



AlphaGo：不懂圍棋的棋王

「廣、準、深」的機器棋王養成之路。

近日來「人工智慧」再度成為炙手可熱的議題，主要因為網路巨擘 Google 發展出來的程式 AlphaGo，在圍棋比賽中以 4：1 打敗了實力九段的南韓棋王李世乭。

要知道 AlphaGo 為什麼可以贏過人類，首先得剖析它下棋的方式：與人類棋手類似，一個優秀的電腦弈棋程式需要具備「廣、準、深」三大要素。「廣度」意思是多方思考許多不同棋步的可能性，而非局限於少數棋步；「準度」是指對於每一步的優劣勢都能夠判斷得很精準；「深度」則意指能夠預估後續的很多步，模擬對手的選擇以及自己相對應的下一步、下下步等。

那麼，AlphaGo 是如何訓練出「廣、準、深」三大要素呢？它的訓練模式跟一般人類棋手截然不同。大致來說可分為三個階段，我各稱為「天下棋手為我師」、「超越過去的自己」、「機率算盡太聰明」。

在廣度方面，AlphaGo 爬梳了 KGS 圍棋伺服器資料庫裡許多圍棋高手走過的共約 3000 萬棋步，學習他們面對各種盤面時會走的「下一步」。在這個階段，電腦會碰到很嚴峻的挑戰：它所看到的 3000 萬棋步，相對於 250^{150} 的總盤面數是微乎其微的，如何利用相對如此少的資料學到「廣度」，一直是電腦弈棋領域裡非常難解的問題。所幸近年來人工智慧學家發展出成熟的「深度學習」技術，能夠從盤面產生一些「抽象化」的隱藏特徵來增加學習的精確度（參見 2015 年 1 月號〈深度學習的深度〉）。

然而，學習高手的棋步並不能保證勝利。僅完成第一階段的電腦，連取勝二段的棋手都不易。AlphaGo 第二階段的重點是「自我訓練」，在這個階段，它會用第一階段學到的下棋法跟自己對弈，然後，藉由對弈最後的勝負，修正第一階段產生的分數。也就是勝利的 AlphaGo 程式所經歷的每個盤面，其對應下法的分數會上升；相對地，失敗的 AlphaGo 走過的盤面，分數就會下降。

AlphaGo 的第三階段，在預估每個盤面的「價值」。所

謂盤面的價值，就是從這個盤面開始，假設對手下了勝率最高的一步時，AlphaGo 的獲勝機率有多高。當然我們無法知道無懈可擊的對手會怎麼下，所以退而求其次，AlphaGo 模擬跟自己下的過程，來決定每個盤面的勝率有多少，然後用有限的盤面與勝率當成訓練資料，去訓練深度學習的模型，來估計所有盤面的價值。至此，AlphaGo 已經具備了「準度」。

真正跟對手下棋時，AlphaGo 沉浸在自己的世界裡。每當對手下完一步，它就開始模擬接下來可能的數步（模擬的時候就是假設對手是自己，所以對手是誰，對 AlphaGo 而言完全不重要），在模擬的過程中，重新對接下來可能面對的棋局估算勝率，最後選擇最穩定、最有可能獲勝的棋步，這是 AlphaGo 棋風深度的展現。

了解 AlphaGo 下棋的原理，我們才能進一步討論「人工智慧」與「人類智慧」的差異。從頭到尾，AlphaGo 完全沒有圍棋的知識（例如優勢、虧損、打劫等）。對它而言，下棋就是在預估每一手獲勝的機率，而這個機率並非根據圍棋邏輯與智慧得來，而是統計分析曾經下過的數千萬種棋步的結果。AlphaGo 的勝利或失敗，不在於它對圍棋的知識與經驗，完全取決於這些機率的估計是否準確。

人工智慧解決問題的方式與人類的智慧大相逕庭。除了不懂圍棋的人工智慧棋手 AlphaGo 之外，Google 也創造了像「Google 翻譯」這樣可以翻譯超過 100 種語言、卻不懂文法的翻譯家。用類似的概念，機器翻譯的過程就如同尋找機率最高的一些「符號」組合，於是電腦不需要太了解人類的「文法」就可以從事翻譯，但是因為翻譯所需的搜尋空間比圍棋更大，目前電腦的表現仍無法達到專業水準。未來，可以想見另一種模式的「智慧」將與人類的智慧並存，如何善用彼此的優點、補足彼此的不足，將是個重要的課題。

林守德是台灣大學資訊工程系教授。

SA

影像來源：林守德