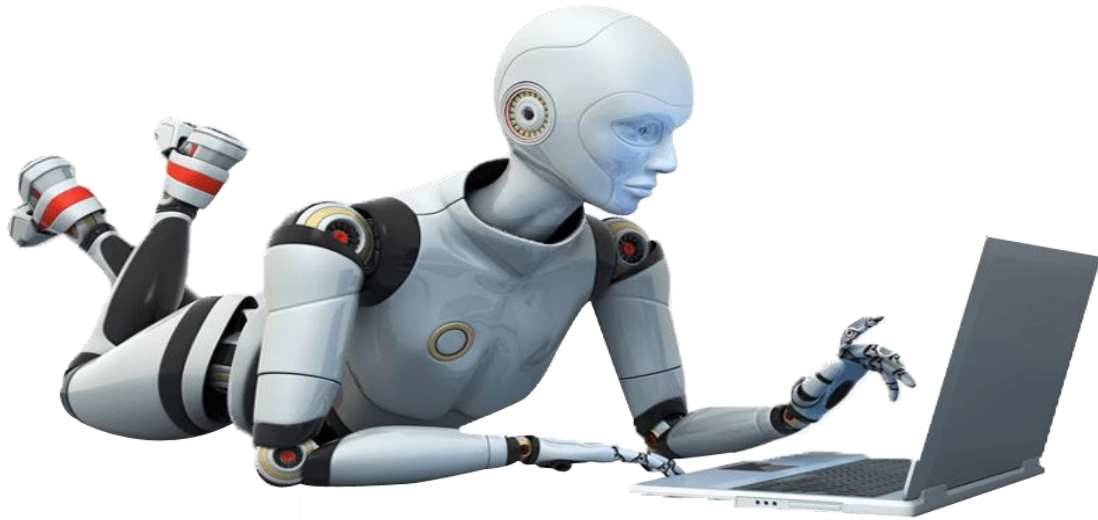


# 【新興領域：5月焦點 8】行雲流水的操作，只有你辦的到：AI 打電動

關鍵字：AI、人工智慧、Artificial Intelligence

日期：2019.5.15



被譽為電腦之父的 Alan Mathison Turing 在 1950 年提出著名的「Turing test」，該測驗基於機器是否可以像人一樣思考？是否表現的與人類沒有差異。近年來我們看到了 AlphaGo 擊敗人類、電競選手輸給 AI，這都指出人工智慧的定義在現今已經不同了，人工智慧已經融入我們日常生活當中，甚至在某方面表現遠遠超過人類。

## 一、前言

從電腦發明以來，就替人類解決不少問題。1852 年有位英國數學家思索著一個問題：「在一幅平面的地圖上，只要兩國有共同的邊界就要塗不同顏色，是否最少需要四種顏色，才能塗任何地圖？」，也就是鼎鼎大名的四色問題，在困惑了數學家超過一百年之後，在 1976 年電腦的計算輔助之下，這個問題被證明出來了。

由於高速的運算能力，電腦也被用來尋找質數。「Great Internet Mersenne Prime Search」是一個由一群志願者發起的合作計畫，目前找到最大的質數是在 2018 年 12 月發現的  $2$  的  $82,589,933$  次方  $-1$ ，共有 24,862,048 位數。或許有人會覺得無聊找什麼質數，但質數在密碼學應用廣泛，由於其無法輕易找到因數的特性，而適合當作密碼。(肯定比 `ji32k7au4a83` 這組密碼靠譜)

在近年由於人工智慧的強大，甚至直接挑戰人類的智慧了，在棋類、電競領域上，人類都不是 AI 的對手了，這對於人們來說是一大威脅還是利多？只能說在科技演進的過程中，質疑的聲音不曾消失過。但毋庸置疑的是，人類的生活福祉始終受惠於科技的進步，技術革新的益處是不

可忽視的。

## 二、IBM 深藍對戰世界西洋棋王



圖片來源：維基百科

圖 8-1、IBM 深藍與西洋棋王

1996 年 2 月 10 日，IBM 開發的超級電腦深藍(Deep Blue)，首次對決西洋棋世界冠軍卡斯帕洛夫(Garry Kasparov)，雖然起初以 2 比 4 落敗。但後來經過研究小組的改良後，並於 1997 年 5 月再度挑戰棋王，最終以 3.5 比 2.5 成績擊敗人類選手，從此聲名大噪。

在這當中也有不少爭議，如深藍輸入了過去卡斯帕洛夫所有的公開棋賽資訊，但關於深藍的棋賽資訊就是沒有的，使得雙方在一開始的資訊量就不對等。另外由於西洋棋很考驗體力與意志力，而電腦對比人類就沒有疲累問題，因此也被質疑其公平性。

那麼深藍的歷史最早可以追溯到哪時候呢？1985 年，畢業於台大電機的許峰雄在美國卡內基美隆大學攻讀博士，並在此時研究的 ChipTest 就已經能計算出棋路，雖然此階段還不足以挑戰人類棋手，但也因為該研究成果獲得了美國電腦協會的獎章以及資金，才得以繼續投入研發，此時許峰雄也期許有朝一日，能開發出可擊敗世界西洋棋手的電腦機器。

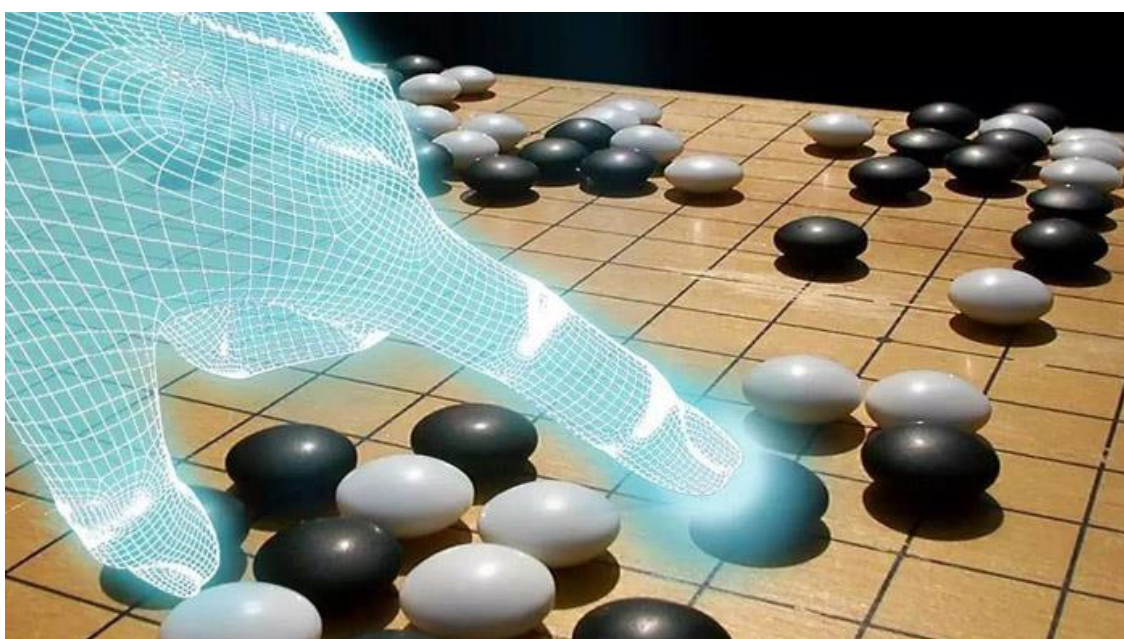
到了 1988 年，許峰雄與其團隊開發出比上一代更具智慧的深思(Deep Thought)，每秒鐘已經能分析出 70 萬步棋、能夠推算出 10 步後結果，此時的戰力僅略輸西洋棋大師了。也因為其出色

的運算能力，此時深思也吸引了 IBM 的注意，也因此邀請剛獲得博士學位的許峰雄加入 IBM 研究團隊。

此時的深思實力也備受肯定，對戰丹麥西洋棋高手本特拉爾森(Bent Larsen)並取得勝利，但是很可惜的是，在 1989 年 10 月首次對上當時的西洋棋世界冠軍卡斯帕洛夫(也就是之後由深藍所擊敗的對手)落得慘敗，於是許峰雄博士與團隊們閉關改進，想打造出更厲害的超級電腦。

於是在團隊廢寢忘食的開發下，1995 年因應企業形象也將深思改名為深藍(Deep Blue)，此時的深藍不僅在計算能力上強化，每秒鐘能夠分析 2 億步棋、推算出 12 步後結果，更請到四位西洋棋特級大師為深藍做「特訓」，不斷進行修正。在 1996 年雖然再度敗給棋王，但在 1997 年的那一戰之後，就讓它從此成名了。

### 三、人工智慧近一步挑戰下棋



圖片來源：<https://pse.is/H4YRE>

圖 8-2、人工智慧下圍棋

在 AlphaGo 擊敗李世石後，關於人工智慧的討論熱度沒有間斷過，網路上也充斥著相關 AlphaGo 的討論，因此此文僅針對最新一代、最強的「AlphaZero」作介紹。

AlphaZero 有多厲害？相較於歷代的 AlphaGo，AlphaZero 在沒有進行太大的改變之下，使用相同的演算法、類神經網路結構，得到了良好的相容性且能夠順利運行在將棋、西洋棋上頭，且經過訓練後，2 小時擊敗最強將棋 AI、4 小時擊敗最強國際象棋 AI、8 小時擊敗最強圍棋 AI(當初擊敗李世石的 AlphaGo)。

不同於傳統的電腦下棋方法，AlphaZero 透過一種叫強化學習(reinforcement learning)的類神經網路不斷地自行進行博弈對戰，並且基於試誤法(Trial and Error)來不斷修正參數而強化，這項技術現在也普遍用在電競上(如 AlphaStar、OpenAI 的 Dota2)。其所訓練的神經網路選用的為蒙特卡羅樹搜尋(Monte-Carlo Tree Search)，以便在每局採取最好的行為。

AlphaZero 的下棋風格不像歷代一樣保守求勝，棋風大膽、激進且富有創意，能夠替人類在棋類上開拓視野，作為日後訓練各種棋路的好夥伴。例如在將棋中，把「王將」移到棋局中心似乎違背人類直覺，使 AlphaZero 陷入危險當中，但最終遊戲節奏仍在其掌握中，依舊能夠以獨樹一格的方式取得勝利。

由人工智慧下棋似乎替人類開啟未知的領域，讓我們知道人們還是存在著諸多進步的空間。AlphaZero 除了展示其在靜態資訊完全賽局下強大的應用能力，也在其他領域如蛋白質結構預測的 AlphaFold 等發揮功效，在未來科學研究的演進上將功不可沒。

#### 四、人工智慧打電動囉-Dota2



圖片來源：OpenAI

圖 8-3、人工智慧玩 Dota2

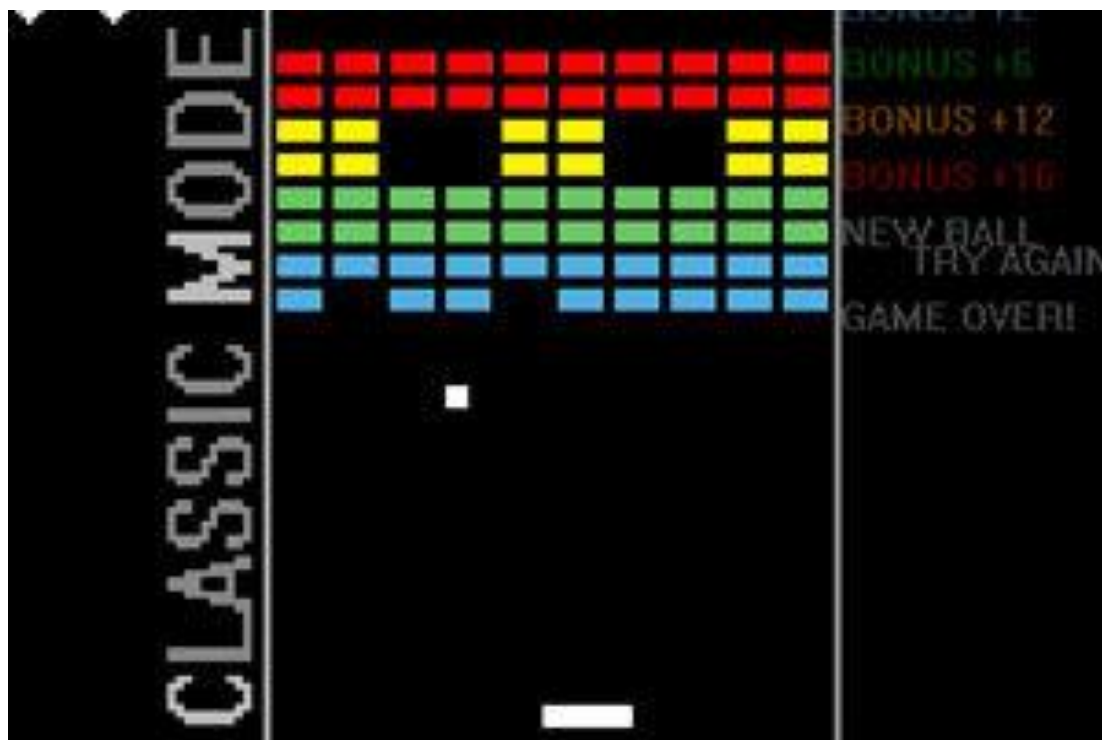
在聊聊 AI 打電動之前，我們必須先有個觀念。不同於下棋屬於靜態資訊完全賽局，也就是在下棋時雙方掌握同樣的檯面資訊，且是你下一步棋子我動一步棋子。不像是電競是屬於動態資訊不完全賽局的，有打過電動的人都知道，每個細微的操作都會影響到你之後所面對的局勢，雙方所掌握的遊戲視野也不相同，因此電競是高難度的遊戲(必須分配資源、部隊調控、走位等)。

首先簡單介紹一下 Dota2 遊戲，Dota2 與英雄聯盟(League of Legends, LOL)屬於同性質的鬥塔遊戲，都是由拆除對方主堡贏得勝利。一般而言為 5vs5 的對打情況，遊戲當中涉及最大化金錢資源、購買裝備、團隊合作、閃躲操作等技巧。

在 2017 年，由馬斯克所支持的 OpenAI 開發出能夠 1vs1 單挑的 AI，對決頂級電子競技選手 Dendi 並贏得勝利，但此階段開發的 AI 還無法有效地進行 5vs5 團隊合作對上人類隊伍。那麼此時的 AI 是什麼運作的呢？

由於 1vs1 的 AI 僅會玩特定的角色，但由於 dota2 遊戲的進行一般都是 5vs5，所以 1vs1 又要是特定的角色的對戰數據難以取得(而且還要是厲害玩家所玩的)，因此訓練 AI 時不像圍棋一

樣能夠有大量的對戰數據參考。那麼在缺少訓練資料的情況下，仰賴的方法是自我博弈「self-play」的強化學習，也就是 AI 不斷與自己進行對戰，從中修正而進化，以下圖較簡易的例子說明。



圖片來源：Breakout 官方

圖 8-4、打磚塊遊戲

以大家小時候都玩過的打磚塊為例，在強化學習中要訓練 AI 玩遊戲的話，就是不斷不斷地讓它一直試，想辦法找出能夠達到最高分數的策略。打磚塊遊戲玩家能操作的，僅有往左或往右移動板子，最終希望能消除越多磚塊越好。一開始 AI 會隨機的操控方向左或右，之後判定採取某個策略(移動到哪)的話，所得到的報酬最高，並且在確定哪個策略能夠使總報酬最高後，就鎖定此策略，再不斷地重複迭代上述流程。其實最簡單的就是，把它想成具有記憶性、能夠改善且不斷變強的「無限猴子」。

但是在 Dota2 電玩中，有些非常基本的動作(如補刀、閃躲)，如果藉由上述流程學習的話(也就是先隨機操控)會非常慢，得藉由一連串的巧合才學習到。比較快的方式是，一開始就引入「外界先驗知識」也就是你小時候所對戰的那些機械式 AI，這些機械式的操作都是被事先下好指令的，遇到什麼情境時觸發某個動作，將這些已經事先寫好「劇本」的機器人，結合強化學習能夠不斷自我強化的人工智慧，將會得到一個擁有基礎電玩底子且能夠透過訓練不斷增強的 AI。

上述的討論還是在 1vs1 的情況下，進展到 5vs5 的時候，此時個別 AI 需要能夠學會團隊合作，以團隊總報酬作為衡量標準，也使得變數更多、更複雜。2018 年 OpenAI Five 開發出來，已有職業級水準，更在今年四月開放全球玩家與其對抗，人間一天相當於 OpenAI Five 訓練 250 年，此時的 AI 已經有經過 45,000 年的訓練成果了。想當然，最終在 7,215 場比賽中僅輸掉 42 場，勝率高達 99.4%。由於遊戲的複雜度，因此十分適合作為 AI 研究的基礎，在未來應用至現實世界時的會是什麼情境，值得我們期待。

## 五、愛玩電動的人工智慧-StarCraftII



圖片來源：DeepMind

圖 8-5、人工智慧玩 StarCraftII

DeepMind 公司在 AlphaGo 闖出名堂後，也在電競領域開發出 AlphaStar 遊玩星海爭霸二，超越人類玩家表現。星海爭霸是一款即時戰略遊戲，與 Dota2 不同的是，需要一次操動多隻角色。玩家需要同時在打仗與採集資源取得平衡，需要透過策略性的多工方式(一方面作資源效率分配與戰鬥)取得勝利，而由於上述動作皆藉由鍵盤與滑鼠執行，因此 APM(Actions Per Minute)是一衡量遊戲表現操作速度的指標，APM 越高代表著玩家短時間能進行越複雜的操作。

為了使得 AlphaStar 學會如何操作星海爭霸，研究人員以職業選手的電玩比賽影片，讓 AlphaStar 模擬這些手法，並且使不同版本的人工智慧之間互相對戰而學習強化，各策略之間有其優缺點可互補，沒有一套武功吃天下的道理。所以這次他們分別更新、記錄許多不同版本的策略，使得模型能快速學習到頂尖玩家所採用的策略與操作。

AlphaStar 的獲勝也引起玩家們的熱烈討論，多數人將 AI 的獲勝歸於其高速操作。一般職業選手 APM 平均為 300 左右，但 AlphaStar 某些時候能夠爆發至 1,000 以上 APM，也讓人質疑其公正性，且 AI 操作電玩存在著「AI 視野」，不像人類僅能看到局部視野。

縱使 AI 擊敗電競選手存在諸多爭議，但不可質疑的是近期 AI 所能執行的工作也日趨複雜，代表著人工智慧在未來生活的重要性是值得期待的。

## 六、結語

AI 在電競上的應用除了上述幾種外，還有應用在撲克牌、第一人稱射擊、超級瑪莉等，應用層面廣泛。而其中的強化學習技術套用到商業面的話，不僅常見於機器人效能的優化、金融交易的應用、室內導航，也因此可以說 AI 玩電競不是虛度光陰，而是能夠受惠於科技技術發展，並且增進生活福祉。

---

劉聖元 ( 台灣經濟研究院研六所助理研究員 )

A general reinforcement learning algorithm that masters chess, shogi and Go through self-play

AlphaZero: Shedding new light on the grand games of chess, shogi and Go

<http://sc.piee.pw/HUB8N>

<https://pse.is/GQ8F6>

DeepMind、OpenAI