

ПРИМЕНЕНИЕ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ИССЛЕДОВАНИЯ НАСТОЛЬНОЙ ИГРЫ HEX

Бикмурзин М.А.¹, Мазуренко В.В.¹

¹⁾ Уральский федеральный университет имени первого Президента России
Б.Н. Ельцина, г. Екатеринбург, Россия
E-mail: mikejob@bk.ru

APPLICATION OF DEEP LEARNING METHODS TO RESEARCH OF THE HEX BOARD GAME

Bikmurzin M.A.¹, Mazurenko V.V.¹

¹⁾ Ural Federal University, Yekaterinburg, Russia

Basing on deep learning techniques, we created a neural network agent and trained it to play Hex. The training took place in accordance with modified Google Deep mind training algorithm for Alpha Go Zero.

Игра Хех была придумана в середине XX века независимо друг от друга двумя людьми – датским ученым Питом Хейном и американским математиком Джоном Нэшем. Игра происходит на гексагональной доске ромбической формы. Два игрока по очереди занимают свободные клетки с целью первым построить соединяющий кластер между сторонами своего цвета. В данной игре невозможна ничья. Джон Нэш доказал, что игрок, начинающий первым, имеет форсированный выигрыш, однако в общем случае он не найден. Хех не является новинкой для научного мира. На данную тему было написано множество статей, в которых ученые применяли различные алгоритмы для исследования этой замечательной игры, как, например, H-search [2], ХН-search [3], Monte Carlo Tree Search [4] и многие другие. Однако, нейросетевые подходы для изучения стратегий игры Хех ранее не применялись, что являлось основной целью данной работы.

Для реализации была выбрана архитектура сверточной нейронной сети с двумя выходами. После подачи на вход некоторой игровой позиции на одном выходе нейронная сеть предсказывает оценку позиции от -1 до 1, а на другом - распределение вероятности по всем возможным ходам из этой же позиции. Для обучения нейросети используется алгоритм Monte Carlo Tree Search, для которого каждая позиция на доске представляется в виде узла в древе поиска. Нейросеть играет сама с собой некоторое число партий, для каждой позиции запоминая победителя и распределение дочерних узлов в древе поиска по посещениям. Функция ошибки требует совпадения этих величин с оценкой позиции и распределением вероятности, предсказанными нейросетью. Обучение нейросети осуществляется методом градиентного спуска. Этапы обучения и накопления обучающей выборки цикличны. Такой обучающий алгоритм впервые был представлен Google Deepmind для обучения нейронной сети AlphaGo Zero [1].

Мы провели модернизацию и адаптацию исходного алгоритма AlphaGo Zero к игре Нех, которая заключается в следующем. Вместо одной нейросети, мы используем две, по одной на каждого из игроков. Результат их обучения положительный. Нейронные сети успешно используют геометрические построения из теории игры Нех, такие как ромб и вилка, хотя ничего кроме правил игры не было предоставлено им перед началом обучения.

Нейросеть обучалась Нех полностью с нуля и демонстрируют очень сильный уровень игры, согласующийся с теоретическим решением [5].

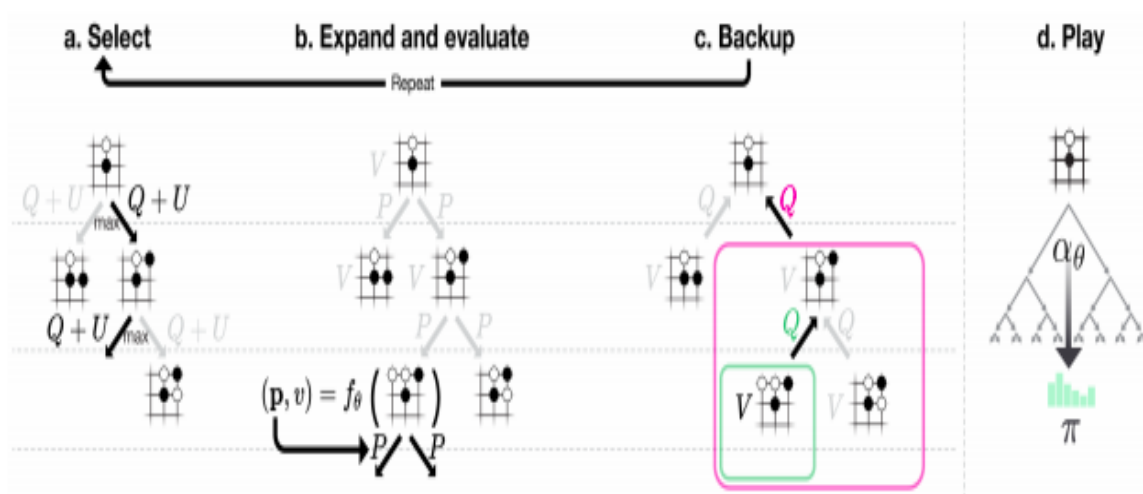


Рис. 1. Алгоритм поиска по дереву методом Монте-Карло, использующийся при обучении нейронных сетей.

1. Silver D. et al. Nature, 550, 7676, 354-359 (2017).
2. Anshelevich V. V., Artificial Intelligence, 134, 1-2, 101-120 (2002).
3. Henderson P., Arneson B., Hayward R. B. Hex, Braids, the Crossing Rule, and XH-Search. In: van den Herik H.J., Spronck P. (eds) Advances in Computer Games. ACG 2009. Lecture Notes in Computer Science, vol 6048. Springer, Berlin, Heidelberg (2010).
4. Arneson B., Hayward R. B., Henderson P. IEEE Transactions on Computational Intelligence and AI in Games, 2, 4, 251-258 (2010).
5. Hayward R. et al., Theoretical Computer Science, 349, 2, 123-139 (2005).