

ФОРМИРОВАНИЕ ПОВЕДЕНИЯ ИГРОВОГО ИИ: ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ ДЛЯ АВТОМАТИЧЕСКОЙ ГЕНЕРАЦИИ ИЕРАРХИЧЕСКИХ СЕТЕЙ ЗАДАЧ

Костюк Д.С. (ИТМО)

Научный руководитель – Бравичев К.А. (ИТМО)

Введение. Разработка игрового искусственного интеллекта (ИИ) требует значительных ресурсов, особенно при создании сложных моделей поведения. Поведение в игровом ИИ – это совокупность действий агента, определяемых правилами, логикой и алгоритмами, направленными на достижение поставленных целей в виртуальном мире. Существует множество типов моделей поведения: конечные автоматы, деревья решений, системы на основе правил, целенаправленное планирование действий, а также иерархические сети задач (HTN). HTN представляют собой метод декомпозиции сложных задач на более простые подзадачи, что делает их удобным инструментом для создания комплексных систем поведения. Однако процесс ручного проектирования HTN сложен и трудозатратен, что ограничивает их применение в крупных игровых проектах. Автоматическая генерация HTN способна снизить затраты на разработку игрового ИИ. Одним из перспективных решений является использование методов машинного обучения, в частности обучения с подкреплением (RL). Данный метод позволяет агентам обучаться на основе взаимодействия с окружающей средой, формируя стратегию поведения на основе получаемых вознаграждений.

Основная часть. Поведение игрового ИИ определяется совокупностью решений, принимаемых агентом в ответ на изменения окружающей среды. Различают несколько типов систем управления поведением, включая конечные автоматы, деревья поведения и HTN.

Иерархическая сеть задач (HTN) – это метод планирования в области искусственного интеллекта, который используется для решения сложных задач путем декомпозиции их на более простые подзадачи. HTN принимает на вход задачу и предоставляет на выходе серию шагов, которые ее решают. В терминах иерархических сетей задач серия шагов называется планом. Основными элементами HTN являются:

- 1) Задачи (Tasks) – основные цели, которые необходимо достичь;
- 2) Методы (Methods) – способы декомпозиции задач на подзадачи;
- 3) Операторы (Operators) – конечные действия, которые может выполнить агент;
- 4) Сеть задач (Tasks network) – график, представляющий иерархическую структуру задач и подзадач, иллюстрирующий взаимосвязи и зависимости между ними;
- 5) План (Plan) – последовательность операторов и подзадач, которые приводят к решению поставленной задачи.

С того момента как в 1975 году [1] было предложено планирование HTN, оно применялось во многих областях, включая:

- 1) Разработку планов действий правительственными ведомствами [2];
- 2) Управление интеллектуальными роботами [3];
- 3) Управление поведением персонажей в шутерах от первого лица (FPS) и стратегиях в реальном времени (RTS) в играх [4].

Обучение с подкреплением (RL) представляет собой метод, при котором агент взаимодействует с окружающей средой, формируя поведенческую стратегию на основе получаемых вознаграждений. Цель RL-алгоритма заключается в том, чтобы найти такие правила поведения, которые позволяют достичь максимальных ожидаемых результатов в окружающей среде, в которой работает алгоритм. Основные подходы к RL:

- 1) Модели на основе таблиц (Q-Learning, DQN) – обучение через оценку ценности состояний или действий;
- 2) Политико-ориентированные методы (REINFORCE) – обучение непосредственного выбора действий;

3) Иерархическое обучение с подкреплением (HRL) – позволяет агенту обучаться на нескольких уровнях абстракции.

В работе Hogg et al. (2010) [5] рассматривается интеграция HTN с методами RL, что позволяет улучшить качество планирования и адаптивность ИИ. Авторы предлагают использование метода Q-SHOP, который сочетает HTN и RL для создания высококачественных планов. Обучение с подкреплением помогает агенту выбрать оптимальные методы декомпозиции задач, что улучшает структуру планов. Эксперименты показывают, что такая интеграция позволяет создавать более эффективные и адаптивные планы, превосходя классические методы планирования.

Недавние исследования, такие как работа Dworzanski и Hlavacs (2023) [6], предложили подход, при котором автоматическая генерация деревьев поведения осуществляется с помощью методов обучения с подкреплением. Это позволяет учить агентов формировать стратегии поведения на основе ценности состояний или действий через Q-learning, что ускоряет процесс создания ИИ-агентов. Такой подход автоматически строит дерево поведения, снижая трудозатраты на ручную настройку и обеспечивая адаптивность моделей к меняющимся игровым условиям.

Применение RL для генерации HTN позволит не только автоматизировать процесс создания иерархических структур, но и адаптировать их к различным игровым условиям. Чтобы применить этот метод для генерации HTN:

- 1) Исследуются существующие методы построения HTN и их использование в игровых проектах, рассматриваются алгоритмы обучения с подкреплением, анализируются ключевые параметры, влияющие на успешное обучение, включая структуру вознаграждений;
- 2) Разрабатывается алгоритм автоматической генерации HTN, который позволяет генерировать структуру задач из данных, полученных в процессе обучения агентов;
- 3) Создаётся игровой сценарий, в котором агенты проходят обучение, а затем их данные используются для построения HTN;
- 4) Для оценки эффективности сравниваются агенты, использующие сгенерированную HTN, с агентами, использующими традиционные методы.

Выводы. Проведен анализ возможности автоматической генерации HTN с применением методов обучения с подкреплением. Рассмотрены преимущества данного подхода, позволяющие сократить трудозатраты на разработку игровых ИИ, а также устранить необходимость ручного проектирования сложных моделей поведения.

Список использованных источников:

1. Sacerdoti E.D. The nonlinear nature of plans // Stanford Research Inst. – 1975. – 100 p.
2. Wang H.W., Qi C. Hierarchical task network planning based emergency response decision making theory and method // Beijing: Science Press. – 2015. – 200 p.
3. Milot A., Chauveau E., Lacroix S. et al. Market-based multi-robot coordination with HTN planning // Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems. – 2021. – P. 2606-2612.
4. Sun L., Zhu A.S., Li B. et al. HTN guided adversarial planning for RTS games // Proc. of the IEEE International Conference on Mechatronics and Automation. – 2020. – P. 1326-1331.
5. Hogg C., Kuter U., Muñoz-Avila H. Learning Methods to Generate Good Plans: Integrating HTN Learning and Reinforcement Learning // Twenty-Fourth AAAI Conference on Artificial Intelligence. – 2010. – № 1. – P. 1530-1535.
6. Dworzanski R., Hlavacs H. Shaping AI behavior: a Q-learning driven approach to automatic behavior tree creation [Электронный ресурс] // University of Vienna. – 2023. – Режим доступа: <https://theses.univie.ac.at/detail/68504>, свободный. – (Дата обращения: 12.02.2025).