

# AlphaGo 和深蓝在算法上有什么不同

1997 年 IBM 公司研制的深蓝首次在正式比赛中战胜了曾经十次获得世界冠军的国际象棋大师卡斯帕罗夫，时隔近 20 年后，DeepMind 公司研制的 AlphaGo 在 2016 年首次战胜了多次获得世界冠军的韩国围棋棋手李世石九段，2017 年又战胜了围棋第一人，中国的柯洁九段，轰动了世界。同样是下棋，AlphaGo 和深蓝在算法上有哪些不同呢？



图 1： 1997 年 IBM 深蓝对战卡斯帕罗夫

我们先看看人是如何下棋的。人在下棋时会根据当前局面考虑几种可能的走法，再对每种可能的走法考虑对方会如何走，再考虑自己会如何应对……高手会这样往前看很多步，最后根据每种走法带来的结局判断如何走棋。这一策略不仅考虑了自己的走法，同时也考虑对方的应对，然后选择那种即便对方能够正确应对，对自己依然最有利的，或者损失最小的那种走法。总结来说，这一策略里包括两个关键点：一是“搜索”，即往前多看几步棋，二是“判断”，即评估多步走棋后形成的局面对谁有利。

深蓝基本是模仿这种搜索+评估的走棋方式。一般来说，搜索的越深（往前看的步数越多），走棋的质量越高。然而，因为每步走棋都有很多种可能性，因此每增加一步搜索都要会显著增加计算量，从而导致“组合爆炸”问题。据深蓝开发者估算，如果每一步搜索都考虑所有可能的走法，即便只往前考虑十步，每走一步棋深蓝也需要“思考”17 年。事实上，人在下棋时并不会（也不可能）考虑所有可能的走法，而是根据自己的经验考虑几种典型的走法去尝试。为模拟人的这种选择性思考能力，深蓝采用了一种称为 $\alpha$ - $\beta$ 剪枝的搜索算法，利用已有的搜索结果，剪掉一些不必要的候选走法，从而达到快速“思考”的目的。在“判断”过程中，深蓝采用的是基于知识的方法，利用国际象棋大师总结的知识给局面打分，分数的大小代表局面

对“我方”的有利程度。



图 2:2016 年 AlphaGo 对战李世石

这种搜索+评估的方式也可以用在围棋上,事实上很多研究者也是这样做的,但直到 AlphaGo 之前,这些尝试都不是很成功。为什么呢?很多人总结为围棋的状态多,搜索空间大。状态多虽然是个问题,但不是决定性的原因,最主要的原因是围棋对局面的判断不是那么容易,而局面判断是否准确对 $\alpha$ - $\beta$ 剪枝方法来说至关重要。

为解决这一问题,AlphaGo 采了一种称为“蒙特卡洛树搜索”的新方法。这一方法的思想很简单,在当前棋局下,随机地模拟双方走步,直到分出胜负为止。通过多次模拟,即可计算出每个可能的落子点的获胜概率,以此作为对局面的判断。利用随机模拟来计算获胜概率,从而实现棋局评估,这是 AlphaGo 能够取得成功很关键的一点。然而,这种评估方式依赖大量走棋模拟,且每次模拟都要走到完局,效率还是非常低的。为了解决这个问题,AlphaGo 利用深度学习方法,通过学习专业棋手以往的棋谱,对每个可能的落子点预测走棋概率。在模拟时,只选择那些概率比较高的落子点,从而极大地提高了随机模拟的针对性,提高了计算效率。

从以上介绍可以看出,深蓝和 AlphaGo 采用了很不同的算法,深蓝依靠的是国际象棋大师总结的知识和 $\alpha$ - $\beta$ 剪枝算法战胜了卡斯帕罗夫,AlphaGo 利用的是专业围棋棋手的下棋数据、深度学习算法和蒙特卡洛树搜索算法,战胜了李世石和柯洁。后来 AlphaGo 又进一步升级为 AlphaGoZero。AlphaZero 完全摒弃了人类棋局,仅依靠自我对弈产生的数据,锻炼出了和人类走棋方式完全无关的 AI 棋艺。