

# Дослідження методів штучного інтелекту в мультиагентних системах



Автор роботи:  
Студентка гр. ДА-61  
Полозюк Катерина Олександрівна



## **Об'єкт дослідження**

мультиагентні системи

## **Предмет дослідження**

мульттагентні системи, в основі яких лежать алгоритми глибокого навчання



## Мета та завдання

**Метою** даної роботи є:

- аналіз існуючих рішень – дослідження наукових праць останніх років у галузі МАС з використання глибинного навчання з підкріпленням та визначити загальні підходи щодо створення даного типу МАС.
- визначити задачу для розв'язку та провести власне дослідження різних методів.

**Результатом** проведених досліджень є створені моделі на основі глибинного навчання для вирішення задачі sudoku та подальший аналіз результатів їх роботи.



## Актуальність роботи

- Актуальність даної роботи обумовлена розвитком апаратного забезпечення, що дозволив науковій спільноті досить швидко розробляти та навчати агентів на основі глибинного навчання з підкріпленням, а отже – проводити велику кількість досліджень з ними, що привело до появи великої кількості новітніх досліджень в даній області.
- Особливо слід звернути увагу на дослідження, пов'язані з використанням інтелектуальних агентів в мультиагентних системах.

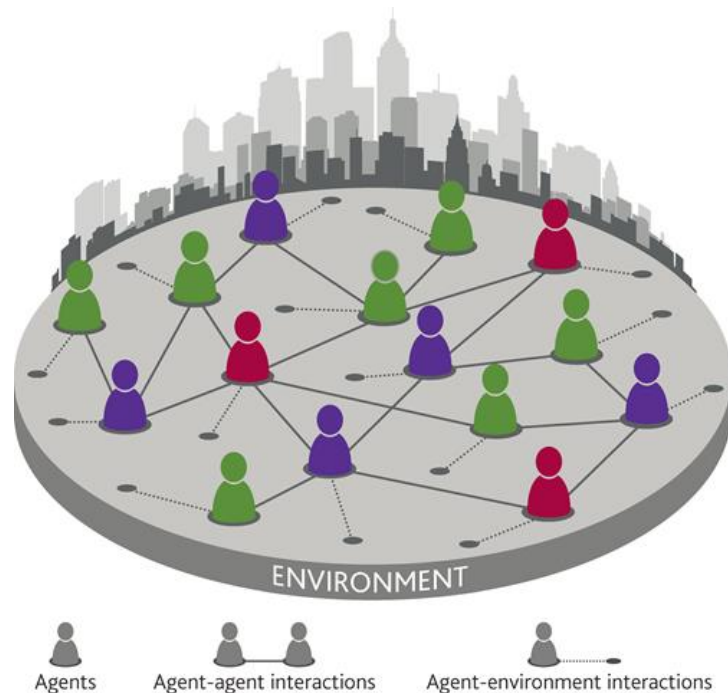


# Мультиагентна система

**МАС** – система, котра складається з:

- Множини агентів – **I**;
- Середовища – **E**;
- Множини станів середовища – **S**;
- Та множини дозволених дій – **A**,

в котрій агенти з множ. **I** виконують дії з множ. **A**, переводячи середовище між станами з множ. **S**, до тих пір поки **E** не досягне термінального стану. Тобто до тих пір, поки агенти не виконають поставлену перед ними задачу.





## Дослідження наукових праць

- Q-Learning, що добре лягає в парадигму МАС, де є середовище та зворотній зв'язок від нього.
- Різні власні реалізації елементів МАС, більшість з них закрита, що не дає можливості іншим дослідникам продовжити експерименти.
- Використання мови програмування Python, та фреймворку OpenAI Gym для створення середовищ.



# Використані методи для навчання агенту

1 DQN

3 TD

5 PPO

2 DDQN

4 TD3DPG

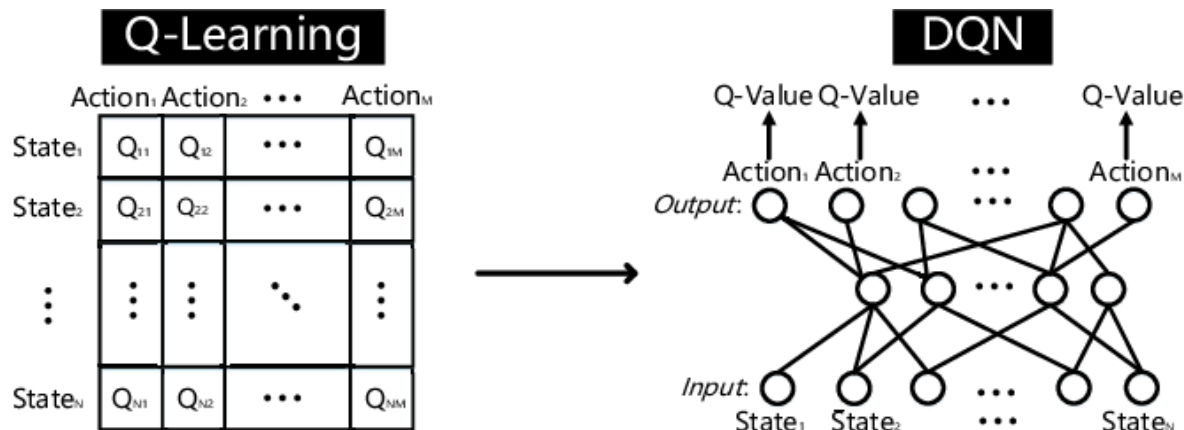
6 NeuralMCTS



# Метод DQN

**DQN** – алгоритм ГН з підкріпленням, що базується на Q-learning та Q-table.

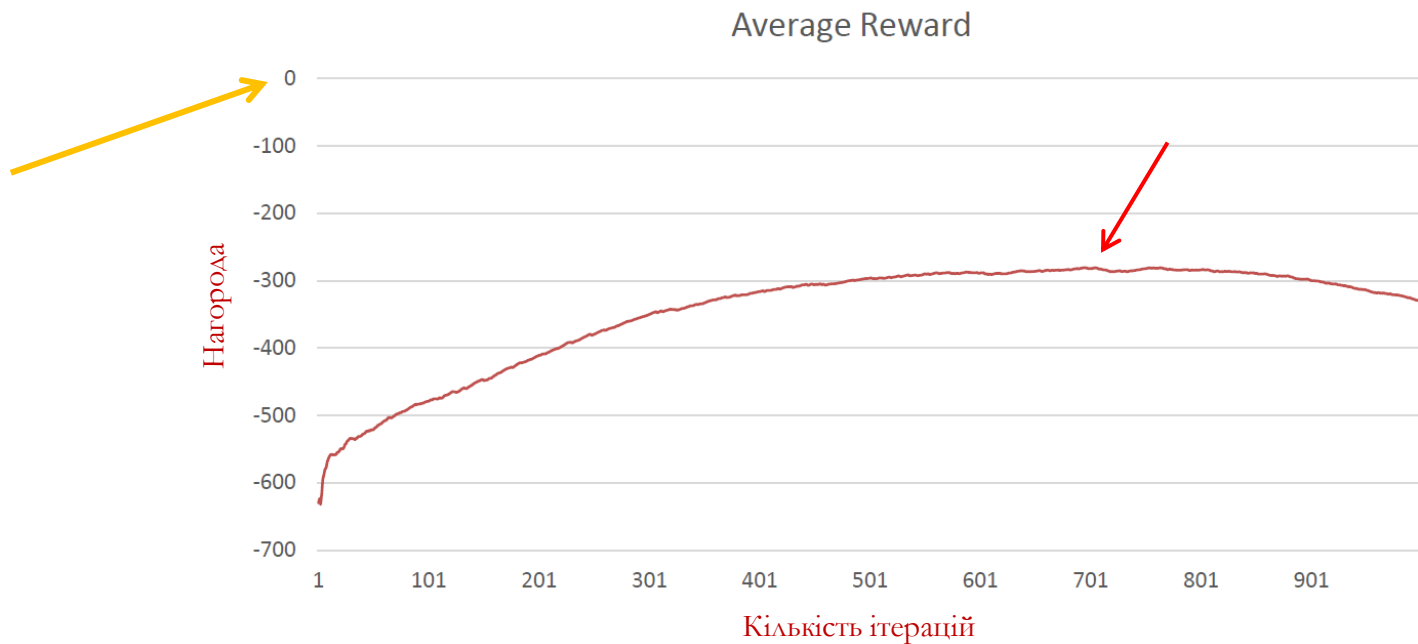
- В даному типі навчання головну роль відіграє оцінка якості дій агента через значення Q-value.
- Q-table – масив записів виду MDP (Markov Decision Process).
- НМ в даному методі займається апроксимацією формули Беллмана.







# Результат роботи DQN



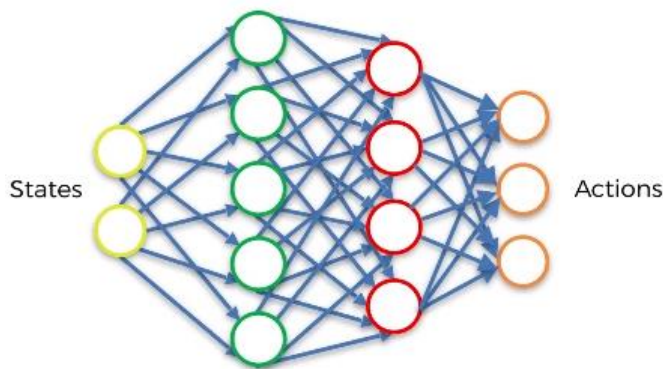
Середні значення нагороди для DQN агента (9x9, 1000 ітерацій)



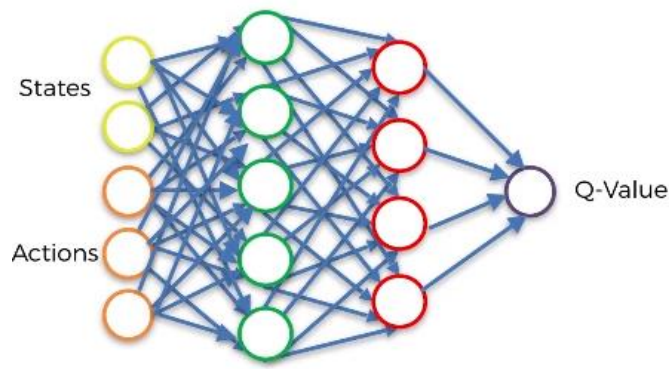
# Метод DDQN

**DDQN** – алгоритм ГН з підкріпленням, що базується на DQN з використанням моделі Actor-Critic.

- В якості функції втрат використовуую середньоквадратичну похибку (MSE).



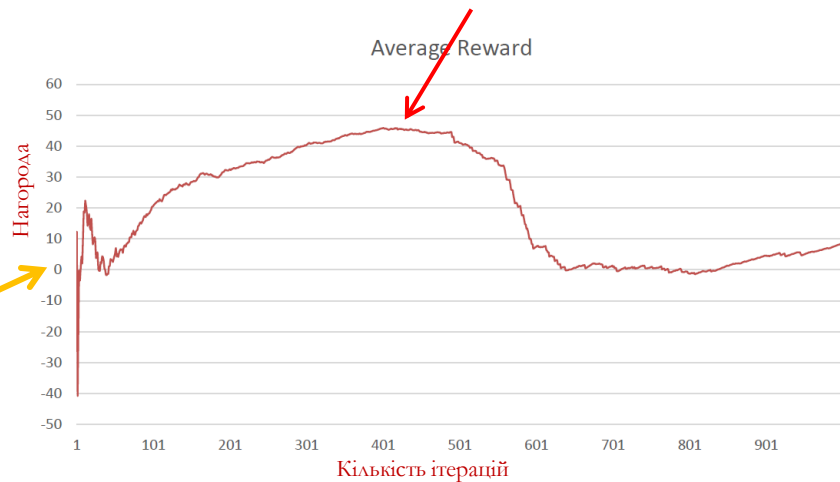
**Actor**



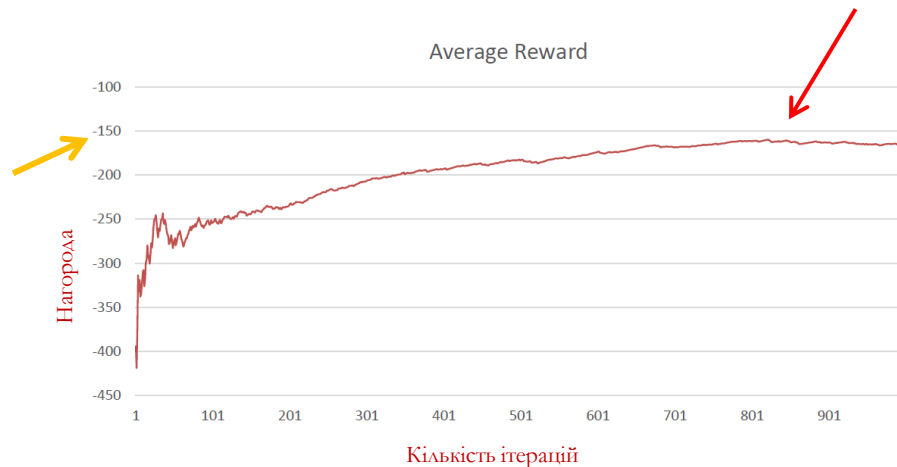
**Critic**



# Результат роботи DDQN



Середні значення нагороди для DDQN агенту (2 x 2, 1000 ітерацій)



Середні значення нагороди для DDQN агенту (4 x 4, 1000 ітерацій)



# Результат роботи DDQN



Середні значення нагороди для DDQN агенту (9 x 9, 1000 ітерацій)



## Метод TD

**TD** – модель ГН з підкріпленням, що базується на DDQN з використанням алгоритму оптимізації TD замість MSE.

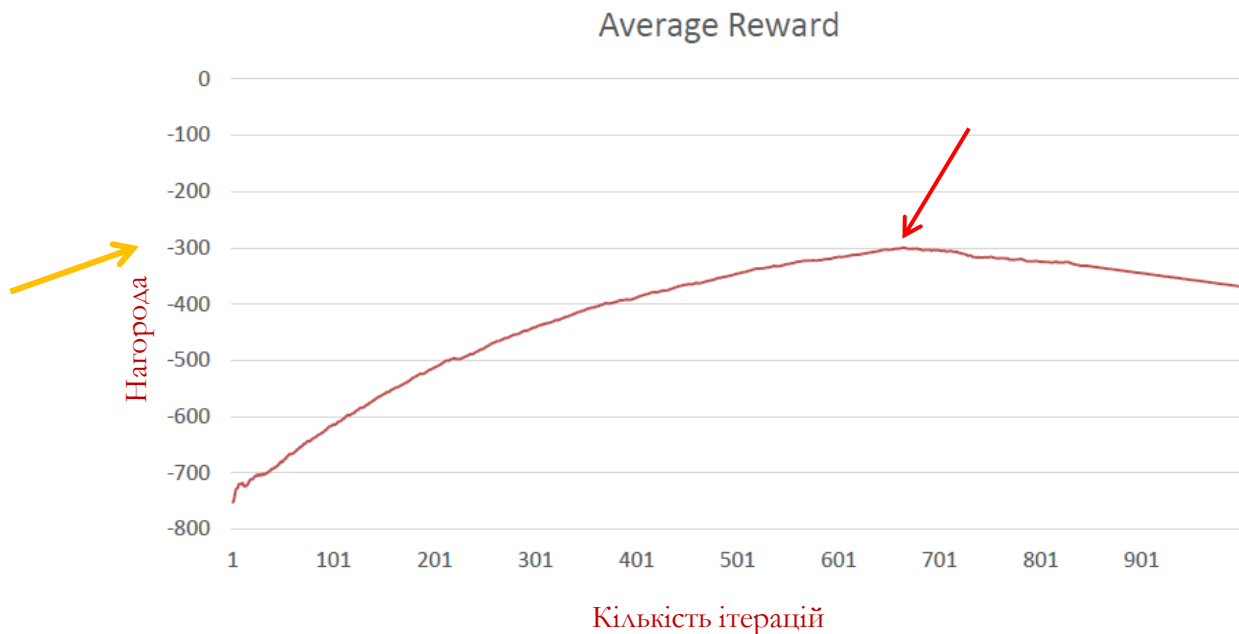
- Існує декілька методів реалізації TD. Найбільш відомим є TD(0). В даній роботі використовується найбільш ефективний – SARSA (State – Action – Reward – State - Action).

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

Також розглядався метод TD3DPG, проте його результати були гірші ніж у всіх інших моделей.



# Результат роботи TD



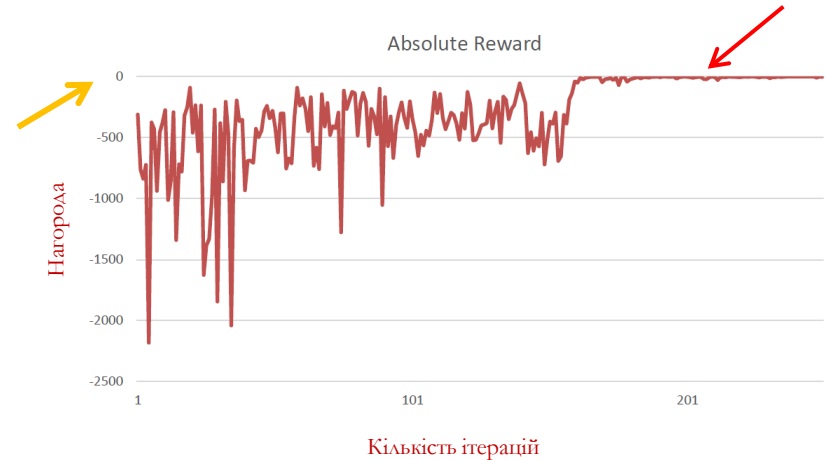
Середні значення нагороди для TD агенту (9 x 9, 1000 ітерацій)



# Метод PPO

**PPO** — метод на основі DDQN з використанням TRPO (Trust Region Policy Optimization).

- Використовує динамічну оцінку кроку в алгоритмі оптимізації (градієнтний спуск).



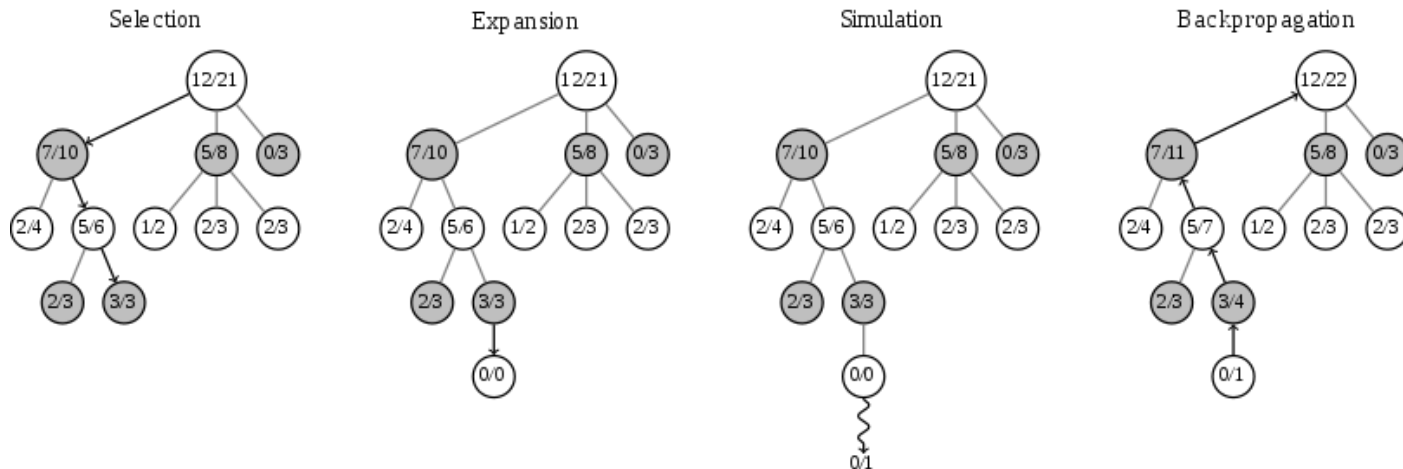
Середні(зліва) та абсолютні(зправа) значення нагороди для PPO агенту (9 x 9, 250 ітерацій)



# Метод MCTS

**NeuralMCTS** – метод на основі MCTS з використанням нейронної мережі для оцінки ймовірностей.

- Представляє собою ймовірнісний обхід дерева, де кожна гілка це дія з вагою, що рівна ймовірності, а вузол це стан середовища.



Приклад роботи MCTS





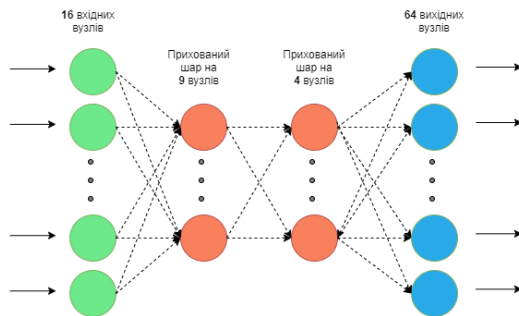
# Результат роботи MCTS

## Судоку 4 x 4

Solution Rate



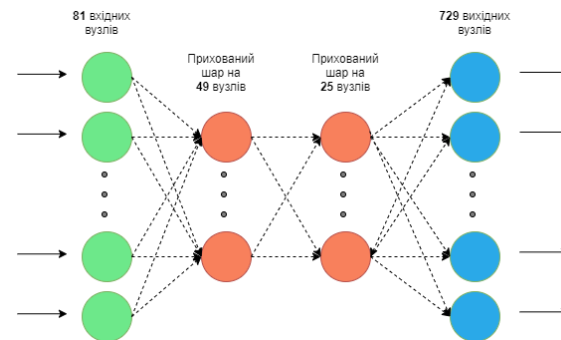
Точність вирішення судоку 4 x 4 MCTS в залежності від епізодів (всього - 10)



Архітектура НМ для MCTS для вирішення судоку 4 x 4

## Судоку 9 x 9

- Проведено 2 повних епізоди (по 1000 ітерацій).
- Точність розв'язку становила 35% та 46% на відповідних ітераціях.

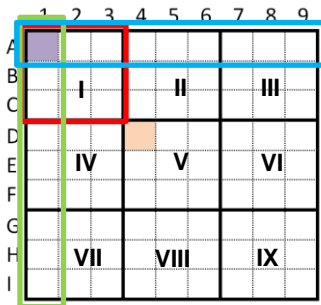


Архітектура НМ для MCTS для вирішення судоку 9 x 9

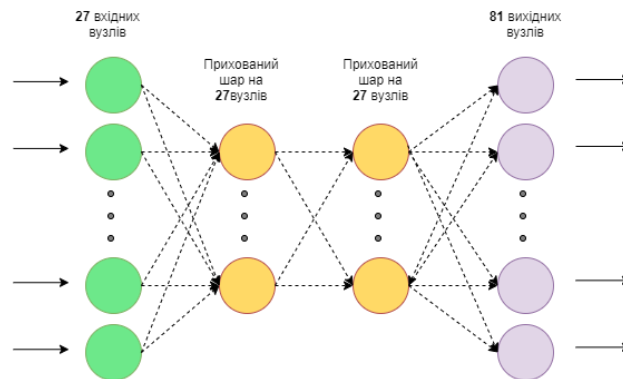
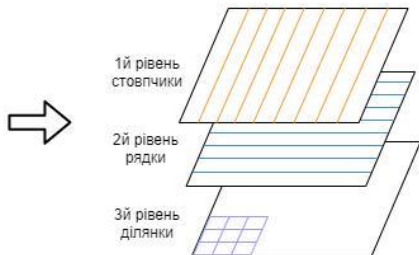
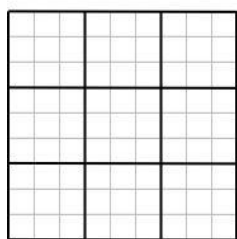


# МАС на основі МСТS

Останнім експериментом є імплементація **МАС** для **судоку 9x9** з використанням **27** агентів, кожен з яких відповідає за колонку, лінію чи квадрат, та повинен обмінюватися інформацією з іншими агентами для успішного заповнення клітинок.



Класичне судоку 9x9



- Середній час навчання одного з агентів – **20 год**;
- досягнення позитивного росту – з **3-4 ітерації** (час всього навчання 3.5 днів);
- Середня точність – **коливається на рівні 40-50%**.



## План подальших робіт

Наступними кроками в дослідженні мають бути:

- розрахунок на більш потужних комп'ютерних кластерах, що дозволить робити більше експериментів та прискорити час обчислення однієї ітерації.

На даний момент нестача обладнання є основною проблемою, котра не дозволяє провести більшу кількість експериментів та отримати більш точні результати.



## Висновки

- Наукоємною складовою роботи є створення моделей на основі DQN та Neural MCTS з їх подальшим аналізом та порівнянням на обраній раніше задачі (судоку).
- За результатами дослідження виявлено, що методи на основі Q-Learning, незалежно від способу апроксимації формули Беллмана (Q-value), погано вирішують комплексні задачі в середовищах з великою кількістю станів.
- Для розв'язку подібних задач успішним є використання методів на основі MCTS.
- За допомогою методів MCTS можна побудувати успішну Мультиагентну систему.
- Єдиною перешкодою для вирішення комплексних задач є обмеження наявного апаратного забезпечення.



**Дякую за увагу!**

Студентка групи ДА-61  
Полозюк Катерина Олександрівна