

МОДЕЛЮВАННЯ ПОВЕДІНКИ БОТІВ НА ОСНОВІ МЕТОДІВ АГЕНТО-ОРІЄНТОВАНОГО ПРОГРАМУВАННЯ

Терпіль Дмитро Олександрович

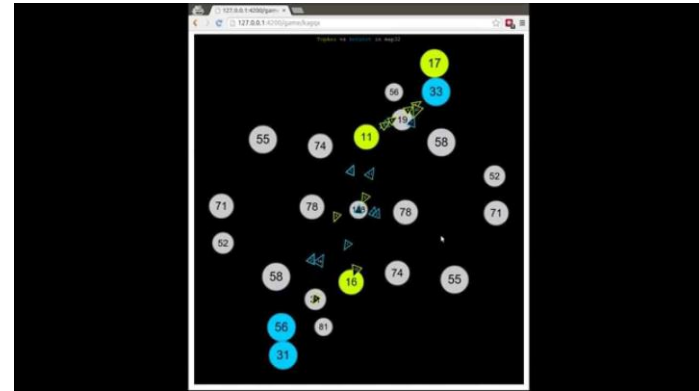
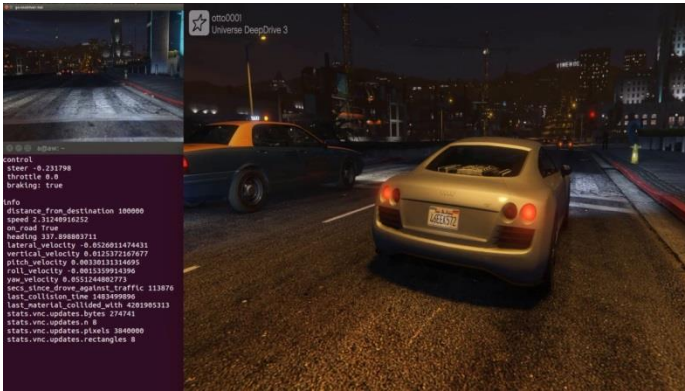
Керівник: к.т.н., доцент Шаповалова

Світлана Ігорівна

Актуальність



Приклади практичних застосувань Deep Q-Network



Мета дослідження

- Розробити алгоритмічну базу для донавчання штучних нейронних мереж під час експлуатації в змінних умовах навколишнього середовища

Задачі дослідження

- Провести аналіз нейромережових підходів для функціонування ботів в середовищах зі змієними характеристиками
- Дослідити методи донавчання нейронних мереж
- Дослідити алгоритми функціонування нейронної мережі DQN та її вдосконалення
- Дослідити сучасні програмні середовища навчання інтелектуальних агентів
- Розробити алгоритмічну базу доповнення знань вже навчених інтелектуальних агентів
- Розробити загальне рішення середовища навчання при динамічно змінюваних умовах
- Експериментально перевірити запропоновані рішення

Об'єкт дослідження

- Навчання штучних нейронних мереж Deep Q-Network

Предмет дослідження

- Донавчання DQN в змінних умовах навколишнього середовища

Наукова новизна

- 1. Запропоновано програмний комплекс для самостійного навчання з підкріпленням (reinforcement learning, RL) бота в умовах динамічної зміни характеристик навколишнього світу для віртуальної підготовки інтелектуальних ботів бортових систем.
- 2. Удосконалено ігровий рушій Unity можливостями моделювання візуальних ефектів, та реалізації фізичних сил для моделювання навколишнього середовища.
- 3. Набуло подальшого розвитку технологія Lightmapping за рахунок оптимізації представлення результатів рендеренгу освітлення в сценах ігрового світу. Для прискорення та зменшення ресурсів пам'яті.

Схема навчання з підкріпленням

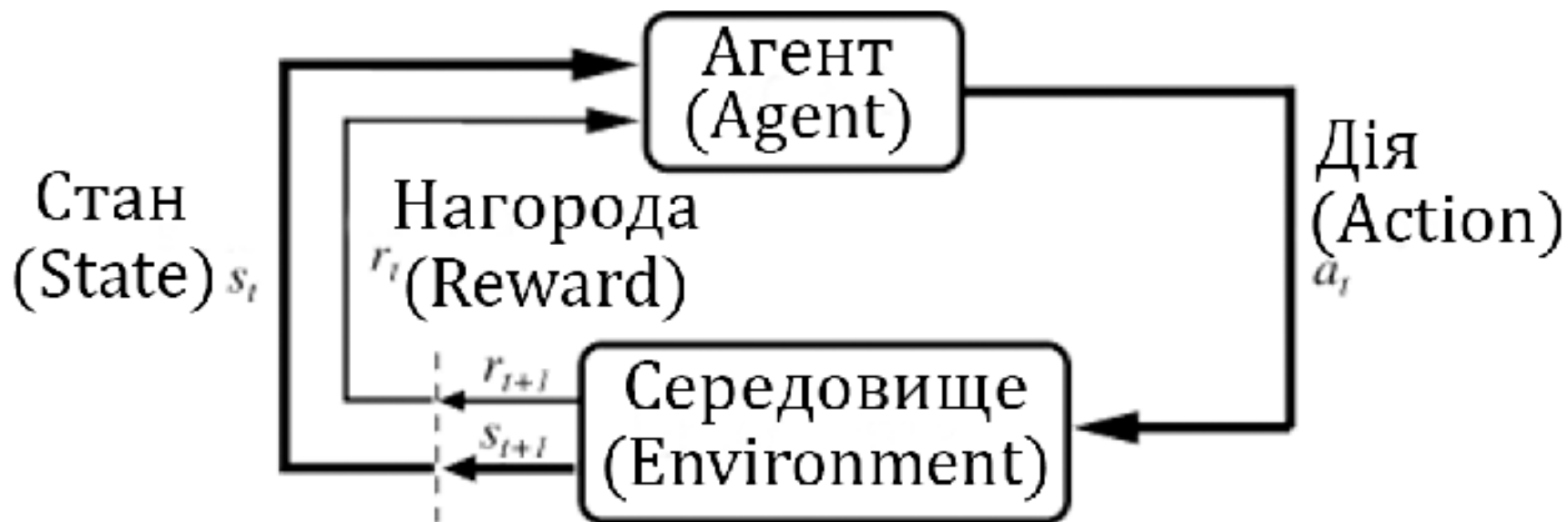
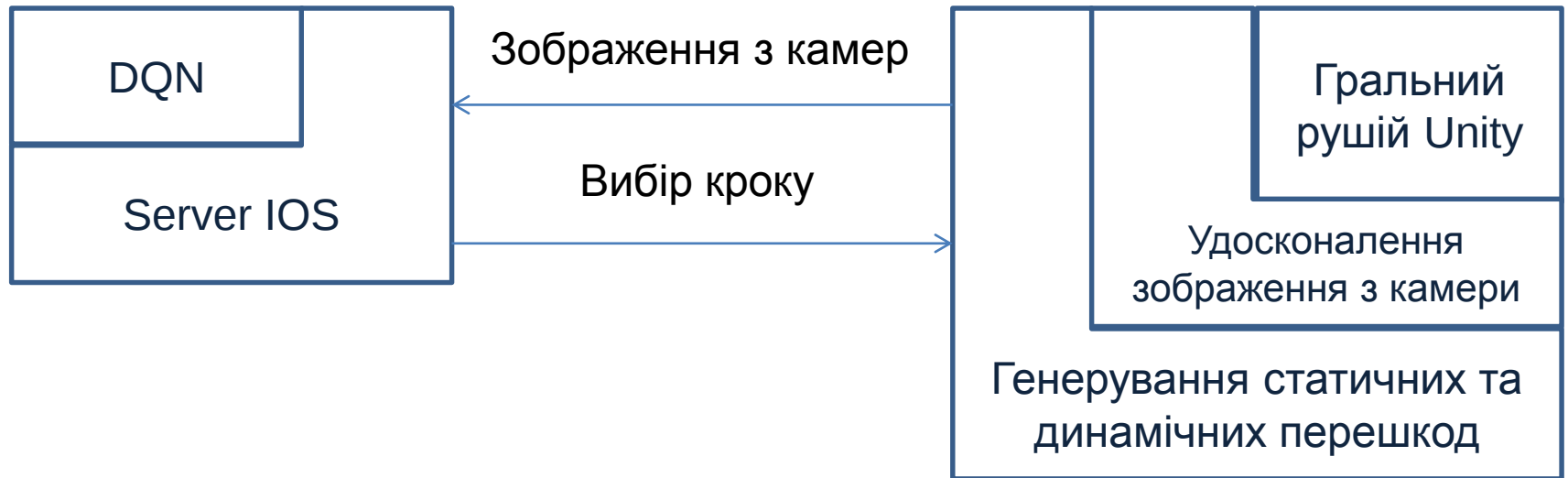
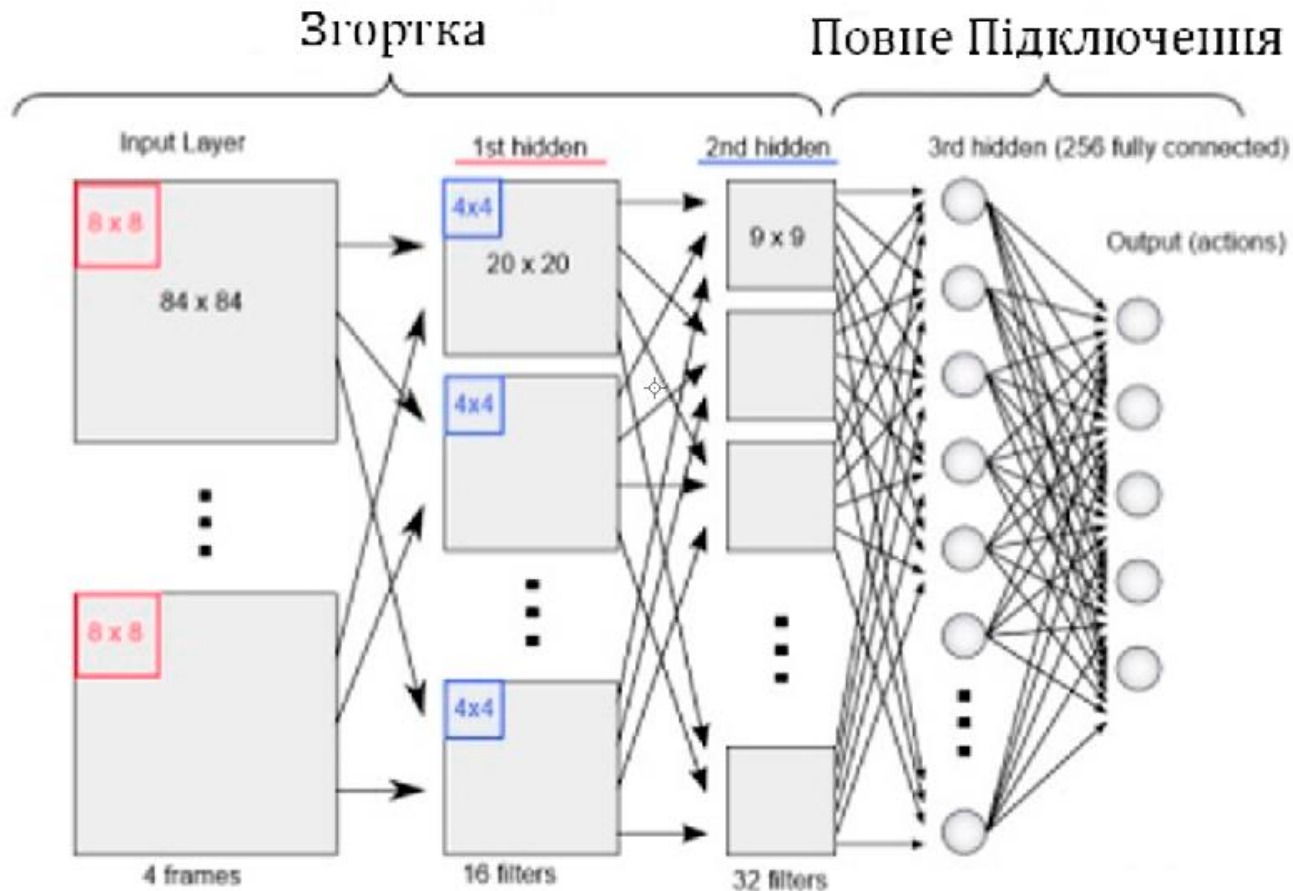


Схема навчання в змінному середовищі



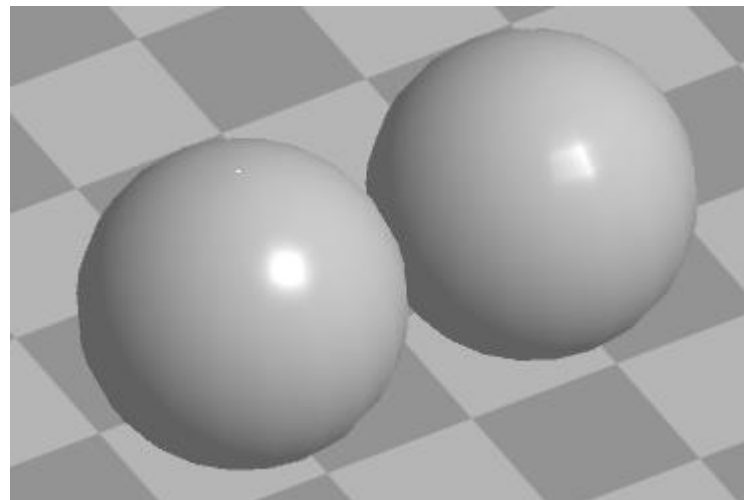
Архітектура Deep Q-Network



Вдосконалення Unity: реалістичне відтворення тіней



Вдосконалення Unity: реалізація оптичних блоків



Відтворення статичних перешкод фізичного світу



Відтворення динамічних перешкод фізичного світу



Тестова задача 1: базовий фізичний світ

Характеристики середовища:

- Симуляція сили тяжіння,
- Симуляція освітлення,
- Генерація винагород для робота.

Характеристики робота:

- Прийняття рішення на кожному кроці,
- Можливість руху, обертів на 45 градусів, бездіяльність.
- Задача робота в найкоротший час знайти і зібрати всі винагороди.

Тестова задача 2: фізичний світ з статичними перешкодами

Характеристики середовища:

- Симуляція сили тяжіння,
- Симуляція освітлення,
- Генерація винагород для робота.
- Генерація статичних перешкод.

.

Тестова задача 3: фізичний світ з статичними та динамічними перешкодами

Характеристики середовища:

- Симуляція сили тяжіння,
- Симуляція освітлення,
- Генерація винагород для робота.
- Генерація статичних перешкод.
- Генерація динамічних перешкод.

Результати випробування

Задача	Перенавчання			Донавчання		
	Час навчання	Частота невдач	Кількість кроків	Час навчання	Частота невдач	Кількість кроків
Базовий фізичний світ	12 годин	2%	20-30	12 годин	1%	
Фізичний світ зі статичними перешкодами	18 годин	3%	40-50	15годин	2%	30-45
Фізичний світ зі статичними і динамічними перешкодами	24 години	5%	60-70	15 годин	4%	60-70

Висновки

1. Проведено аналіз концепцій навчання штучного інтелекту для прийняття рішень з досягнення кінцевої цілі у середовищі зі змінними характеристиками. Обгрунтовано використання нейронної мережі Deep Q-Network;
2. Проведено аналіз існуючих інтелектуальних агентів, реалізованих на основі Deep Q-Network. Визначено недоліки середовищ, в яких проводилось навчання, та напрями їх вдосконалень;
3. Представлено концепцію функціонування нейронної мережі Deep Q-Network та алгоритм навчання з підкріпленням, які стали основою інтелектуального агента.
4. На основі ігорних середовищ Atari визначено фізичні характеристики навчання штучного інтелекту в реалістичних умовах.
5. Визначено недоліки ігрового рушія Unity та додаткові можливості, які потрібні для реалізації середовища навчання.
6. Набула подальшого розвитку технологія Lightmapping для оптимізації використання обчислювальних ресурсів за рахунок оптимізації представлення результатів рендеренгу освітлення в сценах ігрового світу.
7. Розроблено представлення фізичного світу та цілі навчання/донавчання робота.
8. Проведено обчислювальні експерименти навчання і донавчання інтелектуального агента при зміні умов середовища. Випробовування довели, що донавчання ефективніше повторного навчання за швидкістю та зменшенню кількості помилок при прийнятті рішень.