Introduction – Games, Computers, and Artificial Intelligence

- Taxonomy of Games
- AI and Game Playing
- Shifting Goals
- Early Ages
- History
- Resources
- References
  - Jonathan Schaeffer, Jaap van den Herik (2002). Games, computers, and artificial intelligence. Artificial Intelligence 134,

Acknowledgement:

- The slides of this chapter are modified from Prof. (Tsan-Sheng) Hsu's teaching material and Slides of Prof. (Shun-Chin) Hsu's and Prof. dr. H. Jaap van den Herik, under their courtesy. http://www.iis.sinica.edu.tw/~tshsu/tcg2007/index.html

Taxonomy of Games

- According to number of players
  - Single player games: puzzles
  - Two-player games
  - Multi-player games
- Elements of chances
  - Deterministic vs. Stochastic
    - Chances from outcome uncertainty
      - e.g., rolling dice or dealing cards
  - Perfect information vs. Imperfect information
    - Chances from state uncertainty
      - e.g., invisible opponent positions or actions

Examples of Game Classification by Chance

	Perfect Information	Imperfect Information
Deterministic	Checkers, Chess, Chinese Chess, Connect6, Go, Hex, Shogi	Duplicate Bridge, Kriegspiel, Phantom Go, Stratego
Stochastic	Backgammon, Chinese dark chess, EinStein Würfelt Nicht!, 2048	Mahjong, Poker, Rummikub, Scrabble

What Is Artificial Intelligence?

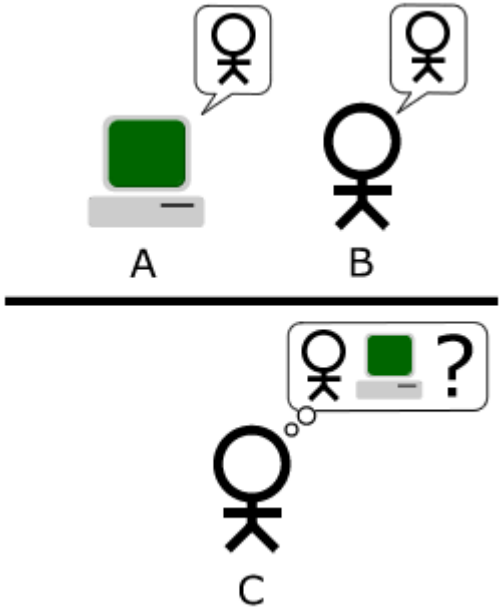
- Professor John McCarthy (Father of AI) :
  - The science and engineering of making intelligent machines, especially intelligent computer programs.
- Russell and Norvig (Artificial Intelligence: A Modern Approach):
  - Thinking Humanly, Thinking Rationally
    - 人性化的思維、理性化的思維
  - Acting Humanly, Acting Rationally
    - 人性化的行為、理性化的行為
- Definition of Kai-Fu Lee (李開復)
  - 感知 (視覺、語音、語言)
  - 決策 (識別、推薦、預測；如人臉辨識、下棋、股市預測)
  - 回饋 (學習、邏輯推論、機器人、自動化)

Artificial Intelligence

- Patrick Henry Winston [1984]
  - Artificial Intelligence (AI) is the study of ideas that
    - **enable computers to be intelligent**.
  - One central goal of AI is
    - **make computers more useful**.
  - Another central goal is
    - **understand the principles that make intelligence possible**.
      - **Making computers intelligent helps us understand intelligence**.
      - **Intelligent computers are more useful computers**.

Turing Test (Humanly)

- How to define “artificial intelligence"?
  - If a machine is intelligent, it cannot be distinguished from a human.
- "Can machines think?" – Turing’s question in 1950.
  - But, hard to define. So, proposed the Turing test as follows.
    - A human judge engages in a natural language conversation with one human and one machine,
      - each of which tries to appear human.
    - All participants are placed in isolated locations.
    - If the judge cannot reliably tell the machine from the human, the machine passes the test.
    - The conversation is limited to a text-only channel
  - Computer passes Turing Test for the first time, 2014.
- "Are there imaginable digital computers which would pass in the Turing test"
  - Turing's new question.



Game Playing and AI (Rationally)

- Elaine Rich [1983]
  - Intelligence requires knowledge.
  - Game playing and AI.
    - Games hold an inexplicable fascination for many people,
    - The notion that computers might play games has existed at least as long as computers.
  - Reasons why games appeared to be a good domain in which to explore machine intelligence.
    - They **did not obviously require large amount of knowledge** (when compared with other AIs).
    - They provide a structured task in which **it is very easy to measure success or failure**.
- Schaeffer & Herik [2002]:
  - **"Chess is to AI as drosophila (the fruit fly) is to genetics"**

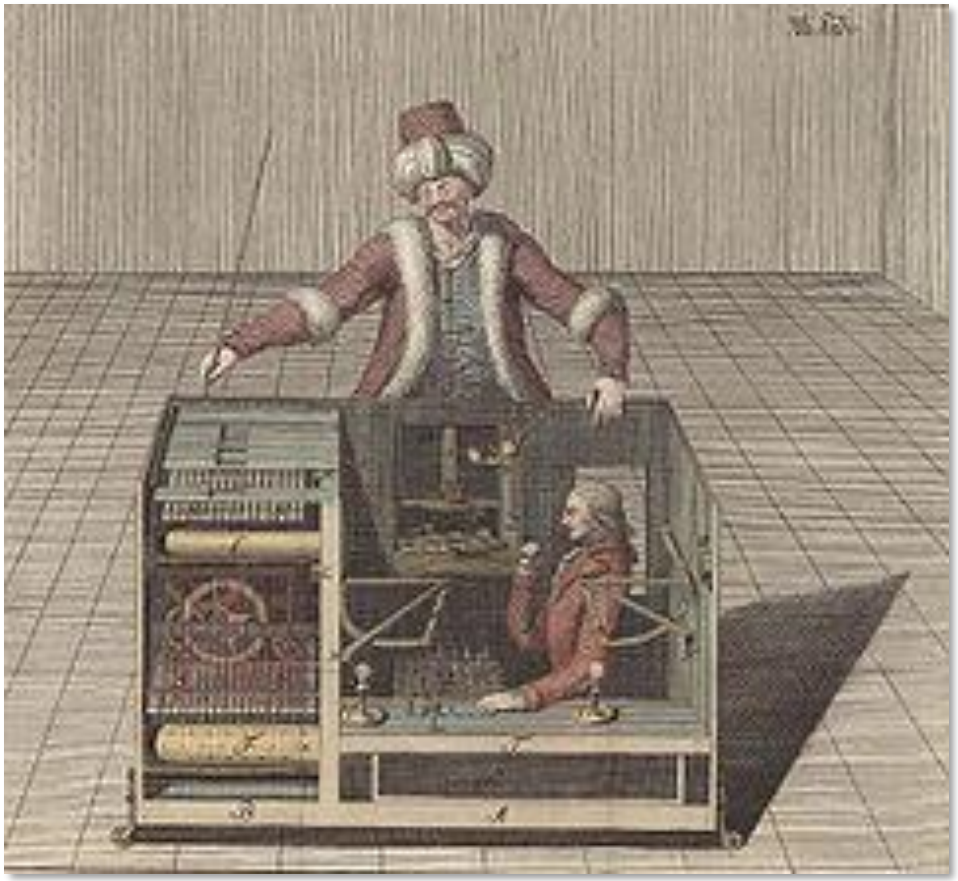
"Chess as the Drosophila (the Fruit Fly) of AI"
– Alexander Kronrod, mid-1960s
– Also John McCarthy, 1990

More Accurately,
"Go is the fruit fly of AI"
"Computer Games are the fruit fly of AI"

Early Ages: The Maelzel's Chess Automaton

- Late 18th century.
  - The Turk, or Automaton Chess Player,
    - Constructed by a Hungarian named Wolfgang Von Kempelen.
    - Exhibited from 1770 for over 84.
    - Defeated many challengers, including Napoleon and Franklin.
    - Operated by a concealed human chess-master. (Hoax)
    - Destructed by fire in 1854.
    - The hoax was revealed in 1857.
    - "Recently" reconstructed in Californian.
  - "Arguments" made by the famous writer Edgar Allen Poe in "Maelzel's Chess Player".
    - It is as easy to design a machine which will invariably win as one which wins occasionally.
    - Since the Automaton was not invincible it was therefore operated by a human.

1769 Wolfgang von Kempelen

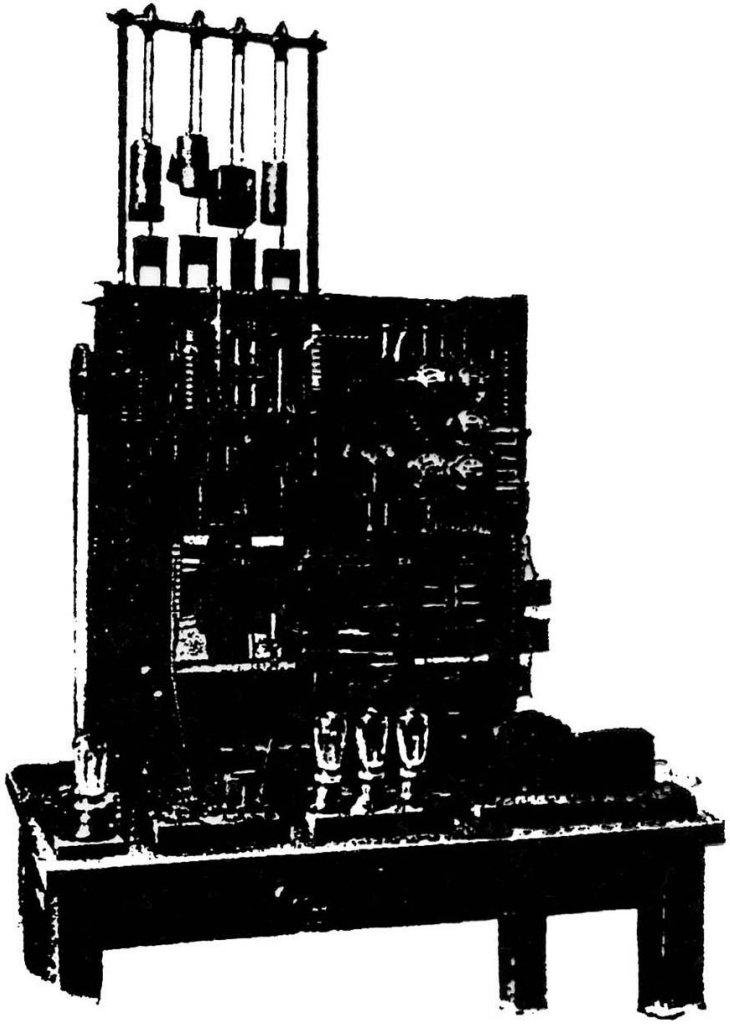


1770 De Turk

"I will invent a machine for a more compelling spectacle within half a year"

Early Ages: Endgame Chess-playing Machine

- 1912
  - Made by Torres y Quevedo.
  - Plays an end game of king and rook against king without human.
  - The machine played the side with king and rook and would force checkmate in a few moves however its human opponent played.
  - An explicit set of rules are known for such an endgame.
  - Very advanced automata for that period of time.




EL AJEDRECISTA (1912)
Leonardo Torres y Quevedo

History

- 1950s and 1960s: as AI were plagued by over-optimistic predictions.
  - C.E. Shannon, 1950, Computer Chess paper
  - Arthur Samuel began his 25-year quest to build a strong checkers playing program at 1952
  - Alan Turing, 1953, "Faster than thought", chapter 25
- 1970s and 1980s.
  - Concentrated on Western chess and Brute-force approach
    - The CHESS series of programs by the Northwestern University.
    - Analysis of Alpha-Beta pruning by Knuth and Moore at 1975.
  - Building faster search engines ➔ Chess-playing hardware.
  - Quantified by Ken Thompson.

Recent History

- 1980s.
  - Advances in theory of heuristic searches.
    - Conspiracy numbers
    - Search enhancements such as null moves and singular extensions
    - Learning ideas
- 1990s
  - Witness a series of dramatic computer successes against the best of humanity.
    - CHINOOK, checkers, 1994.
    - DEEP BLUE, chess, 1997.
      - Beat World Chess Champion Kasparov
    - LOGISTELLO, Othello, 1997.
      - Beat World Othello Champion Takeshi Murakami

1996: First Match Kasparov – Deep Blue: 4 - 2

世界日報
中華民國八十五年二月十一日 星期日

電腦人腦棋賽・機器贏了頭盤

世界棋王下得辛苦・國際網路「實況轉播」

【本報記者李勇費城報導】國際商業機器公司（IBM）的「深藍」西洋棋計劃挑戰棋王蓋瑞・柯斯巴魯夫（Garry Kasparov）的比賽，首場IBM「深藍」旗開得勝。

棋王柯斯巴魯夫是在下棋三小時後，知道大勢已去，雖然整個賽程有七個小時，但他決定放棄比賽，承認失敗。

這場電腦與人腦棋賽於十日下午三時準時在費城市會議中心展開，操縱電腦的是「深藍」計劃五人小組成員許峰雄，他與棋王對坐在棋桌兩端，開賽前五分鐘，兩人迅速走動，容許新聞記者從不同角度拍攝兩人的照片，五分鐘後棋賽進入情況，需要思考才可移動棋子，站在旁邊的裁判立即命記者向後退回記者席，不讓他們照相以免擾亂棋「王」的思考。

IBM「華生研究中心」數十名研究員在「深藍」計劃負責人譚崇仁安排下，全部出動配合支援並接待三百多名購票入場的電腦專家及西洋棋愛好者。讓他們在另一大廳面對電視螢光幕把棋賽進行看得一清二楚。

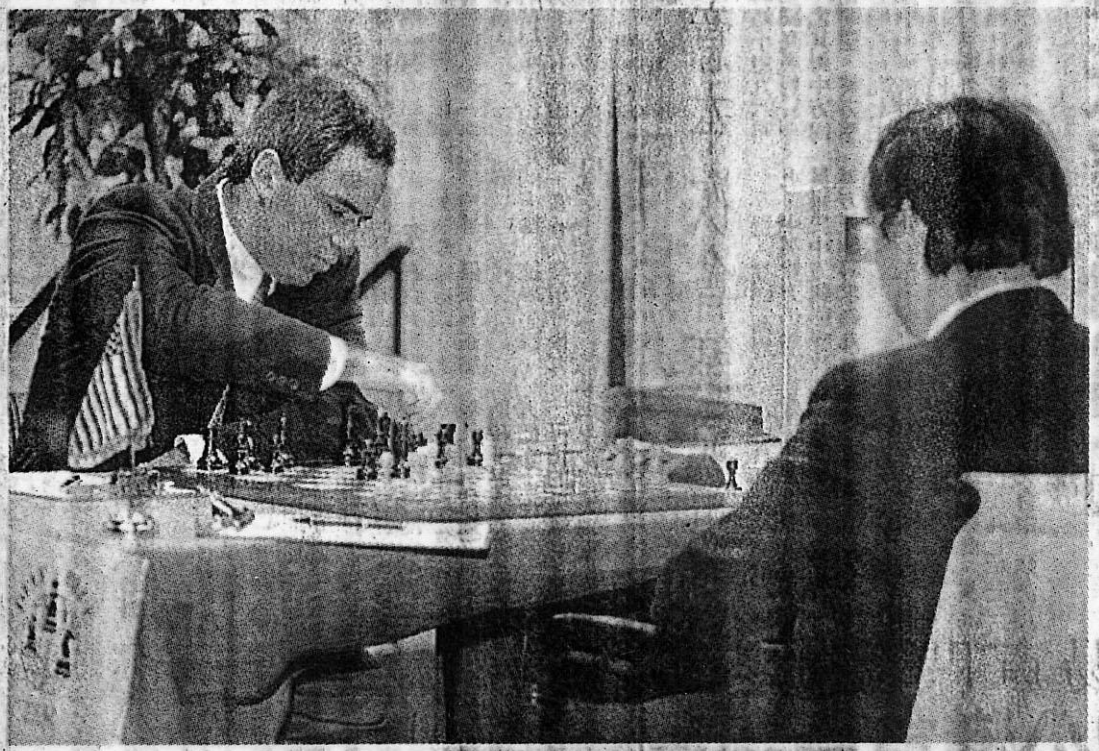
觀看棋賽的大廳與比賽場地完全隔離，IBM特別請了美國棋王耶舍・斯來維及一名黑人西洋棋高手向觀眾解釋戰況，說明棋政、分析優劣。預測勝負。

另外一個小組研究人員在另一個房間把比賽現場的實況透過國際電腦網路立即傳送到世界各地，讓關心這場賽事的人，不論在任何地方，都可立即從網路上知道戰況與結果。

棋賽由「IBM」「深藍」先下，華裔研究員許峰雄在電腦鍵盤上按了幾下，按照電腦的指示先走第一步，然後以筆記下來，棋王立即回應第二步棋，又以筆記下來，棋賽開始廿分鐘後，「棋王」開始進入長考，不久便離座到後面休息十分鐘後才出來走下一步棋，電腦的反應很快，跟著又下一步，於是棋王又開始長考。

根據耶舍・斯來維的講解，電腦在開始一個小時的攻擊凌厲，因此棋王應戰得十分辛苦，他之所以離席休息，目的是想出一個可以突破電腦系統的方法。果然，他休息出來後，連續下了幾步狠棋，直搗對方陣地，顯示出棋王不凡的思考與傑出的棋藝。但是電腦快速冷靜的技巧明顯超過棋王，經過三小時的鏖戰之後，棋局上雖看不出誰有敗象，但棋王心裡有數，宣布投降。

「深藍」計劃的經理譚崇仁在賽後興奮地說：他們只要贏了第一場與第二場的比賽，整個賽局就可控制下來，他仍然估計，這次六場比賽，IBM會以四比二的贏率擊敗棋王，為電腦分析寫下新的一頁。



世界西洋棋冠軍柯斯巴魯夫（左）十日不敵電腦，他在費城舉行的與美國電腦計算協會的國際電腦西洋棋冠軍IBM的「深藍」系統對弈六局中，第一局提前認輸。

圖爲比賽的實況，代表「深藍」系統在棋盤上移動棋子者是該系統的設計人之一許峰雄。（美聯社）

DEEPBLUE (IBM) Beats Kasparov by 3 ½ - 2 ½

Deep Blue Beats Kasparov

- Feng-Hsiung Hsu (The Man Behind Deep Blue)

戰挑腦電敗打、擊追勝乘王棋

定肯獲亦現表劃計藍深、功成冕衛負一和二勝三

【本報記者李勇十七日費城報導】世界西洋棋王柯斯巴魯夫，今天下午六時四十六分在與IBM電腦挑戰決賽中，以凌厲的攻勢乘勝出擊，再度打敗「深藍計劃」，以三勝二和一敗的戰績衛冕成功，並贏得「國際電腦計算協會」（ACM）頒給的獎金四十萬元，IBM則分得獎金十萬元。

這場人腦與電腦較量的西洋棋決賽是十七日下午三時準時開始，棋王一上桌就穩紮穩打，小心謹慎，快到中局時，突然發動凌厲攻勢，明顯佔盡先機，當棋子下到第四十三步時，棋王的優勢十分明顯，「深藍計劃」的研究人員雖明知還可以繼續下一段時間，但不願拖延時間，承認失敗，電腦於是輸了最後一場。

IBM「深藍計劃」以譚崇仁為首的研究員，最後悔是十六日的第五場比賽沒有接受棋王的平手要求，結果敗北，使棋王的信心恢復，譚崇仁承認他們估計錯誤。

保持王座的「棋王」柯斯巴魯夫在棋賽結束後，休息了廿幾分鐘，於七時卅分在比賽場地—賓州會議中心與譚崇仁共同舉行了一個簡短的記者會。

他承認這次與「深藍計劃」對陣，是他下西洋棋十幾年來最艱辛的一次。

柯斯巴魯夫承認「深藍計劃」電腦的棋技高強不能輕視，他與電腦較量，吸收了很多經驗，增長了自己的棋技，使他的棋藝更上一層樓。

柯斯巴魯夫說：他在最後這場決賽中，不敢躁進，不敢冒險，不讓自己重蹈第一場失敗的覆轍，相反，電腦運作一開始就暴露了弱點，給棋王抓住弱點猛攻而敗北。

譚崇仁在接受記者訪問時，承認「深藍計劃」雖集中了一百年來無數西洋棋賽冠軍的棋譜分析，集合了西洋棋高手的智慧，結果還是棋王得勝。棋王雖證明了他的棋技高強，但電腦「深藍計劃」也證明了功力不凡，替電腦科技打開了一扇大門，使電腦應用有了一個巨大的突破。

譚崇仁說，他所領導的「深藍計劃」這次挑戰「棋王」目的是考驗他們花了六年時間研究出來的成果，現在，電腦雖以一勝二和三敗而輸棋，但卻證明了電腦科技向前邁進了一大步，這是他這個部門深以為榮的事，也是IBM從事電腦研究最大的一次收穫。

世界西洋棋王柯斯巴魯夫過經六場比賽，分積的二比四以城費在日七十，IBM 敗擊，分積的二比四以城費在日七十，賽比場六過經夫魯巴斯柯王棋棋洋西界世
智才的測可不深類人對面技科腦電新最明證度再也，照拍者記受接並，「劃計藍深」的展發
。及不有仍，力潛
（社聯美）

柯斯巴魯夫與譚崇仁在記者會結束後，到另一個觀棋賽聽講評的房間與兩百多名觀眾見面，接受他們的提問。

棋王在回答觀眾提問中承認他通過與電腦下棋，學了很多東西，使他終生難忘，今後如有機會，他還願意接受挑戰，不過，他的估測還像以前一樣：在本世紀結束之前，電腦不可能擊敗他。

棋賽結束後，IBM人員及棋王一同到鄰近的馬里奧酒店參加盛大的頒獎典禮，IBM所有高級人員全部出席盛會，大家都為他們的研究成果高興。

譚崇仁說，五十年前二月十四日，美國第一部電腦在費城開始運作，所以「國際電腦計算協會」（ACM）特別選費城作為電腦挑戰人腦的賽場，最後人腦雖勝，但電腦功能也得到肯定。

對弈第六回合 電腦妙算十九步定江山

深藍挑戰成功

棋王輸在動了情緒

由來自港、台華裔電腦專家等所領導的I[illegible]「深藍」電腦經過六局鏖戰後擊敗世界西[illegible]王卡斯帕洛夫，創下電腦首度打敗人類棋[illegible]歷史性紀錄，並造成卡斯帕洛夫稱霸西洋[illegible]以來首次嘗到敗績。

「深藍」十一日下午只花了一小時就擊[illegible]斯帕洛夫，以六局總積分三點五比二點五[illegible]績，迫使卡斯帕洛夫稱臣。第六局最後一[illegible]鍵決賽進行時，現場及透過電視螢幕前的[illegible]棋迷屏息觀戰，棋賽在第十九手分出勝負[illegible]王棄子投降。

卡斯帕洛夫認輸後表情懊惱、震驚，這位去年曾以四比二戰績在費城擊敗「深藍」並誇言「廿世紀結束前不可能有人或電腦擊敗他」的棋王，隨後在記者會中說：「我必須向棋迷致歉，我對最後這一局的表現感到慚愧。」

卡斯帕洛夫解釋說，他是人，當他看到有些情況超越他所理解的範圍時，他會感到害怕。他又說，這次棋賽可以說是「世界上棋藝最高的人在壓力之下俯首稱臣。」

專家們認為卡斯帕洛夫是在生理和心理的雙重壓力下，輸掉第六局比賽。在場觀棋的法國西洋棋高手勞特爾說，縱橫棋壇十餘載的棋王在下第十九步棋就俯首稱臣，令人震驚，因為這局似乎沒有輸，過去都是這麼下，這是眾所週知的棋步。勞特爾說，卡斯帕洛夫有一點失常，他一直想調整步伐適應電腦但不夠積極。

在這局持黑子後攻的卡斯帕洛夫原擬迂迴佈陣，但開賽後不久即犯下失誤，讓深藍在第八步棋犧牲騎士換取卡斯帕洛夫的卒子，打亂了佈局，從佈陣轉為交戰狀態。卡斯帕洛夫陸續取下深藍的城堡、主教和騎士，卻在第十八步棋輸掉皇后，但還不至於輸陣。然而，深藍第十九步棋即把卡斯帕洛夫的國王困坐愁城，卡斯帕洛夫被迫棄子稱降。

「深藍」計畫主持人譚崇仁在賽後記者會上表示，「深藍」計畫全體人員對能夠贏得這次比賽勝利，覺得十分驕傲，並且感到榮幸能參與這一歷史事件。譚崇仁對於卡斯帕洛夫這六場的表現給予極高度評價，強調棋王智力過人，了解電腦將可以帶人類發展到什麼起步。

卡斯帕洛夫也抱怨未能取得深藍過去比賽時的棋譜。根據西洋棋比賽規則，雙方可以要求研究對方過去比賽的棋路。卡斯帕洛夫說，他最大的錯誤就是沒有提出條件，讓這場比賽進行更公平些。卡斯帕洛夫不排除再與深藍對弈的可能，但他堅持在不同的條件下對決，例如讓他和電腦先作練習賽。

勝方「深藍」計畫小組將可獲七十萬美元的獎金，IBM公司已決定將這筆獎金用於電腦研究，至於卡斯帕洛夫則可獲四十萬美元的獎金。

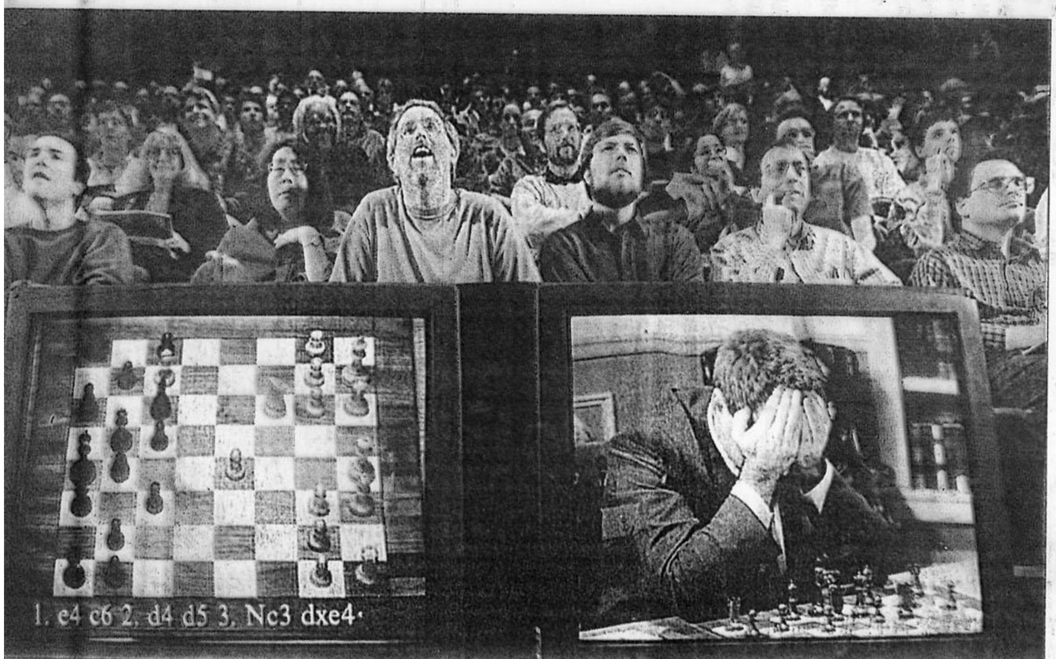
【本報綜合紐約十二日外電報導】超級電腦「深藍」擊敗世界西洋棋王卡斯帕洛夫，不但是一件劃時代的大事，也讓人重新省思人類與電腦之間的關係。不過，卡斯帕洛夫和創造深藍的IBM研究人員對於深藍究竟只是一個龐大的計算機或是一種新的智慧看法互異。

西密西根大學哲學教授麥格魯說，人類一直無法擺脫失去對自己發明物控制權的恐懼，尤其是像深藍這類有思考力的物體。

撰寫有關先進電腦書籍的作家麥柯杜克女士說，人們一直有個迷思，認為下西洋棋和人類智慧的發展息息相關。但是深藍獲勝並不意味它比人類更聰明，只能說這個電腦棋王下了一場精采的棋賽。

加拿大亞伯塔大學電腦學教授薛佛也說，若和醫生以及當醫生所需的技術相比，或是和謝瓦茲卡夫將軍在波斯灣戰爭中的運籌帷幄比較起來，棋術是非常簡單的，只涉及特定數目的物件以及簡單明確的規則。

但卡斯帕洛夫卻無法只把深藍視為計算機。觀察家說，其王輸在動了情緒，他被對手永不



[illegible]深藍的靈魂人物》

「瘋鳥」許峰雄 天才人物

「深藍」一戰成名，經過六場鏖戰，打敗號稱當今最偉大的西洋棋棋王卡斯帕洛夫。「深藍」強大的棋力來自IBM的研究團隊，這個七人小組的靈魂人物則是台大電機系畢業、今年才卅九歲的許峰雄。

「深藍」稱王的原因在於它有超級快速的中央處理器以及特別撰寫的程式，許峰雄就是設計、製造CPU及開發控制程式的主力。

「他實在是一個天才型的人。」台大電機系教授雷欽隆是許峰雄的大學同窗好友，得知電腦打敗棋王時，他的反應與許多計算科學界的人相似，並不感到意外。他說，綽號「瘋鳥」的許峰雄是天才，但也「有一點怪」。

許峰雄在大學時成績就非常好，語文能力也很強，對吸收國外新知有直接幫助，也老是英文歌不離口；但他不喜歡上課。雷欽隆回憶，許峰雄常在上課的時候和同學聊天，聊到別人都快沒有辦法聽課的時候，他會突然問老師一個問題，於是老師「就僵在那裡了」。但他不讀書，也能在電機系一百八十名學生中，搶到第三、四名畢業。

大學時許峰雄就對電腦下棋的研究產生興趣。他的圍棋實力在一段左右，西洋棋只是[illegible]是想教電腦下[illegible]是改以西洋棋[illegible]式，但當時國[illegible]當然也就沒有[illegible]

大學畢業後[illegible]時卡內基美倫[illegible]究院院士孔祥[illegible]念。卡內基美[illegible]坦福並稱美國[illegible]峰雄是該校大[illegible]個給獎學金的[illegible]

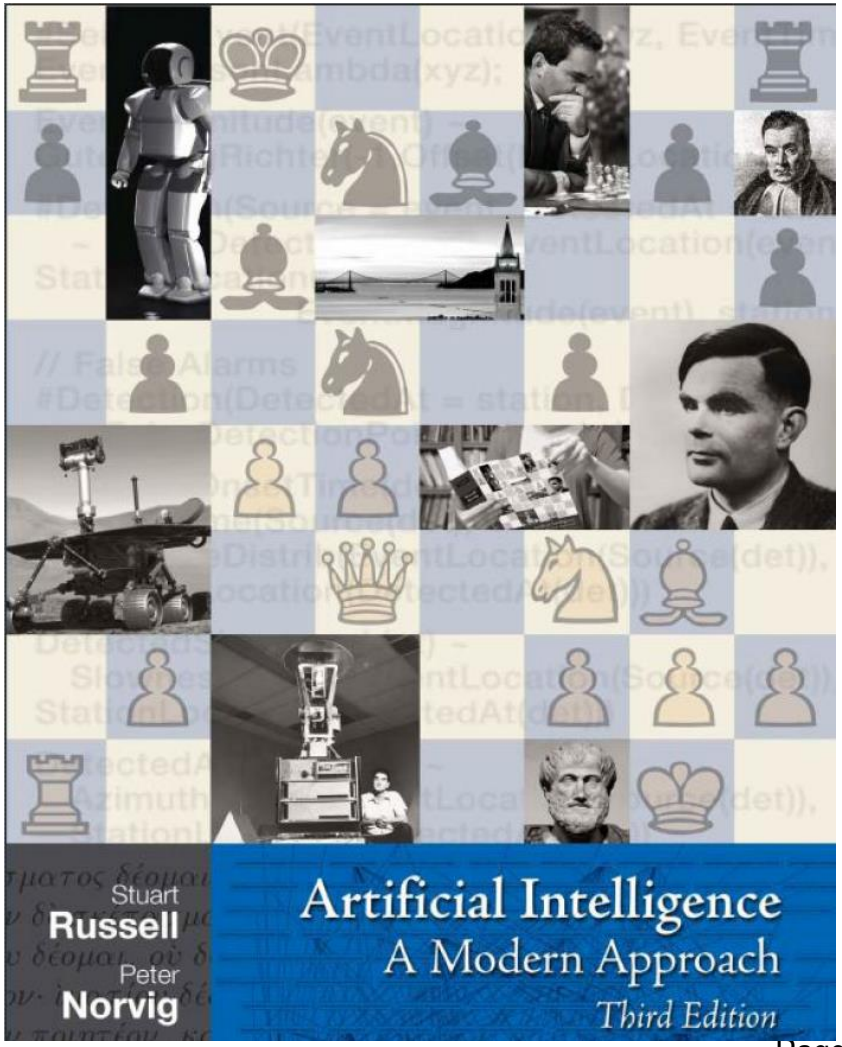
到了美國，[illegible]峰雄的興趣也[illegible]所的時候，他[illegible]的電腦，這是[illegible]思」的計算能[illegible]步，只是「深[illegible]經打敗過不少[illegible]九四年贏得世[illegible]震驚電腦界與[illegible]界談到電腦下[illegible]就是「NO.1」

今年初，許[illegible]系演講的時候[illegible]今年二度與卡[illegible]贏。如今預言[illegible]樂得再度歡唱[illegible]

Crazy Bird (CB)

Cover Page of the AI Bible

- The final position from the decisive game 6 of the 1997 match
  - between chess champion Garry Kasparov (Black) and program DEEP BLUE.
  - Kasparov was forced to resign, making this the first time a computer had beaten a world champion in a chess match.



"Go Has Emerged as a Major Challenge to AI Research"

– Schaeffer & van den Herik 2002

Modern History

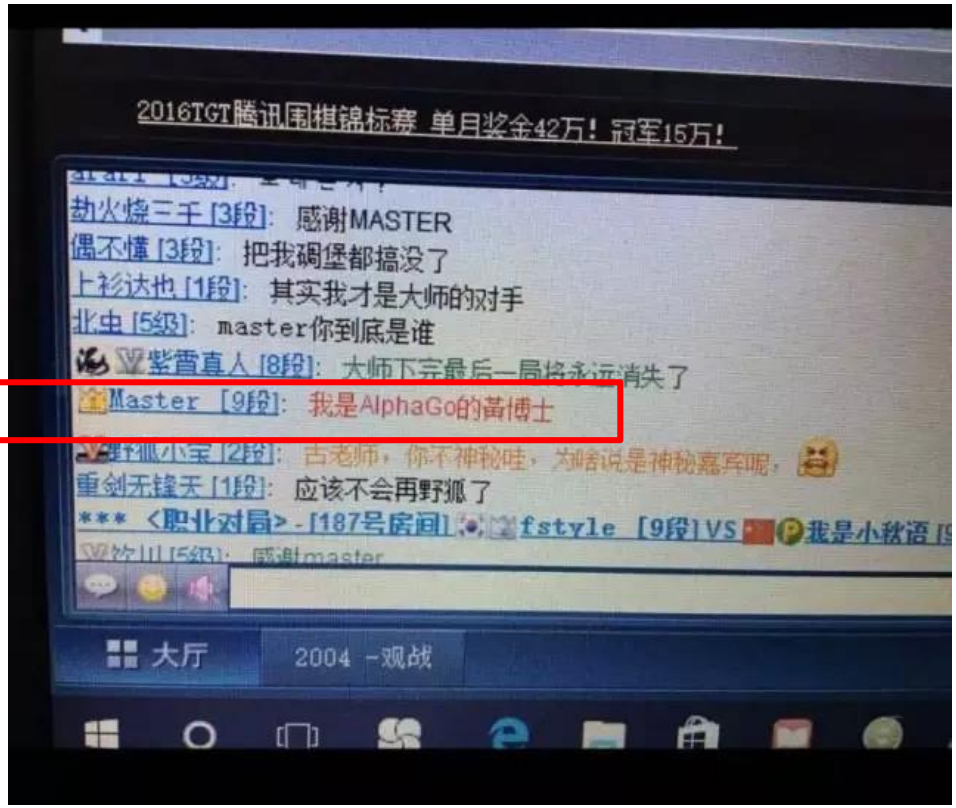
- 2000s (most from 2005).
  - Checkers was solved in 2007.
  - Many more challenging games
    - Connect6, Havannah, Go.
  - MCTS (Monte-Carlo Tree Search): a new breakthrough
    - Go becomes new drosophila.
    - MCTS brought the 9x9 Go program (MoGo) to high dan level in 2007.
    - In 2008,
      - 9x9 Go: Mogo already beat professional 7-dan player in 2008.
      - 19x19 Go: Mogo beat Amateur 6dan with 5-stone handicap in 2008.
    - In 2012, the best Go program is ranked 6 dan.
      - In 1998, very strong players were able to beat computer programs at handicaps of 25–30 stones.
      - In contrast, in the 1994 World Computer Go Championship, the winning program, Go Intellect, lost all 3 games against the youth players on a 15-stone handicap.

AlphaGo vs. 李世石

- 4:1 (left: 黃士傑, right:李世石)

Master Beat Go Champions/Grand Masters on 2016/12/30 ~ 2017/1/4 (not official)

60 (master) : 0 (human)



AlphaGo

04/01/17

We've been hard at work improving AlphaGo, and over the past few days we've played some unofficial online games at fast time controls with our new prototype version, to check that it's working as well as we hoped. We thank everyone who played our accounts Magister(P) and Master(P) on the Tygem and FoxGo servers, and everyone who enjoyed watching the games too! We're excited by the results and also by what we and the Go community can learn from some of the innovative and successful moves played by the new version of AlphaGo.

Having played with AlphaGo, the great grandmaster Gu Li posted that, "Together, humans and AI will soon uncover the deeper mysteries of Go". Now that our unofficial testing is complete, we're looking forward to playing some official, full-length games later this year in collaboration with Go organisations and experts, to explore the profound mysteries of the game further in this spirit of mutual enlightenment. We hope to make further announcements soon!

DeepMind

The Future of Go Summit – 人機圍棋最終決戰

- AlphaGo vs. 柯潔(世界排名第一): 3:0

The Future of Go Summit – 團體賽

- AlphaGo (win) vs. 陳耀燁、周睿羊、芈昱廷、時越、唐韋星（均為9段，曾獲世界冠軍）

Impact of AlphaGo

Not just Go Community!

Not just Computer Go Community!

Not just Artificial Intelligence Community!

Not just Computer Industry!

Even for the Whole Human Community!

The Impact is clearly much higher than Deep Blue.

Why?

- When compared with Deep Blue,
  - "Not much Go domain knowledge is used."
    - Use many general machine learning techniques
      - Deep learning (DL, 深度學習)
      - Reinforcement learning (RL, 強化式學習)
      - Combine DL+RL
  - A big gap to beat human Go champions which most people thought a decade away.
    - As you have seen.

➔

  - More inspiration
  - Higher impact

Open Questions Listed in 2015

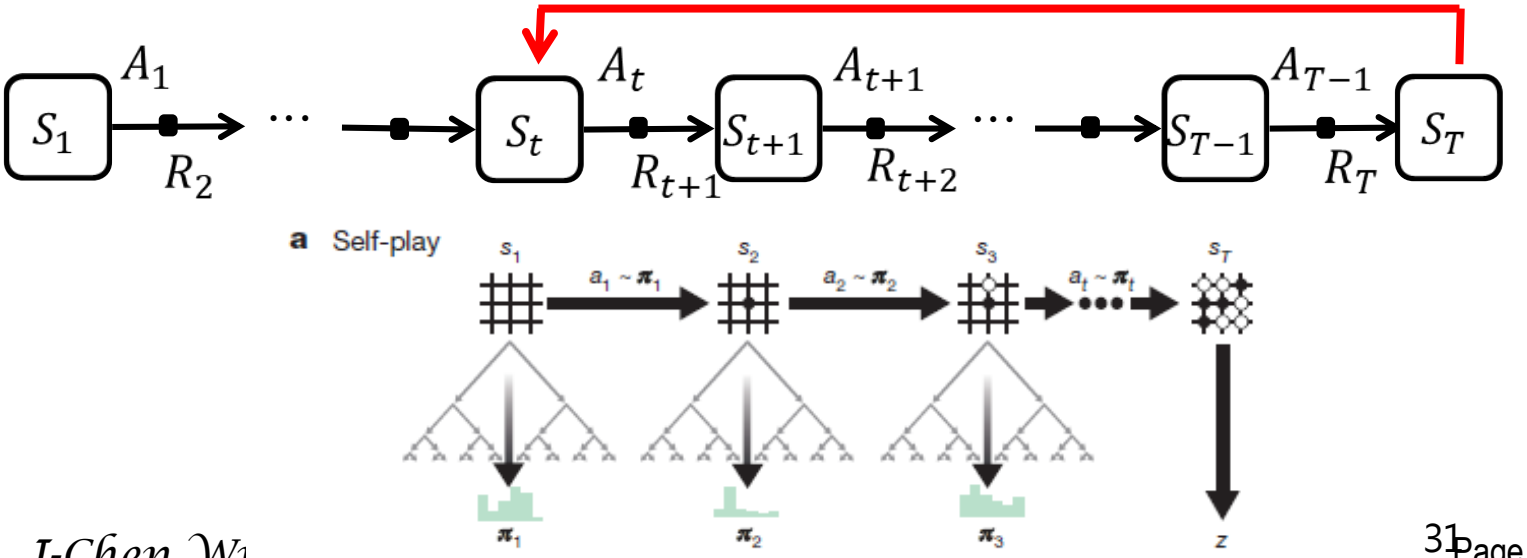
- When can a Go program reach pro (professional) 1-dan?
- When or whether a Go program can beat pro 9-dan or beat human in general ?
- More experiences for other games like Connect6, NoGo, will help solve Go or AI in general?
- How to apply the successful experiences of Go to other domain applications?
  - MoGo team is applying MCTS to
    - Playing Mario, Pac-man, running cars, etc.
    - Solving mathematical optimization, e.g.,
      - Unit commitment problem (UCP): for optimizing electricity power plants. (A successful work by Prof. Kao and by Dr. Teytaud)
      - Flexible Job-Shop Scheduling Problem (FJSP): (A successful work by us)
  - … (expected to have much more)

Major Research Topics in Our Lab

- **Most research topics are related to Deep / Reinforcement Learning (DRL)**
- **Three major classes of DRL applications:**
  1. **Lightweight-Model Applications**
     - Board/Card Games with Zero
  2. **Complex-Model Applications**
     - Video Games Applications
  3. **Real-World-Model Applications**
     - Robotics Applications

Class 1: **Lightweight-Model Applications**

- Properties:
  - Model is well known or tractable
    - E.g., branching factor is limited.
  - Simulator exists (allow backtracking)
- Applications: Games, Education, etc.
- Possible Solutions: AlphaZero-like.

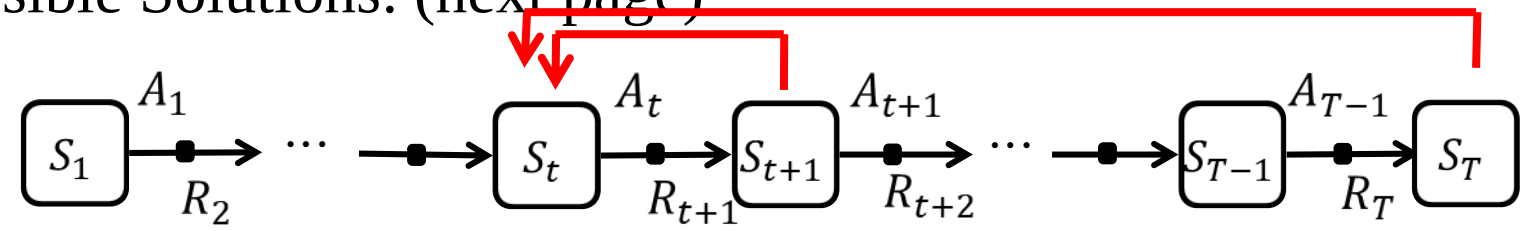


Related DRL Techniques

- Temporal Difference (TD) Learning
- Monte-Carlo (MC) Learning
- Partial Observable Markov Decision Process (POMDP)
- Monte-Carlo Tree Search (MCTS)
- AlphaZero
- …

Class 2: **Complex-Model Applications**

- Properties:
  - Model is well known, but may be complex or intractable
    - E.g., # of actions and environment dynamics are huge or continuous.
  - Simulator exists. (backtracking is hard or costly)
- Applications: Video Games, Robots with Simulator, etc.
- Possible Solutions: (next page)

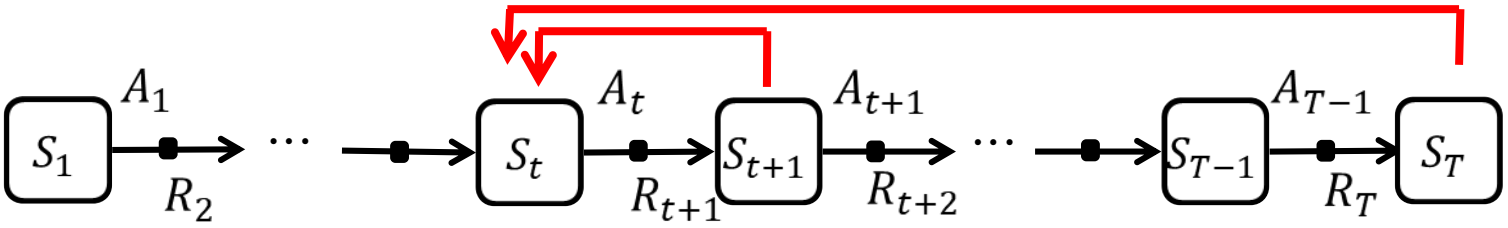


Related DRL Techniques

- Value Based
  - DQN (Deep Q-Network)
  - DDQN (Double DQN)
  - DRQN (Deep Recurrent Q-Network)
  - Dueling Network (with Advantage)
  - Distributional RL (C51, QR-DQN, IQN)
  - Ape-X
- Policy-based & Actor-Critic
  - Actor-Critic (Discrete actions)
  - A3C (Asynchronous Advantage Actor-Critic)
  - DDPG (Deep Deterministic Policy Gradient)
  - TRPO & PPO
- Misc.
  - Random Network Distillation (RND) & Intrinsic Curiosity Module (ICM)
  - GAIL & infoGAIL

Class 3: **Real-World-Model Applications**

- Properties:
  - Model is unknown or too complex
  - Simulator does not exist or runs with expensive costs.
    - So, it is hard to produce a large data set.
- Applications: Robots, Drone, Auto-driving, etc.
- Solutions: (see next page)



Related DRL Techniques

- Curriculum learning
- Imitation Learning
- Behavior Cloning
- Dagger
- Transfer Learning (sim2real)
- Hindsight Experience Replay (HER)
- Meta Learning (one-shot/few-shot)
- …