МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ» (ФГБОУ ВО «ПГТУ»)

Кафедра иностранных языков и лингвистики

**Микриуправление в Starcraft с использованием обучения с подкреплением и постепенного трансферного обучения**

Перевод

Kun Shao, Yuanheng Zhu, Member, IEEE and Dongbin Zhao, Senior Member, «StarCraft Micromanagement with Reinforcement Learning and Curriculum Transfer Learning», принята IEEE Transactions on Emerging Topics in Computational Intelligence, 3 апреля 2018 г, [Электронный ресурс] URL: https://arxiv.org/pdf/1804.00810.pdf

Выполнил: студент гр.ПСм-21 Колчин И. А. Проверил: к.пед.н., доцент кафедры ИЯиЛ О.В. Филипчук.

Йошкар-Ола

2019

Оглавление

[Аннотация 3](#_Toc27040009)

[Оригинал 3](#_Toc27040010)

[Перевод 3](#_Toc27040011)

[Перевод статьи 5](#_Toc27040012)

[1. Введение 5](#_Toc27040013)

[2. Постановка проблемы и предпосылки 5](#_Toc27040014)

[2-a. Постановка проблемы 6](#_Toc27040015)

[2-b. Обучение с подкреплением 6](#_Toc27040016)

[2-c. Постепенное трансфертное обучение 6](#_Toc27040017)

[3. Обучение модели микроменеджменту 6](#_Toc27040018)

[3-a. Представление многомерного состояния 6](#_Toc27040019)

[3-b. Определение действий 6](#_Toc27040020)

[3-c. Архитектура нейросети 6](#_Toc27040021)

[4. Метод обучения микроменеджменту 6](#_Toc27040022)

[4-a. Много-агентный алгоритм с градиентным спуском и разделением параметров Сарса(λ) 6](#_Toc27040023)

[4-b. Функция вознаграждения 6](#_Toc27040024)

[4-c. Пропуск кадров 6](#_Toc27040025)

[5. Постановка эксперимента 6](#_Toc27040026)

[5-a. Сценарии микроуправления в Starcraft 6](#_Toc27040027)

[5-b. Обучение нейросети 6](#_Toc27040028)

[6. Результаты и обсуждения 6](#_Toc27040029)

[6-a. Микроуправление малого масштаба 6](#_Toc27040030)

[6-a1. Голиафы против Зилотов 6](#_Toc27040031)

[6-a2. Голиафы против Зерглингов 6](#_Toc27040032)

[6-b. Микроуправление большого масштаба 6](#_Toc27040033)

[6-b1. Морпехи против Зерглингов 6](#_Toc27040034)

[6-c. Анализ стратегий 6](#_Toc27040035)

[6-c1. Разделение врагов 6](#_Toc27040036)

[6-c2. Сохранение команды 7](#_Toc27040037)

[6-c3. Бей и беги 7](#_Toc27040038)

[6-c4. Существующие проблемы 7](#_Toc27040039)

[7. Заключение и будущие работы 7](#_Toc27040040)

[Приложения 7](#_Toc27040041)

[Терминологический словарь 7](#_Toc27040042)

# Аннотация

## Оригинал

Real-time strategy games have been an important field of game artificial intelligence in recent years. This paper presents a reinforcement learning and curriculum transfer learning method to control multiple units in StarCraft micromanagement. We define an efficient state representation, which breaks down the complexity caused by the large state space in the game environment. Then a parameter sharing multi-agent gradient-descent Sarsa(λ) (PS-MAGDS) algorithm is proposed to train the units. The learning policy is shared among our units to encourage cooperative behaviors. We use a neural network as a function approximator to estimate the action-value function, and propose a reward function to help units balance their move and attack. In addition, a transfer learning method is used to extend our model to more difficult scenarios, which accelerates the training process and improves the learning performance. In small scale scenarios, our units successfully learn to combat and defeat the built-in AI with 100% win rates. In large scale scenarios, curriculum transfer learning method is used to progressively train a group of units, and shows superior performance over some baseline methods in target scenarios. With reinforcement learning and curriculum transfer learning, our units are able to learn appropriate strategies in StarCraft micromanagement scenarios.

Index Terms — reinforcement learning, transfer learning, curriculum learning, neural network, game AI.

## Перевод

Стратегические игры в реальном времени были важной областью игр с искусственным интеллектом в последние годы. Эта работа представляет обучение с подкреплением и метод обучения с откладыванием плана, чтобы управлять множеством боевых единиц в StarCraft. Мы определяем эффективное представление состояний, которое устраняет сложность, вызванную большим пространством состояний в игровой среде. Затем для обучения юнитов предлагается алгоритм Sarsa (λ) (PS-MAGDS) с градиентным спуском для совместного использования параметров. Политика обучения распространяется среди наших подразделений, чтобы поощрять совместное поведение. Мы используем нейронную сеть в качестве аппроксиматора функции для оценки функции «действие-значение» и предлагаем функцию вознаграждения, чтобы помочь подразделениям сбалансировать их движение и атаку. Кроме того, метод трансферного обучения используется для расширения нашей модели до более сложных сценариев, что ускоряет процесс обучения и повышает эффективность обучения. В мелкомасштабных сценариях наши юниты успешно учатся сражаться и побеждать встроенный ИИ со 100% вероятностью выигрыша. В крупномасштабных сценариях метод обучения с откладыванием плана используется для постепенного обучения группы единиц и показывает превосходную производительность по сравнению с некоторыми базовыми методами в целевых сценариях. Благодаря обучению с подкреплением и обучению с откладыванием плана наши подразделения могут выучить подходящие стратегии в сценариях микроуправления StarCraft.

Ключевые слова: обучение с подкреплением, трансферное обучение, обучение по плану (или постепенное обучение), нейронная сеть, игровой ИИ.

# Перевод статьи

# 1. Введение

# 2. Постановка проблемы и предпосылки

## 2-a. Постановка проблемы

## 2-b. Обучение с подкреплением

## 2-c. Постепенное трансфертное обучение

# 3. Обучение модели микроменеджменту

## 3-a. Представление многомерного состояния

## 3-b. Определение действий

## 3-c. Архитектура нейросети

# 4. Метод обучения микроменеджменту

## 4-a. Много-агентный алгоритм с градиентным спуском и разделением параметров Сарса(λ)

## 4-b. Функция вознаграждения

## 4-c. Пропуск кадров

# 5. Постановка эксперимента

## 5-a. Сценарии микроуправления в Starcraft

## 5-b. Обучение нейросети

# 6. Результаты и обсуждения

## 6-a. Микроуправление малого масштаба

### 6-a1. Голиафы против Зилотов

### 6-a2. Голиафы против Зерглингов

## 6-b. Микроуправление большого масштаба

### 6-b1. Морпехи против Зерглингов

## 6-c. Анализ стратегий

### 6-c1. Разделение врагов

### 6-c2. Сохранение команды

### 6-c3. Бей и беги

### 6-c4. Существующие проблемы

# 7. Заключение и будущие работы

# Приложения

## Терминологический словарь

Обучение с подкреплением(reinforcement learning) – один из методов машинного обучения, в ходе которого испытуемая система (агент) обучается, взаимодействуя с некоторой средой.

Трансферное обучение(transfer learning) - это метод машинного обучения, в котором нейросеть, разработанная для одной задачи, повторно используется в качестве отправной точки для другой нейросети, которая решает схожую задачу.

Постепенное обучение(или обучение по плану)(curriculum learning) - обучение нейросети определённой последовательности постепенно усложняющихся задач, которые помогут в достижении конечной цели.