

제2장 알파고

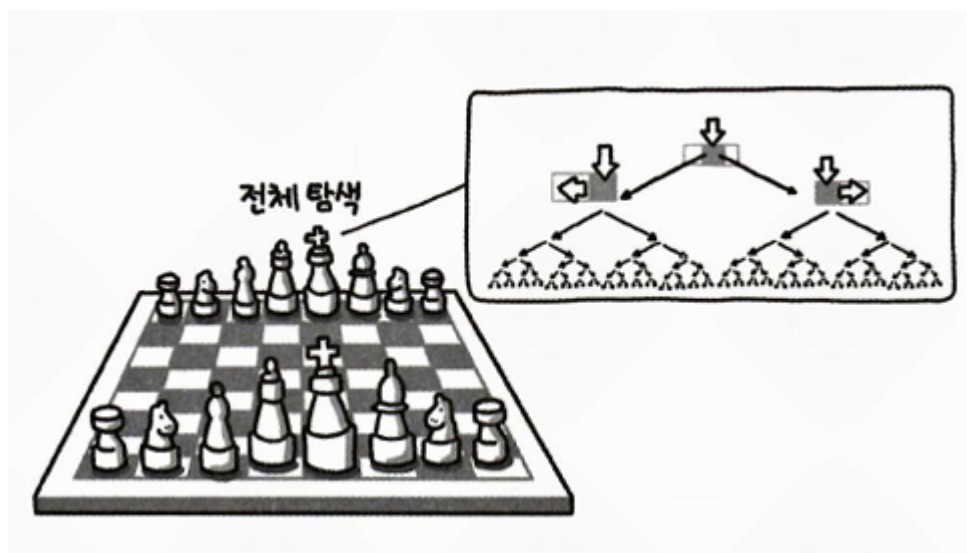
2.1 딥 블루는 어떻게 체스 챔피언이 되었을까?

- 1997년 인간 대표 카스파로프와 컴퓨터 대표 딥 블루(IBM의 슈퍼 컴퓨터)의 체스 대결에서 딥 블루가 승리하였는데 컴퓨터는 어떻게 인간을 능가했을까?

딥 블루는 컴퓨터의 강력한 계산 능력에 의존해 모든 가능성을 탐색합니다. 초당 2억 번을 계산할 수 있었고 그래서 모든 계산을 20초 만에 끝낼 수 있었다. 또한 딥 블루 곁에는 카스파로프의 약점을 가르칠 수 있는 여러 체

스 전문가를 두었습니다. 카스파로프 공략법을 엔지니어들이 규칙으로 입력하기도 했다.

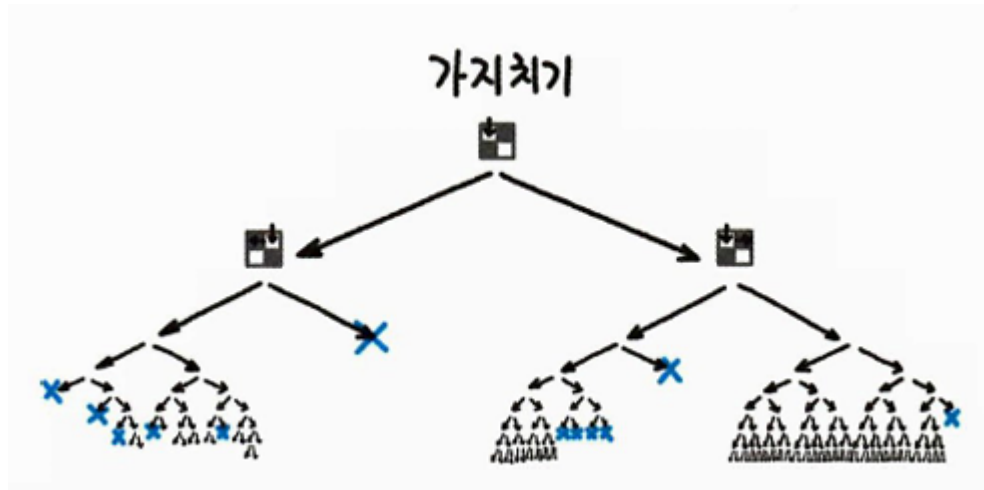
체스에서 경우의 수를 탐색하는 과정을 살펴보면 경우의 수가 확장되는 모습은 마치 나무가 뿌리에서 출발해 가지로 이어지는 모습과 비슷한데 이런 구조를 영어로 트리(Tree)라고 표현한다



또한 이처럼 게임에서 가능한 모든 경우의 수를 트리 형태로 나타낸 것을 게임 트리(Game Tree)라고 부른다

딥 블루는 엄청난 계산 능력으로 가능한 수 대부분을 계산해냈지만, 게임 트리 전체를 탐색할 수는 없다

모든 경로를 탐색하는 것을 완전 탐색(Exhaustive Search)라고 하고 좀 더 효율적으로 탐색하기 위해, 한 번 탐색해보고 성과가 없다면 그쪽은 더 이상 탐색하지 않도록 표시해두는 것을 컴퓨터 과학에서는 가지치기(Pruning)라고 한다



2.2 바둑 필요한 모든 수를 계산 할 수 있을까?

- 가로세로 19줄, 총 361개의 점으로 이뤄진 바둑판에서 가능한 수를 계산해보면 361팩토리얼($361!$) = 10의 768승 채스가 10의 120승 이니 엄청난 차이다. 실제로 유효한 값만 대상으로 정교하게 계산한 결과, 바둑의 게임 트리 크기는 10의 360승 이지만 그래도 엄청난 차이다

항목	크기
우주의 원자 수	10^{80}
체스의 게임 트리 크기	10^{120}
바둑의 게임 트리 크기	10^{360}

2.3 인간을 대표하는 천재 기사 이세돌

- 이세돌은 열일곱 살의 어린 나이에 창의적인 묘수로 바둑의 고정관념을 깨트려나갔고, 무려 32연승이라는 대기록을 세우며 '불패소년'이라는 명성까지 얻었다

2.4 도박의 확률을 이용하는 몬테카를로 방법

-인공지능으로 두는 바둑은 1960년대부터 꾸준히 시도되어 왔으나 몬테카를로 방법Monte Carlo Method을 도입하면서 비로소 본격적으로 가능성이 보이기 시작한다.

몬테카를로 방법이란 계속 도전하면 결국은 원래 확률 만큼의 결과를 얻을 수 있는데, 이런 식으로 값을 계산하는 알고리즘을 몬테카를로 방법이라고 한다.

바둑 인공지능이 가장 풀기 어려운 문제 중 하나는 현재 국면이 어떤 상태인지 한눈에 알아 내기가 어렵다는 것인데, 몬테카를로 방법은 이를 확률적인 방법으로 해결한다.

어차피 어느 한쪽이 유리한 입장에 있다면 그다음부터는 무작위로 바둑을 둔다고 해도, 유리한 쪽이 결국 이길 가능성이 높다고 여긴다는 것이다.

2006년부터는 게임 트리 탐색에 몬테카를로 방법을 접목한 몬테카를로 트리 탐색(Monte Carlo Tree Search)이 라는 알고리즘을 고안해 바둑에 적용했고, 이때부터 바둑 인공지능의 실력이 급상승한다.

2.5 정책망, 어디에 돌을 내려놓을까?

- 정책망은 사람이 만든 기보(바둑의 순수를 기록한 문서나 기록)를 이용해 학습합니다. 학습에 사용한 데이터는 KGS라는 바둑 사이트에서 가져왔으며, 6단 이상인 고수의 기보만 사용한다

알파고의 전체적인 수준은 5단 정도에 불과했다. 6단 이상의 데이터로 학습했는데 최종 결과물은 오히려 5단인 이유는 바둑 기사들의 기풍 때문이다. 기풍이란 바둑을 두는 스타일이라고 한다.

그래서 알파고는 이를 보완할 수 있는 다양한 방법을 찾았다. 우선, 각각의 상황에 따라 정책망 3가지를 만들었다.

첫째 망은 바로 앞서 살펴본 사람의 기보를 이용해 학습한 정책망(기보학습 정책망)

둘째 망은 롤아웃 정책망이다 기보학습 정책망과 비슷하지만 훨씬 작고 가벼운 망입니다 성능은 떨어지지만 계산 속도가 빠르기 때문에 탐색을 빠르게 진행할 때 많은 도움이 된다

마지막 셋째 스스로 대국하며 강화학습을 수행한 정책망(강화학습 정책망)이다.

이 망은 알파고끼리 대국을 치르면서 스스로 실력을 향상시킨 망이다

경기에서 이기면 승리한 측의 모든 수의 점수가 더욱 강화되는 형태입니다. 그래서 강화학습(Reinforcement Learning)으로 부르는 것이다

	정책망	특징
1	사람의 기보를 이용해 학습한 정책망(기보학습 정책망)	57% 정확도. 5단 수준
2	롤아웃 정책망	24% 정확도. 1,500배 빠름
3	스스로 대국하며 강화학습한 정책망(강화학습 정책망)	기보학습 정책망과 대전하면 80% 확률로 승리

2.6 가치망, 형세를 판단하다

- 가치망은 앞서 살펴본 정책망과 구조가 조금 다르고 훨씬 더 단순하다. 가치망은 현재 국면에서 승패 여부를 예측하는 망이다. 가치망은 오로지 승리 할 가능성만을 계산한다.

따라서 알파고는 고수의 기보로 지도 학습을 진행한 다음, 어디에 돌을 내려놓을지 판단하는 정책망을 만들고, 그리고 정책망을 이용해 스스로 대국을 두어 강화학습을 진행한 다음, 승리할 확률이 얼마나 되는지 알아내는 가치망을 만들어낸다. 여기까지가 바로 알파고의 학습 과정이다.

2.7 알파고가 수를 두는 방법

-이제 알파고가 학습한 것을 바탕으로 어떤 과정을 거쳐 착점을 결정하는지 봐보자

몬테카를로 트리 탐색을 활용하는데, 몬테카를로 방법에도 엄연히 한계는 있다. 이른바 '묘수'를 간과할 수

있다는 점이다. 무작위로 시뮬레이션하는 몬테카를로 방법은 아주 작은 확률인 상대의 묘수를 놓칠 수도 있었다.

1. 어떤 수에서 시작할지 승리 할 가능성 이 높아 보이는 수를 선택 합니다.
2. 기보학습 정책망을 이용해 다음 수를 어디에 둘지 확장합니다.
3. 롤아웃 정책망을 이용해 끝까지 시뮬레이션 합니다.
4. 가치망의 점수를 50% 반영한 승패 여부를 모든 수 에 업데이트합니다.
5. 1~4의 과정을 계속해서 반복한 후 최종적으로 가장 많이 진행한 수를 다음 수로 결정 합니다.

2.8 신의 한수

-알파고도 '신의 한 수'인 78수를 막아내진 못했다. 왜 신의 한 수를 허 용하고 말았을까?

알파고의 몬테카를로 트리 탐색은 유망한 수를 중심으로 꼼꼼하게 탐색해나간다고 했습니다.

그러니까 확률이 높은 쪽을 향해 더 많이 더 깊게 탐색해 나가고 가장 신뢰가 높은 지점에 착수를 하는 원리다. 하지만 이세돌이 둔 신의 한 수 지점에 착수할 확률을 알파고는 1 만 분의 1로 매우 낮게 예측 했고, 알파고는 설마 그 지점에 둘 줄은 몰랐기에, 충분히 탐색하지도 않았다. 당연히 그 지점이 묘수인지 아닌지조차 알아내지 못하고, 결국 78수 다음에 대국이 어떻게 흘러갈지를 알파고는 전혀 알아차리지 못하고 패배하게 된다.

2.9 알파고 제로, 인간은 필요 없다

-구글 딥마인드는 계속해서 알파고를 개선했고, 그렇게 알파고 제로 (AlphaGo Zero) 가 등장한다

완전히 무(無)에서 시작한다고 하여 알파고 '제로'라는 이름을 부여했다. 알파고 제로는 매일 100만 대국을 휴식 없이 치를 수 있다. 결국 알파고 제로가 공개되기 전에는 기존 알파고보다 좀 더 성능이 좋은 알파고 마스터보다 훨씬 더 강력하다고 한다. 이제 인간은 더는 바둑에서 컴퓨터를 이길 수 없게 됐다.

2.10 알파제로, 진정한 인공지능을 향하여

-알파고 제로 이후에 딥마인드는 2018년 알파제로(AlphaZero)를 공개했다. 알파제로도 원리는 비슷 하지만 이름에 바둑의 영문 명칭인 고(GO)가 빠진 것에서 알 수 있듯 알파제로는 이제 바둑뿐 만 아니라 체스, 일본 장기까지 게임 영역을 넓혔다.

그렇다면 이제 컴퓨터는 인간을 능가하는 지능을 갖게 되었을까?

알파고가 인간을 능가했다고 해서 이제 인공지능이 모든 면에서 인간을 능가한다고 할 수는 없다

이미 컴퓨터는 인간을 능가한다. 원래부터 컴퓨터는 인간보다 훨씬 더 빠르고 정확하게 계산했기 때문이다

인간을 능가하는 기계를 두고 "기술을 있는 그대로 바라 보라"고 주문했다.

"지나치게 기술에 의존해서도 안 되며 그렇다고 기술에 공포를 느껴서도 안 된다"면서,

인간을 능가하는 인공지능의 등장은 여전히 먼 이야기이다. 전혀 두려워할 필요가 없다. 무엇보다 컴퓨터에는 인간의 지혜와 능력이 담겨져 있다. 새로운 기술이 등장하더라도 인간의 그림자는 언제나 함께할 것 이다.