

Реализация симулятора кормящегося роя

Выполнил:

Алтухов Александр Дмитриевич, гр. 8303

Руководитель:
ЭВМ

Середа А.-В.И., д.т.н., профессор каф. МО

Консультант:

Азаревич А.Д., мл. инженер ООО «ЯЦТМС»

Актуальность

- Рост спроса на робототехнические решения;
- Задача фуражирования имеет множество вариантов практического применения (сбор ресурсов, поисково-спасательные операции, очистка окружающей среды);
- Изученные подходы на практике применяются неэффективно из-за комплексности задачи;

Решение – дополнительное исследование алгоритмов

Цель и задачи

Цель: реализация симулятора для исследования алгоритмов решения задачи фуражирования несколькими агентами

Задачи:

1. Обзор существующих алгоритмов
2. Обзор существующих симуляторов
3. Формулирование условий окружения
4. Разработка программы для обучения и симуляции
5. Исследование разработанного симулятора

1. Обзор существующих алгоритмов

Задача фуражирования – поиск и сбор объектов с последующей транспортировкой в точку сбора

Алгоритм	Требования к аппаратному обеспечению	Система управления	Эффективность
BFL-PSO	Низкая	Централизованная	Средняя
RGWO	Низкая	Централизованная	Средняя (success rate 80%)
DQN	Высокая	Децентрализованная	Высокая (success rate 90%)
DSQN	Высокая	Децентрализованная	Средняя
NN с феромонами*	Высокая	Децентрализованная**	Высокая

*Модель феромонного следа, оставляемого насекомыми в среде

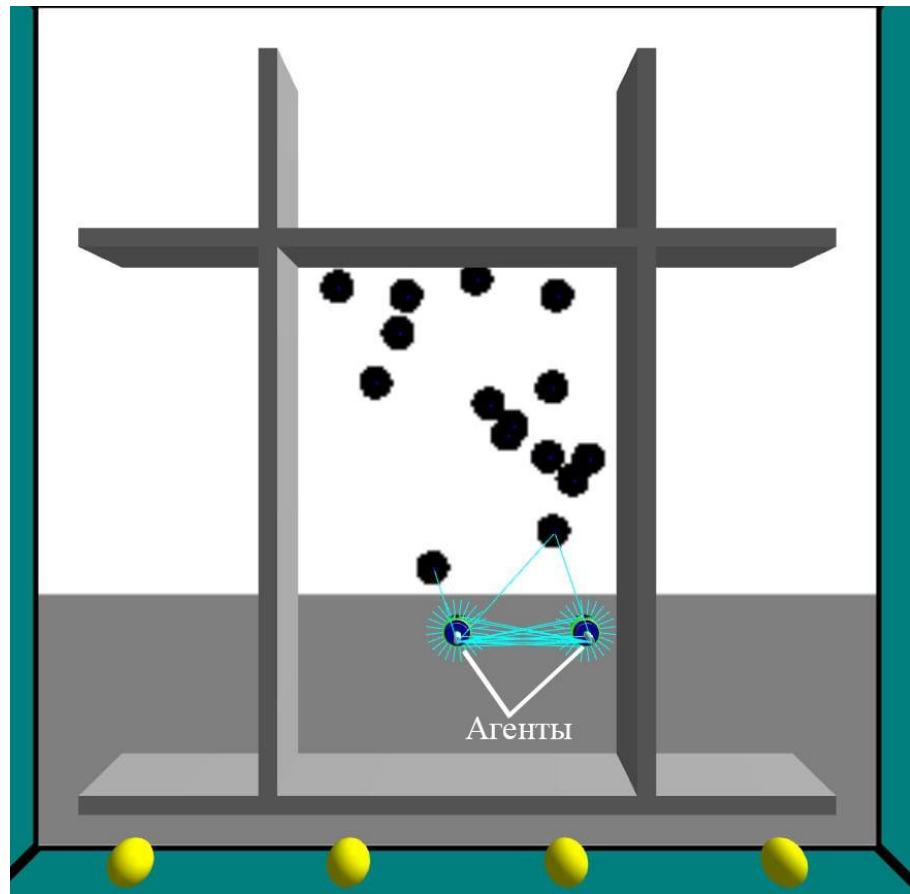
**При нахождении способа оставлять информацию о феромонах в физической среде

2. Обзор существующих симуляторов

Симулятор	Эффективная поддержка мультиагентности	Наличие фреймворка с поддержкой RL-алгоритмов	Наличие шаблона среды
ARGoS	Да	Неактуальные	Да
Gazebo	Нет	Неактуальные	Нет
Webots	Нет	Неактуальные, есть API для Python	Нет
CoppeliaSim	Нет	Неактуальные, есть API для Python	Нет

3. Формулирование условий окружения. Арена

- Арена: область, ограниченная стенами
- Две зоны: рабочая зона (белая) и домашняя зона (серая)
- Домашняя зона (д.з.) освещена
- Объекты для сбора – черные круги, испускающие синий свет
- Агенты испускают зеленый свет



3. Формулирование условий окружения.

Доступные наблюдения

- Данные о позиционировании агента (координаты, угол поворота)
- Текущая зона (рабочая или домашняя) – сенсор поверхностей
- Вектор освещенности – $\sum a_i$, где a_i это вектор света, полученный от сенсора освещения i
- Вектор препятствий – $\sum b_i$, где b_i это вектор препятствия, полученный от сенсора приближения i
- Сообщения других агентов о наличии еды в их области видимости – сенсор коммуникации
- Окружающие объекты (искомые и другие агенты) – камера кругового обзора

3. Формулирование условий окружения. Система наград

Путь до объекта сбора занимает много итераций. Решение: награды за выполнение дополнительных целей для стимуляции агента. Для избежания эксплуатации введены штрафы за невыполнение целей.

Основные действия	Начисляемые очки
Подбор объекта	100
Доставка объекта	300
Столкновение	-15
Передвижение	-0,5

Дополнительные цели	Начисляемые очки
Приближение к объекту	+1/-1
Приближение к домашней зоне с объектом	+1/-1
Отдаление от д.з. без объекта	+1/-1
Непрямое выполнение доп. цели	-0,1

4. Разработка программы для обучения и симуляции. Используемые технологии

- Языки Python 3, C++
- ARGoS – симуляция
- PettingZoo – посредник между симуляцией и обучением
- Tianshou – библиотека алгоритмов RL



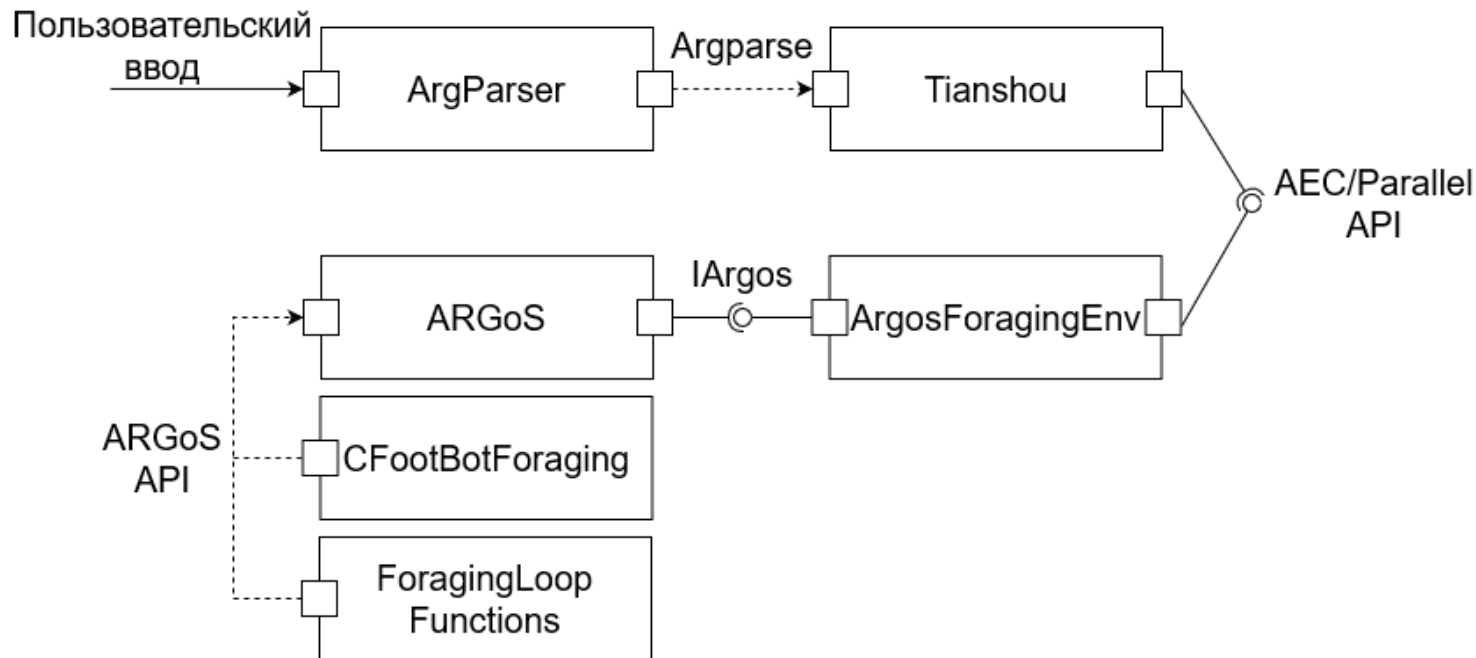
4. Разработка программы для обучения и симуляции. Архитектура

ArgParser – обработка пользовательского ввода

Tianshou – обучение агентов

ArgosForagingEnv – управление обучением

Argos и др. – симуляция и контроллеры роботов и среды



4. Разработка программы для обучения и симуляции. Модель нейросети

Гиперпараметры были подобраны в соответствии с результативностью сети в экспериментах (для каждого параметра взята серия значений из пяти вариаций, выбрано лучшее значение)

Параметр	Значение
Размер скрытых слоев	512-256-256-128
Learning rate	0,0003
Epsilon train/test	0,1/0,05
Количество эпох	150

Значения параметров

Размер пакета	Макс. значение reward
16	631,1
32	859,6
64	1015
128	799,8
256	887,6

Тюнинг размера пакета

5. Исследование разработанного симулятора.

Время работы

Было исследовано время работы различных модификаций DQN. Измерена скорость обновления сети во время обучения агентов в единицах обработанных наблюдений в секунду. Для оценки влияния скорости обновления приведено общее время обучения. Для измерений использовались наблюдения в случайных средах.

Алгоритм	Скорость обновления	Общее время обучения, мин
DQN	297	30,2
DDQN	291	30,16
RainbowDQN	273	34,41

Измерения проводились на ЭВМ с процессором Intel Core i5 12600KF, видеокартой Nvidia GeForce GTX 3080, (CUDA Cores: 8960, VRAM: 12GB)

5. Исследование разработанного симулятора. Сравнение алгоритмов

Алгоритм	Среднее макс. значение reward	Отклонение от макс. значения	Среднее значение reward
DQN	665,2	677,5	256,9
DDQN	870,8	570,4	301,6
RainbowDQN	1015	231,5	441,3

RainbowDQN достигает высшего максимального значения и делает это более стабильно, но проигрывает в скорости работы

Заключение

- Обзор алгоритмов показал эффективность алгоритмов DQN в данной задаче. Решено провести сравнение между модификациями DQN.
- Проведенный обзор симуляторов показал необходимость разработки собственного решения для сравнения алгоритмов. В качестве основы был использован симулятор ARGoS.
- Были сформулированы условия среды: арена со стенами и двумя зонами, проведен дизайн системы наград: награды за доставку объектов дополнены вспомогательными целями для стимуляции агентов
- Была разработана программа, связывающая симулятор с инструментами обучения (PettingZoo, Tianshou), и предоставляющая консольный интерфейс для управления средой и алгоритмами
- Проведенное сравнение показало, что алгоритм RainbowDQN наиболее эффективно показывает себя при решении задачи фуражирования

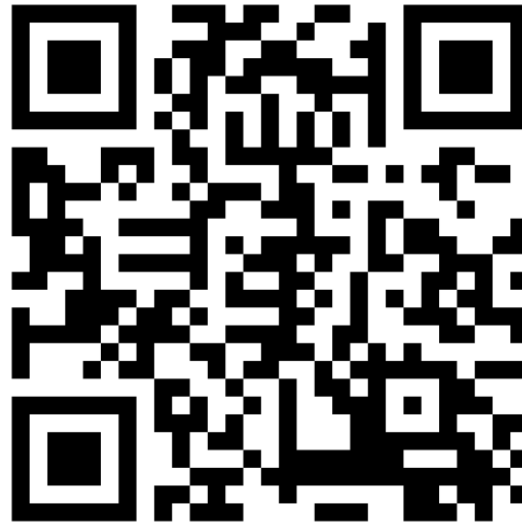
Направления дальнейшего развития:

- Подключение дополнительных библиотек машинного обучения
- Усложнение окружения для фуражирования

Апробация работы

- Репозиторий проекта

<https://github.com/Legendorik/robotic-swarm>



REPO

Запасные слайды

Термины и обозначения

- Термин – обозначение

Архитектура нейросети

Полный набор гиперпараметров

График наград