	40	79	53	117
Предобучение + эпсилон жадная	63	97	34	59	60	134

Метрика: уникальные состояния

Стратегия	applebees	booking	ebay	faceapp	nytimes	wsj
Абстрактные состояния	30	74	41	36	34	36
Сверточная нейронная сеть	27	51	24	37	27	53
Обратная частота нажатий	32	116	67	54	77	96
Количество интерактивных элементов	30	85	45	48	45	117
Обратное количество интерактивных элементов	34	85	40	48	49	85
Расстояние между состояниями	32	86	46	39	74	82
Эпсилон Жадная	35	94	60	63	68	84
Предобучение	68	107	40	79	53	117
Предобучение + эпсилон жадная	63	97	34	59	60	134

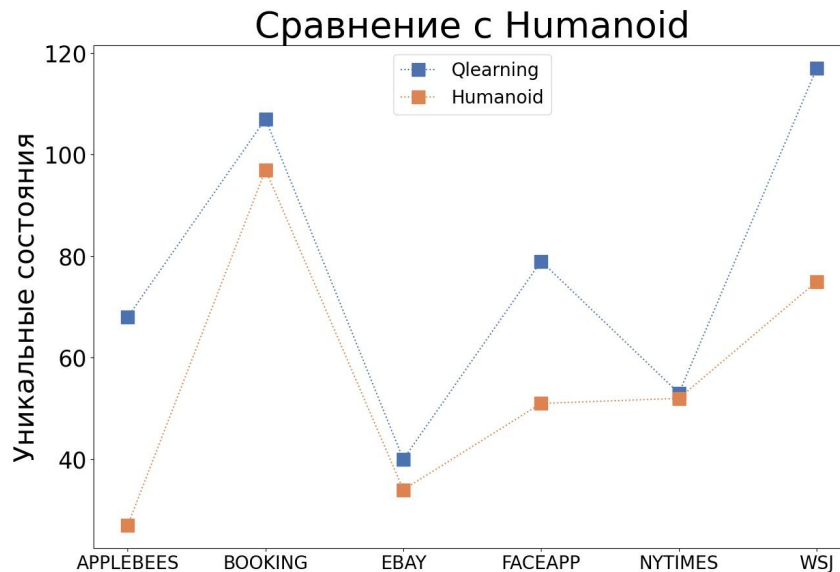
Метрика: уникальные состояния

Стратегия	applebees	booking	ebay	faceapp	nytimes	wsj
Абстрактные состояния	30	74	41	36	34	36
Сверточная нейронная сеть	27	51	24	37	27	53
Обратная частота нажатий	32	116	67	54	77	96
Количество интерактивных элементов	30	85	45	48	45	117
Обратное количество интерактивных элементов	34	85	40	48	49	85
Расстояние между состояниями	32	86	46	39	74	82
Эпсилон Жадная	35	94	60	63	68	84
Предобучение	68	107	40	79	53	117
Предобучение + эпсилон жадная	63	97	34	59	60	134

Метрика: уникальные состояния

Сравнение с Humanoid

- Humanoid: современный инструмент тестирования на основе глубоких нейронных сетей



Результаты и дальнейшая работа

- Изучены подходы обучения с подкреплением в задаче тестирования мобильных приложения
- Реализованы и внедрены несколько стратегий обучения с подкреплением в инструмент DroidBot
- Реализованный Q-learning подход с предобучением Q-таблицы превзошел современный инструмент тестирования Humanoid
- В перспективе можно рассмотреть влияние альтернативного представления состояний приложения (в том числе эмбединги)