|  |
| --- |
| Студент группы ИВТ, Семенов Леонид |

|  |
| --- |
|  |
| **Alpha GO** |
|  |

Оглавление

[Введение 2](#_Toc502678449)

[Саммит «Будущее Го» 3](#_Toc502678450)

[Сравнение версий AlphaGo 4](#_Toc502678451)

[Значение 5](#_Toc502678452)

[Историческое значение 6](#_Toc502678453)

[Практическое применение 7](#_Toc502678454)

[Основные принципы 7](#_Toc502678455)

[Стиль игры 9](#_Toc502678456)

[Аналогичные системы 10](#_Toc502678457)

Введение

AlphaGo — программа для игры в го, разработанная компанией Google [DeepMind](https://deepmind.com/about/) в 2015 году. AlphaGo стала первой в мире программой, которая выиграла матч без гандикапа у профессионального игрока в го на стандартной доске 19×19, и эта победа ознаменовала собой важный прорыв в области искусственного интеллекта, так как большинство специалистов по искусственному интеллекту считало, что подобная программа не будет создана ранее 2020—2025 года. В марте 2016 года программа выиграла со счётом 4—1 у Ли Седоля, профессионала 9 дана (высшего ранга), во время исторического матча, широко освещавшегося в прессе. После победы в матче Корейская ассоциация падук присвоила AlphaGo «почётный 9 дан» за «искренние усилия» программы в овладении мастерством игры.

Победу AlphaGo над Ли Седолем [[1]](#footnote-1)часто сравнивают с шахматным матчем между программой Deep Blue и Гарри Каспаровым 1997 года, где победа программы, созданной IBM, над действовавшим чемпионом стала символической точкой отсчёта новой эпохи, когда компьютеры превзошли людей в шахматах. В отличие от Deep Blue, AlphaGo — это программа, которая не привязана к конкретному аппаратному обеспечению. Кроме того, AlphaGo основана на общих принципах машинного обучения и практически не использует (в отличие от шахматных программ) ни алгоритмов, ни оценочных функций, специфичных для игры в го. При разработке AlphaGo авторы использовали только самую элементарную теорию игры в го, программа достигла высокого уровня игры, обучаясь сама на партиях профессионалов. Таким образом, её методы машинного обучения могут быть использованы в других областях применения искусственного интеллекта. В частности, команда разработчиков планирует применить опыт, полученный при написании AlphaGo, для создания системы медицинской диагностики.

AlphaGo комбинирует технику, которая дала большой прорыв в силе программ для игры в го в 2007—2012 годах (метод Монте-Карло для поиска в дереве), и недавние успехи в области машинного обучения, а именно глубинное обучение с помощью многоуровневых нейронных сетей.

  
Рисунок 1 – Логотип AlphaGo

Саммит «Будущее Го»

23-27 мая 2017 года в китайском городе Вучжен проходил саммит о будущем го, где среди прочих событий AlphaGo провёл несколько показательных партий и победил во всех:

* Мини-матч из 3 партий: AlphaGo против Кэ Цзе.
* Продвинутое го: Человек+AlphaGo против Человека+AlphaGo.
* AlphaGo против объединённой команды лучших китайских профессиональных игроков в го.

Сравнение версий AlphaGo

1. AlphaGo Fan

* 176 GPUs, распределенные вычисления
* 3144 elo
* 5:0 Матч AlphaGo — Фань Хуэй

1. AlphaGo Lee

* 48 TPUs, распределенные вычисления
* 3739 elo
* 4:1 Матч AlphaGo — Ли Седоль

1. AlphaGo Master

* 4 TPUs v2, одна машина
* 4858 elo
* 60:0 против профессиональных игроков го;
* Саммит «Будущее Го»;
* 3:0 Матч AlphaGo — Кэ Цзе

1. AlphaGo Zero

* 4 TPUs v2, одна машина
* 5185
* 100:0 против AlphaGo Lee
* 89:11 против AlphaGo Master

Вероятность победы игрока A с меньшим рейтингом над игроком B с высшим рейтингом:

(1), где

D – разница рейтингов;

a – параметр, зависящий от предыдущего значения рейтинга игрока A

Рисунок 2 – Сравнение рейтинга версий AlphaGo и лучшего игрока в Го  
 на октябрь 2017 г.

Значение

## Историческое значение

До создания AlphaGo являлось одной из немногих традиционных игр, в которые человек играл лучше компьютера.

Го намного сложнее шахмат из-за большего количества возможных позиций. Во-первых, как правило, в го в каждой позиции существует больше разрешённых ходов, чем в шахматах. Во-вторых, партия в го в среднем длится больше ходов. Эти два свойства делают го очень сложной для традиционных методов искусственного интеллекта, таких как альфа-бета-отсечение[[2]](#footnote-2).

Другая сложность го по сравнению с шахматами и многими другими логическими играми заключается в создании оценочной функции, сопоставляющей произвольной позиции определённую оценку: для какой из сторон данная позиция выгоднее и насколько (либо позиция равная). В AlphaGo используются 2 [нейронные сети](https://ru.wikipedia.org/wiki/Искусственная_нейронная_сеть), одна из которых специализируется на оценке позиции. Таким образом, её можно условно назвать аналогом оценочной функции в шахматах.

По этим причинам даже спустя почти двадцать лет, после того как компьютер Deep Blue впервые выиграл у чемпиона мира Каспарова в 1997 году, самые сильные программы, играющие в го, достигли только любительского 5 дана, и не могли выиграть у профессионала на полной доске. Лучшими достижениями искусственного интеллекта в го были победы программы Zen, работающей на кластере из четырёх компьютеров, против профессионала 9 дана Масаки Такэмия два раза с форой в пять и четыре камня в 2012 году, а также победа программы Crazy Stone над Ёсио Исидой (9 профессиональный дан) с форой в четыре камня в 2013 году.

Создание программы, которая хорошо бы играла в го, являлось камнем преткновения в развитии искусственного интеллекта. До 2015 года большинство специалистов по искусственному интеллекту считали, что программа, способная победить чемпиона мира в го, не будет создана, по крайней мере, в течение следующих пяти лет, а многие из них предполагали, что на это потребуется как минимум десять лет. Даже после победы AlphaGo над Фань Хуэем многие эксперты предсказывали, что AlphaGo проиграет Ли Седолю.

После матча с Ли Седолем правительство Южной Кореи объявило об инвестиции 1 триллиона вон (863 миллиона долларов) в исследование искусственного интеллекта в течение следующих пяти лет.

**Практическое применение**

Методы, разработанные для игры в AlphaGo, могут применяться и в других областях искусственного интеллекта, в частности для медицинской диагностики. Как сообщил основатель DeepMind Демис Хассабис, DeepMind заключил соглашение с Национальной службой здравоохранения Великобритании для изучения возможности применения искусственного интеллекта для анализа медицинских данных. Для этого было основано подразделение DeepMind Health.

Основные принципы

Существенной особенностью AlphaGo является то, что она использует общие алгоритмы, которые практически не зависят от особенностей игры го. В алгоритм AlphaGo заложены только базовые принципы игры, с которых начинает изучение игры любой начинающий, такие как подсчёт числа свободных пунктов (дамэ) у группы камней, и анализ возможности захватить камни, используя приём «лестница» (ситё). Остальное AlphaGo выучила сама, анализируя базу данных из 160 тысяч партий общими методами, которые могут использоваться и в других областях искусственного интеллекта. Этим AlphaGo отличается от таких систем, как Deep Blue — шахматного суперкомпьютера. При написании алгоритма для Deep Blue использовалась продвинутая шахматная теория. Так, например, для оценки позиции Deep Blue использовал 8000 настраиваемых признаков, в Deep Blue была заложена огромная библиотека дебютов. AlphaGo — программа, которая может работать на любом аппаратном обеспечении (хотя она использовала процессор, разработанный специально для машинного обучения), и принципы которой можно использовать в других областях. Она не использует библиотеку дебютов и во время игры не обращается ни к какой базе данных ходов непосредственно (а только использует нейронную сеть, натренированную на большом количестве партий). Сами разработчики программы не знают, как AlphaGo оценивает позицию, её ходы представляют собой феномен эмерджентности. Кроме того, Deep Blue был разработан с одной целью — выиграть матч у Гарри Каспарова в шахматы. AlphaGo не была обучена играть специально против Ли Седоля, более того, в базе данных тренировки AlphaGo не было партий Ли Седоля, и во время матча программа AlphaGo не менялась и не подстраивалась под стиль игры Ли Седоля.

Главное новшество AlphaGo заключается в использовании глубинного обучения — метода, успешно применявшегося для распознавания образов (например, для поиска картинок в Google Images). А именно, AlphaGo использует нейронные сети, обученные на большом количестве профессиональных партий, способные предсказывать, какой ход сделает профессионал в той или иной позиции. Сыгранные сами с собой партии дополнительно улучшили нейронные сети. Одно лишь это позволило AlphaGo играть на уровне лучших компьютерных программ 2015 года: без перебора вариантов, используя лишь нейронную сеть для выбора хода, программа смогла выиграть у Pachi (программы, играющей на уровне второго любительского дана).

Как и лучшие предшествующие программы для игры в го, AlphaGo использует также [метод Монте-Карло](https://ru.wikipedia.org/wiki/Метод_Монте-Карло) для поиска в дереве (MCST). Суть этого метода (названного по аналогии с методом Монте-Карло в вычислительной математике) состоит в том, что для оценки позиции программа много раз играет сама с собой случайным образом, доводя каждый раз игру до конца. Потом выбирает ход, при котором было наибольшее число побед.

Стиль игры

Тоби Маннинг, судья в матче AlphaGo с Фань Хуэем, охарактеризовал стиль игры AlphaGo скорее как консервативный, чем агрессивный. По словам генерального директора DeepMind Демиса Хассабиса, AlphaGo пытается увеличить вероятность выигрыша, а не разрыв в количестве очков.

Профессиональные игроки отмечали, что некоторые ходы AlphaGo, первоначально кажущиеся ошибочными, при более глубоком изучении оказывались стратегически важными.

Не Вэйпин (9 профессиональный дан, Китай) после первой игры с Ли Седолем предположил, что AlphaGo играет на уровне 6 или 7 профессионального дана в фусэки и 13—15 дана в тюбане.

В третьей партии выяснилось, что AlphaGo способна управлять ситуацией при ведении ко-борьбы, что ранее считалось значительной слабостью большей части программ, играющих в го.

Некоторые обозреватели охарактеризовали ошибки AlphaGo, приведшие к поражению в четвёртой партии, как типичные ошибки для программы, работающей на основе метода Монте-Карло. Демис Хассабис заявил, что эти ошибки будут тщательно проанализированы, и что, видимо, AlphaGo не знает некоторые классические тэсудзи и совершает тактические ошибки. Впоследствии Айа Хуань (один из разработчиков AlphaGo, который ставил камни за AlphaGo) сказал, что у команды авторов два предположения о причинах этих ошибок: либо AlphaGo просто не хватило глубины просмотра, чтобы проанализировать ситуацию; либо 78-й ход Ли Седоля был настолько необычным (комментаторы его называли «божественным»), что программа при самообучении не встречала подобных ситуаций, в результате её оценочная функция (см. выше) в этой ситуации оказалась слаба. В любом случае, по словам Айа Хуаня, новая версия AlphaGo в этой ситуации уже играет правильно. При этом команда не делала ничего специального, AlphaGo сама в результате обучения перестала делать такие ошибки

После матча Ли Седоль заявил, что был побеждён психологически, но вовсе не технически. Программа показала способность к креативным решениям, что удивило многих игроков (например, ход № 37 во второй партии); некоторые ходы противоречили классической теории го, но в матче доказали свою эффективность, некоторые профессионалы стали использовать эти находки в своих партиях.



Рисунок 3 – Кэ Цзе после осознания проигрыша AlphaGo

Аналогичные системы

Facebook также разрабатывает программу для игры в го, Darkforest, которая тоже основана на машинном обучении и поиске в дереве. На начало 2016 года Darkforest показал сильную игру против других компьютеров, но не смог выиграть у профессионала. По силе Darkforest оценивается на уровне программ Crazy Stone и Zen.

1 марта 2016 года разработчики программы Zen (Ёдзи Одзима и Хидэки Като), компания DWANGO и исследовательская группа глубинного обучения Токийского университета (которая создала программу Ponanza для игры в сёги, победившую человека) объявили о совместном проекте «Deep Zen Go Project», с целью победить AlphaGo в течение 6—12 месяцев. Японская ассоциация го обещала поддержать проект. В ноябре 2016 года Deep Zen Go проиграла со счётом 2-1 самому титулованному игроку Японии Тё Тикуну.

Таблица 1 – Configuration and performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Configuration | Search threads | No. of CPU | No. of GPU | Elo rating |
| Single | 40 | 48 | 1 | 2181 |
| Single | 40 | 48 | 2 | 2738 |
| Single | 40 | 48 | 4 | 2850 |
| Single | 40 | 48 | 8 | 2890 |
| Distributed | 12 | 428 | 64 | 2937 |
| Distributed | 24 | 764 | 112 | 3079 |
| Distributed | 40 | 1202 | 176 | 3140 |
| Distributed | 64 | 1920 | 280 | 3168 |

1. Корейский го-профессионал 9 дана, победитель в командной дисциплине по го на летних Азиатских играх 2010. [↑](#footnote-ref-1)
2. Алгоритм поиска, стремящийся сократить количество узлов, оцениваемых в дереве поиска алгоритмом минимакса. [↑](#footnote-ref-2)