

Создание бота для командной тактико-стратегической игры Dota 2

Кулеш Сергей Павлович, 441 группа

Научный руководитель: ст. преп. Зеленчук И. В.

Рецензент: магистр Кокорев А. Д.

СПбГУ, 2019 г.

Dota 2 и киберспорт

- 1,29 млн максимальный онлайн
- Ежедневный онлайн — 600 000
- > \$25 млн призовой фонд The International 2018
- > \$100 млн суммарный призовой фонд
- Где тренироваться?

Цель

Создание бота с элементами искусственного интеллекта, способного выигрывать у стандартного бота высокой сложности на центральной линии в режиме игры "1 на 1", на котором можно оттачивать отдельные элементы игры.

Постановка задачи

- Изучить предметную область и методы взаимодействия с игровым процессом
- Написать бота без использования машинного обучения, исследовать результат
- Развить модель бота путем внедрения машинного обучения в отдельные области игрового процесса

Defense of the Ancients

- Существует с 2005 года
- Непрерывная симметричная карта
- 2 игрока
- Задача — дважды убить противника или уничтожить башню



API

- С 2016 года находится в открытом доступе
- Список скриптовых функций и переменных на Lua
- Ограниченное время реакции ботов и возможность действовать лишь в рамках правил игры



Существующие решения

- OpenAI
 - 60 000 CPU
 - 300 лет обучения в день
 - 3.3 — размер одного наблюдения
 - > 167 млн наблюдений в минуту
 - Весь код в закрытом доступе
- Боты из мастерской
 - LUA
 - Мастерская — база плагинов
- Стандартные боты
 - C++
 - Код в закрытом доступе

Бот без машинного обучения

- ≈ 2000 строк кода
- Эффективнее стандартного бота высокой сложности, а также всех ботов из мастерской
- Играет на уровне среднестатистического игрока

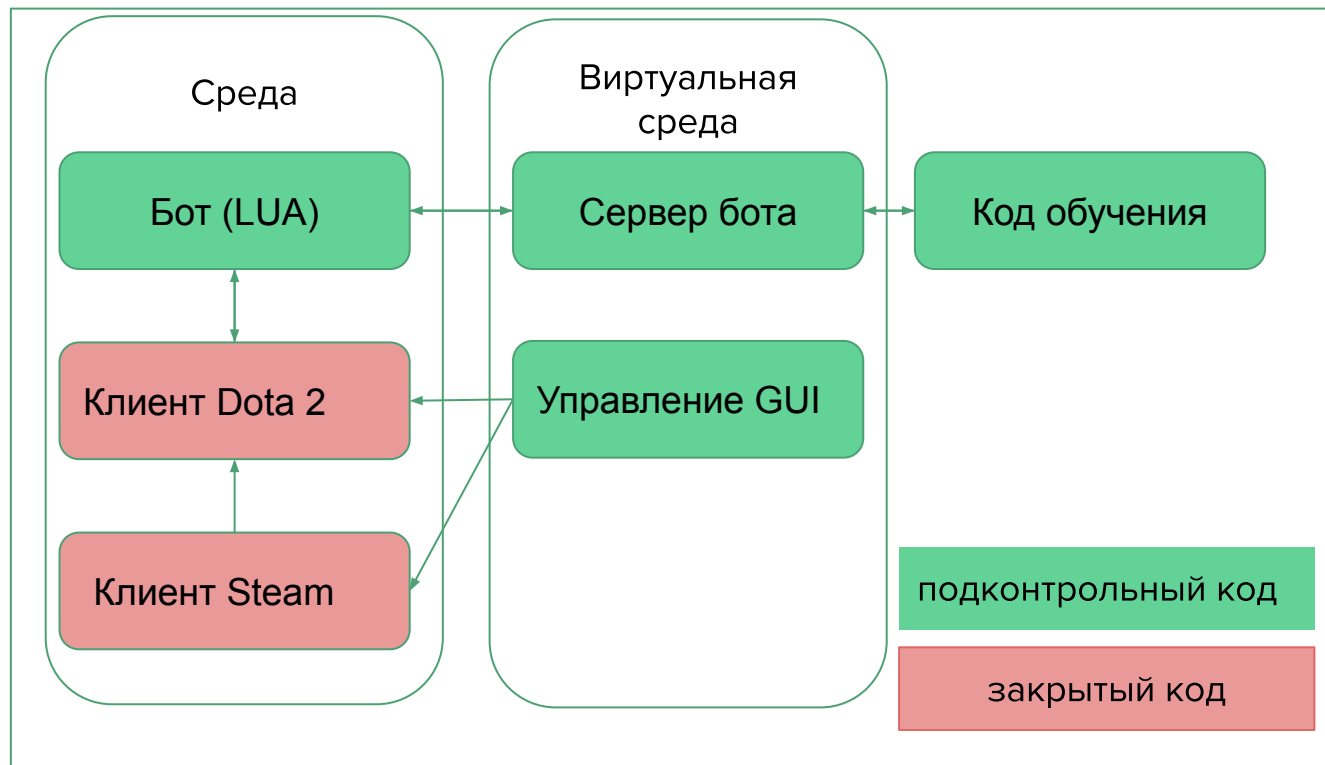
[JPBot] 1vs1 SF mid

StrawberryBot — 1vs1
Mid SF Bot

Mid Only Solo

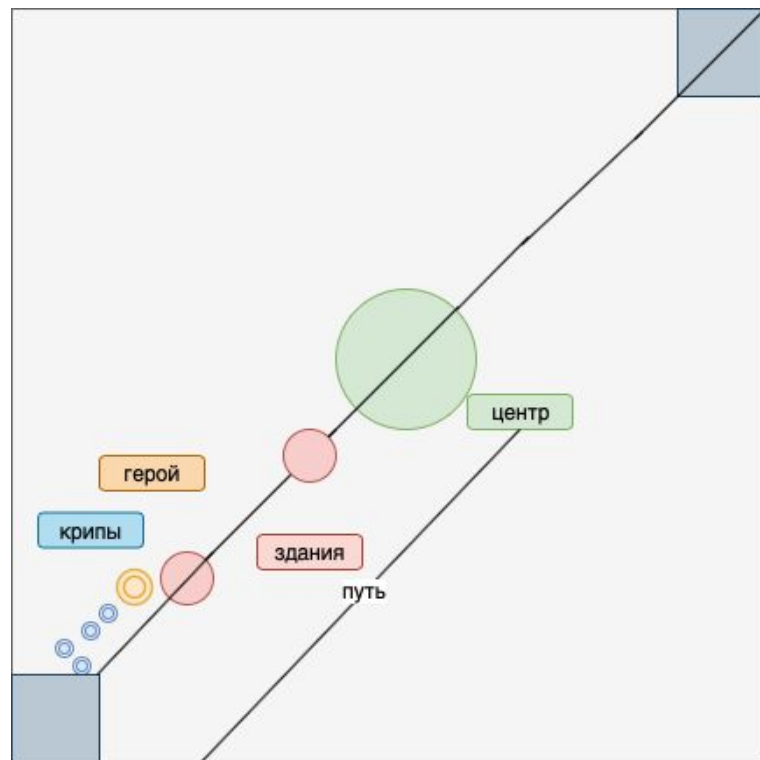
Боты из мастерской

Бот с машинным обучением



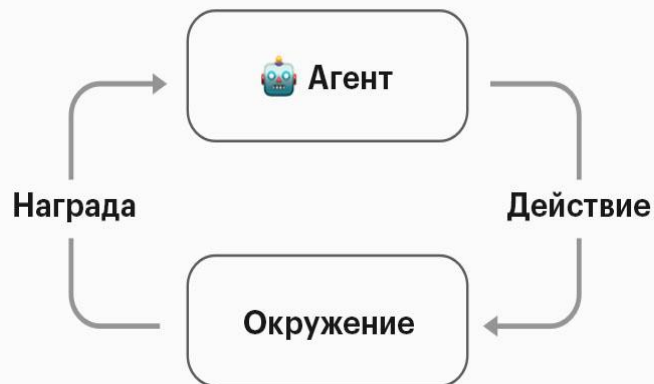
Выбор элемента игры для обучения - bodyblock

- 4 крипа (юнита) и один бот
- Bodyblock (от англ. body — тело, block — блокировать): максимально замедлить крипов до встречи с крипами противника, блокируя им проход моделью персонажа



Обучение с подкреплением

- $\{S\}$ — состояния
- $\{A\}$ — действия
- R — функция награды



Алгоритм

- Обучение с подкреплением (Q обучение, Q - целевая функция)
- Состояния (s) — положения на карте, действия (a) — перемещения бота, бинарная награда (r)
- Политика выбора действий

$$Q(s_{t+1}, a_{t+1}) = Q(s_t, a_t) + \alpha \cdot (r_t + \gamma \cdot \max_a Q(s_{t+1}, a) - Q(s_t, a_t))$$

α — фактор обучения γ — фактор дисконтирования

Метрика

- Метрика — время подхода крипов к центральной башне
- Для сравнения взяты максимальные значения на выборках



OpenAI	Мой бот
21 с.	26 с.

Результаты

- Изучена предметная область и существующие решения
- Реализован бот без использования машинного обучения
- Настроена инфраструктура для обучения бота
 - Реализован bodyblock

Реализация: <https://github.com/Pechckin>

Демонстрация

