

電腦圍棋四十年

作者：顏士淨、嚴初麒、許舜欽

顏士淨為國立東華大學資訊工程系教授，專長：人工智慧、電腦對局、深度學習及機器學習。

嚴初麒為僑光技術學院資訊科技系副教授，研究領域與專長：電腦圍棋、人工智慧、資料結構與演算法。

許舜欽為臺灣大學資訊工程系退休教授，現任長榮大學資訊管理系教授。專長為電腦對局，他也是現任臺灣電腦對局學會理事長。

緒論

自從電腦圍棋程式 AlphaGo 挑戰世界圍棋冠軍李世乜九段的消息發佈後，電腦圍棋的發展再次受到世人矚目。人機大戰對局立刻成為新聞話題，賽前雖然有各式各樣的預測，但絕大多數看好李世乜，甚至很多人猜測李應該會以 5：0 輕鬆取勝。原因之一是設計電腦圍棋程式的難度非常高，雖在過去歷經幾十年的研究探索，圍棋程式感覺上與人類專家之間仍有極大的差距。

另一個原因則是李世乜名氣太大。李世乜棋路變幻莫測、難以捉摸，宛如天外飛仙神來一手，有時會隨便捨棄看來很重要的棋子，讓人大吃一驚，而往往在其後的戰局中，先前的棄子還會參與戰鬥，順勢救出。中國棋壇為李世乜取了綽號「殭屍流」，不只形容他擅長讓「假死」的棋子還魂，也讚嘆圍棋的深奧不可思議。李世乜可稱得上是最具不可預測性的圍棋高手。

這場全球矚目的對局在 2016 年 3 月 9 日拉開序幕。AlphaGo 以穩健而均衡的步伐，絲毫不讓李變幻莫測的棋路佔到便宜，碰到李的無理挑釁時，也毫不客氣予以還擊，而且招法精妙讓人大開眼界。李世乜在第一場錯估電腦實力吞敗之後，接下來的對局心理明顯受到影響，不是太過保守就是太急於求戰。然而電腦不動如山，雖不主動求戰，卻也毫不懼戰，而且戰鬥的表現極為出色。最終電腦程式

AlphaGo 以 4：1 獲勝，也讓世人對人工智慧的表現大為嘆服。

AlphaGo 成了世界最強的圍棋程式之一，谷歌團隊 DeepMind 也將持續發展 AlphaGo。展望未來之餘，本篇文章將回顧電腦圍棋的歷史。電腦圍棋自 1970 年佐布里斯特（Albert L. Zobrist）設計出第一個可與人對奕的程式以來，至今已有 40 多年的歷史，圖 1 為 1985 年之後，電腦圍棋程式進步的情形。

根據這幾十年來電腦圍棋的發展情形，我們可將其分為四個時代：石器時代（1970-1985）、部落時代（1986-1997）、戰國時代（1998-2006）和機器學習時代（2007-）。以下簡述這四個時代的電腦圍棋歷史，來了解電腦圍棋的發展狀況。

石器時代（1970-1985）

將電腦圍棋初期的萌芽階段取名為石器時代，不

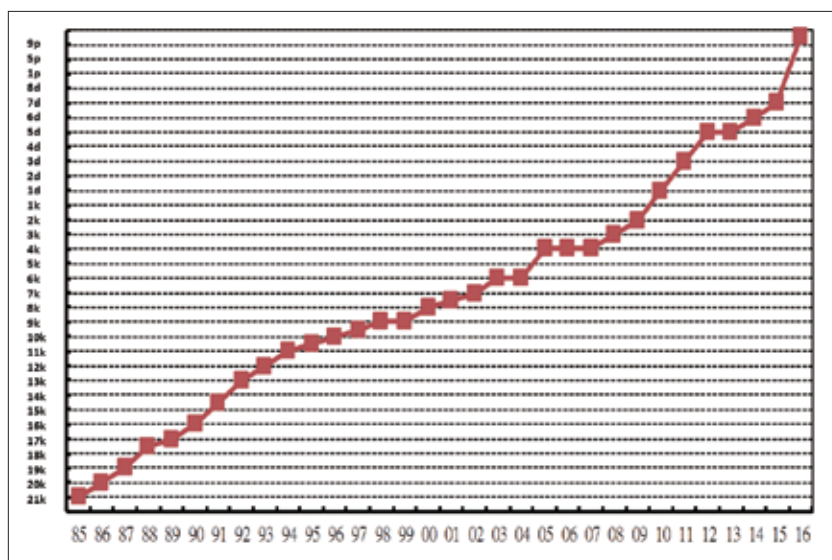


圖 1 | 1985 年之後，電腦圍棋程式進步的情形。

僅因為這是電腦圍棋文明的濫觴，也因為此時的發展情形頗類似人類的石器時代——電腦圍棋文明在世界各地獨立發展，彼此之間的接觸很少，也沒有公開的戰場讓它們比較棋力高下，其中很多發展數年後就夭折了。這個階段的電腦圍棋程式相對於人類棋手，棋力都很差。其中大部分是為了測試某些特別理論而寫的，真正經由整體規劃而完成的完整程式並不多。這個時期的代表性程式有 Zobrist[1]、Ryder、Reitman and Wilcox[2]。正如人類史一樣，現今活躍的圍棋程式中，沒有任何一支程式與這些「古」程式有血緣關係。這個時期的程式棋力約在 35 級到 25 級左右。

其實電腦對局的相關研究很久以前就開始了，不過當時的主要研究對象是西洋棋。最早可溯源至夏農（Claude Shannon）於 1950 年提出利用電腦來下西洋棋的概念，其後電腦對局就成為人工智慧領域中最吸引人的課題之一。這個階段的西洋棋或象棋程式，都已進入段位程度。相對而言，電腦圍棋的進步就顯得緩慢許多，而且一路崎嶇，荊棘滿布。

電腦圍棋自從佐布里斯特於 1970 年設計出第一個可以對奕的程式 Zobrist 之後，有 20 多年的時間進步緩慢，當時軟體環境逐漸從 DOS 轉為 Windows，而硬體支援在容量和速度上均與現在相去甚遠。早先的圍棋程式棋力有些慘不忍睹，因為運用在西洋棋、象棋程式的那一套窮舉搜尋演算法，完全不適用於圍棋。所謂窮舉搜尋，就是電腦會找出當下局面所有可走的棋步，並且嘗試走走看，然後電腦再反過來替對手找出所有可能的應招，而且也嘗試走走看。如此一來一往模擬數個回合之後，電腦程式再根據所走出的局面優劣程度，

來反推最初的最佳著手 [3]。

為何這套演算法不適用於圍棋程式？原因之一是圍棋棋盤比象棋大得多，可選擇著點平均就有 200 多個，這對電腦的負荷量太大了。更關鍵的一點是，圍棋不像象棋可以容易而迅速的評估局勢優劣。象棋或西洋棋都是以「擒王」這個明確目標作為獲勝條件，其他棋子不論偏重攻擊或防守（有時候也攻防兼備），都有其大致上的價值。如果我們把象棋的子力（棋子的價值）予以量化如表 1，那麼要判斷局面的優劣程度，只要將盤面上雙方的子力各自加總做個對比，就可得知。

表 1 象棋子力表

帥	仕	相	俥	馮	炮	兵
將	士	象	車	馬	包	卒
10000	200	200	1000	450	450	100-300

在表 1 中由於帥、將是不能死的棋子，所以給它一個極大值 10,000，其他棋子則依據高手兌子經驗給予相對分數，而兵和卒在過河前與過河後的位置高低價值有異，所以有彈性範圍。透過這樣的子力表，電腦就能用非常迅速的方式計算出局面的大致優劣程度。

對比之下，由於圍棋每個棋子均為等價，判斷優劣所依靠的是棋子排列的陣形，雖然決定圍棋勝負的是圍取地域的多寡，但這當中包含了太多諸如死活、厚勢、味道、地域確定性等等不確定的因素，所以圍棋盤面很難找到迅速有效的方式來判斷局勢優劣。在電腦圍棋的石器時代，軟硬體都相對不足，電腦圍棋只能當作測試理論的工具。

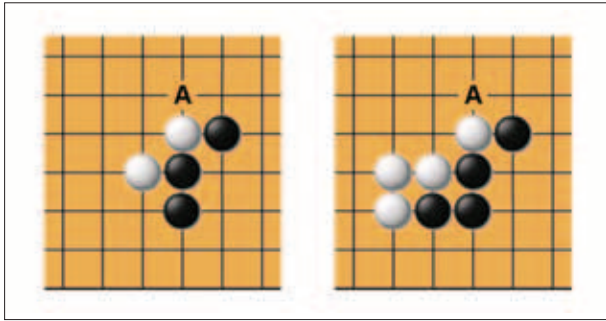


圖 2 棋形的範例。

部落時代（1986-1997）

相對於石器時代，部落時代（1986-1997）的圍棋程式棋力明顯提昇。這個時期促進圍棋程式棋力提昇的主要因素之一為「應氏盃世界電腦圍棋錦標賽」（1985-2000）的成立。應氏盃由應昌棋圍棋教育基金會主辦，是全球第一個世界性的電腦圍棋比賽 [4][5]。這項比賽提供了一個舞台，舞台上的程式約以每年進步一級的速度前進，代表性的程式有臺灣臺大資工的程式 Friday、Dragon、Stone、Archmagic、Jimmy，美國陳克訓的 Go Intellect，英國瑞斯（Michael Reiss）的 Go4++，美國傅蘭（David Fotland）的 Many Faces of Go (MFGO)，以及早期德國布恩（Mark Boon）的 Goliath。這個時期的程式棋力約在 23 級到 9 級之間。

此階段的程式大多利用一些技術來輔助選點下棋，例如在佈局階段輸入定石資料庫，以及利用人造棋形（pattern）資料庫來協助找出好點。例如若事先將圖 2 左圖建立為一個棋形，並指出白 A 長是好點，那麼碰到如圖 2 右圖之類似盤勢時，電腦就能判定白 A 長是好棋。對此也有相關研究論文發表，例如嚴初麒、顏士淨的〈棋形知識庫之設計與製作〉[6]。

除了棋形資料庫之外，當時最常使用到的技術主要是攻殺搜尋，這是一種由追殺某個棋串（此方稱做 killer）以及防禦該棋串（此方稱做 defender）這兩種副程序，互相循環呼叫，最後檢驗能否吃掉該棋串的技術。臺大許舜欽教授的學生劉東岳發展的 Dragon 程式在這方面下過不少功夫，曾有不錯成績。另外也有經過專業分析研究發表的論文，例

如嚴初麒、許舜欽的〈Killer Moves Generator 之研究與設計〉[7]。

這段時期也有不少國外的電腦圍棋程式設計者相繼投入，例如陳克訓的 Go Intellect、傅蘭的 MFGO、山下宏的 Aya、慕勒（Martin Müller）的 Explorer 等等，這些程式和臺灣顏士淨的 Jimmy 都是電腦圍棋界的長青樹，迄今仍在繼續研發中。

儘管使用了這些技術，但想要讓圍棋程式接近業餘高手仍有一段距離，例如攻殺搜尋時發現某個棋串可以吃掉，可是要如何判斷這個棋串是要子？吃掉這個棋串的價值高低？棋形的分數如何給定？這就需要更高層的分析技術與專業知識。圍棋程式作者有史以來棋力最強的鄭銘煌九段，他的程式若葉（Young Leaf），就用到大量專家知識，幾乎不需要搜尋就可以下出不錯的棋步，因此在 1998 年左右，成為日本最暢銷的圍棋軟體。

戰國時代（1998-2006）

由於棋力提昇，使得這些圍棋程式越來越具經濟價值。1997 年，全球各地共有 10 多套商業圍棋程式。1998 年，在日本舉辦的 FOST 盃電腦圍棋比賽中，前 10 名的程式除了 Jimmy 之外，其他均已上市發行。由於世界各國群雄並起，而且在這次比賽中，前八名的程式互有勝負，也就是說，每個程式都有擊敗其他程式的能力，所以我們將這個時期稱為戰國時代。

電腦圍棋發展至此，由於各項比賽的推動，以及商業化的結果，使得每年電腦圍棋的棋力均有穩定的進步。到了 20 世紀末，全球共出現了 10 多組由軟體公司或個人以專職方式發展的電腦圍棋程式，



圖 3 右邊是本文作者之一的許舜欽，中間是曾經稱霸十年之久的手談程式作者陳志行，左邊是本文作者之一的顏士淨。（顏士淨提供）

競爭非常激烈。

廣州中山大學退休教授陳志行所研發的手談（HandTalk），可說是戰國時代的代表程式。陳志行本身有約業餘五段棋力，且周遭不乏圍棋專業人士可供諮詢。他所撰寫的手談，可能是世界上唯一使用組合語言撰寫的圍棋程式。據說陳志行在化學系退休之餘，想開發圍棋程式，詢問資訊科技的同行哪種程式語言執行速度最快，對方推薦組合語言，因此空前絕後的組合語言圍棋程式就此誕生。

因為組合語言的特性，手談程式短小卻棋力強勁。陳志行的方法獨特，尤其擅長分斷對方棋子展開攻勢，且對於死活頗有獨到之處，也建立了相當專業的知識庫，無論在定石、棋形、局部手段、攻殺細算，在不少地方皆可看出其出類拔萃之處，以致手談自出賽以來（1993-2005）奪冠無數，無論應氏盃、FOST 盃或電腦奧林匹亞（Computer Olympiad）均連連告捷。在 FOST 的人機賽中，手談甚至贏過日本業餘三段，成為當時公認最強的圍棋程式。遺憾的是後來陳志行罹癌，手談棋力停

滯不前。2008 年，他辭世後，初段的手談成為最後遺作，但這個程式無疑已在電腦圍棋程式發展史上寫下光輝的一頁 [8]。

機器學習時代（2007-）

電腦圍棋的棋力之所以長期落後人類，最大的原因是人類棋手可以用視覺掃描盤面，快速判斷情勢，而電腦程式則不易找到良好的審局函數。自 2003 年起，布基（Bruno Bouzy）試圖打破這種情況 [9]。他運用蒙地卡羅法（Monte Carlo method）作為評估函數，並使用此評估函數做全局性搜尋，然而在棋力上始終沒有太大的突破。直到 2006 年，同樣使用蒙地卡羅法，庫隆（Rémi Coulom）的程式 Crazy Stone，才在義大利杜林舉行的第 11 屆電腦奧林匹亞的九路圍棋項目中奪得金牌。雖然如此，當時 Crazy Stone 僅在 19 路圍棋項目中奪得第五名，尚未撼動以人類思維為主的圍棋程式在 19 路圍棋的地位。然而在 2007 年，蒙地卡羅法結合上界信賴法（upper confidence bound, UCB）的樹狀搜尋——合稱蒙地卡羅樹搜尋（Monte Carlo tree search, MCTS）——的演算法出現。以蒙地卡羅樹搜尋為基礎的圍棋程式 MoGo，逐漸在一些非正式比賽中嶄露頭角。2007 年 6 月，第 12 屆電腦奧林匹亞於荷蘭阿姆斯特丹舉行。上屆 19 路冠軍 GNU Go、亞軍 Go Intellect 以及前文介紹過的 Crazy Stone 等程式均有參賽。MoGo 在強敵環伺之下，以全勝戰績奪得了 19 路圍棋項目的金牌，Crazy Stone 也拿下第二名，GNU Go 退居第三 [10]。這象徵蒙地卡羅樹搜尋的成功，也代表電腦圍棋的嶄新時代即將到來。

蒙地卡羅樹搜尋是一種使用強化式學習（reinforcement learning）的方法，可以在實戰時，利用自我對奕模擬結果來評估盤面，某種程度解決了先前所提圍棋不易找到良好審局函數的問題。而且這個方法可充分利用機器能力，使用的電腦 CPU 核心（core）數越多，棋力就越強。之後程式更開始利用各種機器學習技術，學習如何找出當前盤面較佳著手的機率分布，乃至於以近年開始發展的深度學習（deep learning），來直接評估盤面與產生著手。這些發展都充分運用了機器的效能，因此我們稱這個時代為機器學習時代。這個時代又可再約略分為四個時期：機器暴力搜尋、離線機器學習、加強模擬準確率、深度學習，以下簡略敘述這四個時期。

機器打群架——機器暴力搜尋時期

機器學習時代一開始由 MoGo 稱霸，MoGo 的作者為希梵（Gelly Sylvain）與王一早等人，後期由戴德（Olivier Teytaud）負責研發，並將此系統平行化，改寫至可以在 500-1,000 核的分散式系統機器上執行 [11]。這個程式是第一個使用蒙地卡羅樹搜尋為基礎的電腦程式，最早的版本在模擬階段是用專家知識撰寫模擬落子順序，並使用 3x3 的人造棋形。之後使用增強式學習改進模擬方法，並且在蒙地卡羅樹搜尋架構底下，加入了 RAVE 法 [12]。RAVE（rapid action value estimation，落子快速估值）可以加速收斂蒙地卡羅樹搜尋的結果，這個方法對於增加棋力可說立竿見影，加入普通圍棋程式，棋力可以馬上進步 200Elo 級分 ❶。

MoGo 建立了現代圍棋程式的典範。蒙地卡羅樹

搜尋程式與以往圍棋程式的行為有很大差異，首先它的佈局很差，可說是沒有佈局。但它擅長近距離接觸戰，也擅長收官，這可說是所有以蒙地卡羅法為基礎的圍棋程式共同的特性：局面優時，下得極端保守，常常最後只贏半目；局面不佳時，則開始想盡辦法翻盤，甚至下出像在耍賴的棋。這種特性在當時各項比賽中，把傳統的圍棋程式整得七葷八素，往往在過去能勝定的棋卻被翻盤，讓作者在一旁看得又好氣又好笑。

蒙地卡羅樹搜尋的最後一個特性是棋力跟使用的計算資源成正比。給予的計算資源愈多，棋力就愈強。因此在這個時期，很多程式的研究重心都放在如何有效平行化蒙地卡羅樹搜尋。其它利用大量核心來做暴力搜尋的程式還有 Fuego 與 Pachi，這兩個程式都是開源碼（open source）程式，對於電腦圍棋的發展有很大的幫助。

自 2008 年開始，MoGo 與臺灣東華大學和台南大學展開合作，開發 MoGoTW 版本，並且在九路圍棋非正式比賽，有擊敗職業棋手的紀錄。九路圍棋正式擊敗職業棋手則是在 2013 年，由東華大學周政緯研發的東華七號（Coldmilk）在七番賽中，以 4:3 勝職業棋士黨希昀。

回家讀書再來拼——離線機器學習時期

機器暴力搜尋的時代，很快被使用離線（offline）機器學習方法的程式所取代。Crazy Stone 的作者庫隆引入特徵強化搜尋（Minorization-

❶ 編註：Elo 是棋弈（甚至一般對局運動或遊戲）的評分系統，發明人為美國物理學家 Arpad Elo。



圖 4 包括 10 年來頂尖圍棋程式的一些作者：後排由左至右分別是 AlphaGo 的黃士傑、Jimmy 的顏士淨、Crazy Stone 的庫隆、Aya 的山下宏，前排左為 Zen 的加藤英樹，右為帶我們參觀日本棋院的七段記者。攝影地點是日本棋院的幽玄之室，只有四大頭銜比賽才會在這個房間舉行。（顏士淨提供）

Maximization, MM) 演算法 [13]，以監督式學習的方式從棋譜自動學習棋形的權重。學到的棋形，不僅可以用在加強模擬的準確性，也可以用在蒙地卡羅樹搜尋展開選點的優先順序。執行模擬棋局時，就分析每個可能著手的特徵，得出其重要性的分數，並依此計算選擇此著手的機率，依機率落子。使用這種新方法，Crazy Stone 對 GNU Go 的勝率，從 38% 提升到了 68%。除此之外，蒙地卡羅樹搜尋展開選點時，可以採用漸進式擴展 (progressive widening) 的技術，減少搜尋樹的分支度。首先，Crazy Stone 先依著手機率排序，隨著模擬次數逐漸增多，再逐漸開放子點以供選擇。此方法使其勝率從 68% 提升到了 90%。這時期另外還有 MFGO 與 Aya 等老牌程式，它們利用以往的基礎，在漸進式展開上面下了很多功夫，在這段期間的比賽成績也相當優異。

千錘百鍊事在人為——加強模擬準確率時期

除了利用機器學習技巧能更有效率的展開搜尋之外，另一個蒙地卡羅樹搜尋的重點就是模擬準確率的加強 [11]。以往利用機器自動學習加強模擬的方法有 RAVE 與 Simulation Balance，但都逐漸遇到瓶頸，後來就漸漸走向 Heavy playout，以較複雜的策略來進行模擬。這方面做得最好的首推日本的圍棋程式 Zen。Zen 的作者為尾島陽兒（暱稱

Yamato）與加藤英樹。尾島原本是日本知名的遊戲設計師，自 2007 年開始投入電腦圍棋。據說他除了每周出外採買糧食之外，大部分時間都足不出戶，所有精力放在程式開發。Zen 的主要成功原因是

尾島將專家知識非常有效率、幾近完美的放在模擬當中。Zen 的商業版本只要進行 1,000 次模擬，就可以到達初段水準，10,000 次模擬就可以接近四、五段，這是以往動輒需要千百萬次模擬的程式無法想像的。從 2009 年到現在，Zen 一直是世界各項競賽的常勝軍，與 Crazy Stone 兩個程式可謂一時瑜亮，同時其商業版軟體也是最暢銷的兩個程式。Zen 的程式還在不斷進步中，目前的版本 Deep Zen，已經在圍棋服務器 KGS 網站有 9d 等級，達到職業水準。

全面模仿人類—深度學習時期

2008 年起，就有人將深度捲積式類神經網絡 (Deep Convolutional Neural Network, DCNN) 運用到圍棋。因為圍棋歷史悠久，累積逾數百萬份專家棋譜，是研究 DCNN 很理想的平台 [14]。2015 年，這個方法已成功將預測率提升到 55% 的正確率，超越以往其他機器學習方法的 45% 水準 [15]。同時，電腦圍棋界也有許多人將 DCNN 直接用在圍棋程式 [16]。趨勢科技董事長張明正在 2014 年，邀請王銘琬與趨勢的工程師和東華大學共同成立 GoTrend 團隊，研究 DCNN 在圍棋的應用。2015 年，GoTrend 在日本 UEC 電腦圍棋比賽表現不錯。2016 年，日本 UEC 比賽的前八名有七隊都採用深度學習技術，其中來自臺灣交通大學吳毅成教授領

導團隊開發的 CGI 程式，更在初賽時一鳴驚人擊敗 Zen，以全勝戰績晉級。值得一提的是，CGI 在 2016 年 3 月就曾有擊敗職業棋士的紀錄。同時，與賽隊伍甚至有三隊只用深度學習網絡，未使用搜尋，也能取得前 16 強的成績。由此可見 DCNN 造成電腦圍棋棋力的大躍進。像東華大學的圍棋程式塔羅圍棋 (Tarogo)，單機版也已經達到業餘七段，超越作者顏士淨本身六段的棋力 [17]。

隨著機器軟硬體的進步，2016 年，擁有龐大計算資源與優秀人才的谷歌 DeepMind 團隊以 AlphaGo 擊敗南韓棋王李世乭九段，震撼了全世界。AlphaGo 主要是以深度學習與蒙地卡羅樹搜尋的技術來開發，並以強化式學習的方式提高棋力。AlphaGo 可分為三部分：蒙地卡羅樹搜尋、策略網絡 (policy network) 與使用強化式學習開發的價值網絡 (value network) [18]。(詳細技術內容請參閱本期〈打開 AlphaGo！〉)

AlphaGo 對圍棋界帶來很大的衝擊，被圍棋界尊稱為 Alpha 老師，形成了 Alpha 棋風，促使人類棋手從新觀念來發展圍棋 [19]。

大體而言，深度學習技術可以將整理專家知識的過程自動化，更精準而有效率地吸收專家知識。電腦圍棋技術到了這個階段，可說是出頭天了。深度學習技術的影響將會非常深遠，不僅只於擊敗人類棋王的勝利。其相關技術在廣泛應用後，將對人類社會產生衝擊，可望促成許多技術與服務的提升，減輕人類整體工作負擔，享受更好的生活品質。

結語

AlphaGo 雖然贏了李世乭，但也顯示目前人類對

圍棋的了解還是太少，DeepMind 團隊仍持續研發電腦圍棋中。被公認為有史以來最強圍棋棋士的吳清源曾表示，圍棋實在太難，他自己對圍棋的了解還不及十分之一。在電腦圍棋 40 多年的歷史當中，研發人員前仆後繼，將電腦圍棋棋力不斷向前推進，持續探索圍棋之道。雖然終於出現戰勝九段棋士的 AlphaGo，但相關人員心裡也都清楚，人類與電腦程式在圍棋上，都還有很大的進步空間。電腦圍棋的研究才正要進入嶄新的時代，在這個時代，電腦將與人類並肩作戰，一起研究圍棋的真理。☺

本文參考資料請見〈數理人文資料網頁〉<http://yaucenter.nctu.edu.tw/periodical.php>

延伸閱讀

► 顏士淨、許舜欽〈電腦圍棋的發展概況〉卷 1 (1997) no. 2, *Communications of IICM*。

http://www.iicm.org.tw/communication/c1_2/page04.html

► 國際電腦遊戲協會 (ICGA) 網站。ICGA 是電腦遊戲程式奧林匹亞的主辦單位。

<http://www.game-ai-forum.org/icga-tournaments/>

► 圍棋棋力評比網站 GoRatings。AlphaGo 在短暫取得冠軍之後，目前又暫居第二，落後柯潔。不過 AlphaGo 自從與李世乭戰後並未再比賽過。

<http://www.goratings.org>