

НАВЧАННЯ НЕЙРОКОНТРОЛЕРА

Нейроконтролер в ігровому середовищі є постійним елементом персонажа. Далі ми обговоримо навчання нейроконтролера в режимі реального часу.

Навчання нейроконтролера полягає в наданні тим, хто навчається, прикладів із невеликої групи бажаних дій. Потім слідує виконання алгоритму зворотного розповсюдження з урахуванням бажаного результату і дійсного результату. Наприклад, якщо персонаж має пістолет, здоровий і бачить одного ворога, бажана дія – атакувати. Проте якщо персонаж здоровий, має ніж, але бачить два вороги, то правильна дія – сховатися.

Дані для тестування

Дані для тестування передбачають декілька сценаріїв з набором дій. Оскільки вимагається, щоб нейроконтролер поведився так само, як справжня людина, ми не будемо навчати його для кожного випадку. Мережа повинна розрахувати реакцію на входи і виконувати дію, яка буде схожою на навчальні сценарії. Приклади, які використовувалися для навчання мережі, представлені в табл. 1.

Дані, що наведені в табл. 1, були передані мережі в довільному порядку під час навчання за допомогою алгоритму зворотного розповсюдження. Графік зниження середньої помилки показаний на рис. 6.

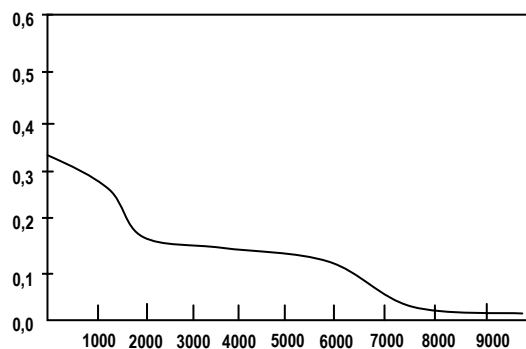


Рис. 6. Приклад виконання алгоритму зворотного розповсюдження для нейроконтролера

Таблиця 1
Приклади, які використовуються для навчання
нейроконтролера

| Здоров'я | Має ніж | Має пістолет | Вороги | Поведінка |
|----------|---------|--------------|--------|-----------|
| 2 | 0 | 0 | 0 | Ухилятися |
| 2 | 0 | 0 | 1 | Ухилятися |
| 2 | 0 | 1 | 1 | Атакувати |
| 2 | 0 | 1 | 2 | Атакувати |
| 2 | 1 | 0 | 2 | Ховатися |
| 2 | 1 | 0 | 1 | Атакувати |
| 1 | 0 | 0 | 0 | Ухилятися |
| 1 | 0 | 0 | 1 | Ховатися |
| 1 | 0 | 1 | 1 | Атакувати |
| 1 | 0 | 1 | 2 | Ховатися |
| 1 | 1 | 0 | 2 | Ховатися |
| 1 | 1 | 0 | 1 | Ховатися |
| 0 | 0 | 0 | 0 | Ухилятися |
| 0 | 0 | 0 | 1 | Ховатися |
| 0 | 0 | 1 | 1 | Ховатися |
| 0 | 0 | 1 | 2 | Бігти |
| 0 | 1 | 0 | 2 | Бігти |
| 0 | 1 | 0 | 1 | Ховатися |

У більшості випадків мережа успішно проходить навчання на всіх представлених прикладах. При деяких запусках один або два приклади призводять до неправильних дій. При збільшенні кількості прихованих комірок навчання проходить ідеально. Проте в цьому випадку для роботи нейроконтролера вимагається набагато більше комп'ютерних ресурсів. Тому були вибрані три приховані комірки з декількома циклами навчання. Для навчання нейроконтролера було виконано стільки запусків, скільки було потрібно для отримання ідеального результату на основі тестових даних.

Щоб протестувати нейроконтролер, мережі були представлені нові приклади. Це дозволило визначити, як мережа буде реагувати на сценарії, про які їй нічого невідомо. Дані тести дають відповідь на питання, наскільки добре нейроконтролер може генерувати і виконувати необхідні дії, реагуючи на непередбачену ситуацію.

Якщо запропонувати нейроконтролеру сценарій, в якому персонаж повністю здоровий, володіє зброєю двох видів і бачить два вороги (тобто 2:1:1:1), нейроконтролер вибере дію “атакувати”. Це розумна реакція на дану ситуацію. Тепер розглянемо сценарій, в якому персонаж повністю здоровий, володіє ножем і бачить трьох ворогів (тобто 2:1:0:3). Нейроконтролер вибирає дію “сховатися” – цілком розумний вибір у даній ситуації. Інші приклади показані в табл. 2.

Таблиця 2

Приклади, що ілюструють правильну генерацію дій

| Здоров'я | Має ніж | Має пістолет | Вороги | Поведінка |
|------------|---------|--------------|--------|-----------|
| Добре (2) | Так | Так | 1 | Атакувати |
| ОК(1) | Так | Так | 2 | Ховатися |
| Погане (0) | Ні | Ні | 0 | Ухилятися |
| Погане (0) | Так | Так | 1 | Ховатися |
| Добре (2) | Ні | Так | 3 | Ховатися |
| Добре (2) | Так | Ні | 3 | Ховатися |
| Погане (0) | Так | Ні | 3 | Бігти |

Таким чином, нейроконтролер правильно генерує дію із заданого набору у відповідь на нові обставини (табл. 2). Хоча його не навчали конкретно для цих прикладів, він здатний правильно на них реагувати.

На прикладах показано тільки невелику галузь застосування методів штучного інтелекту в інформаційних системах.