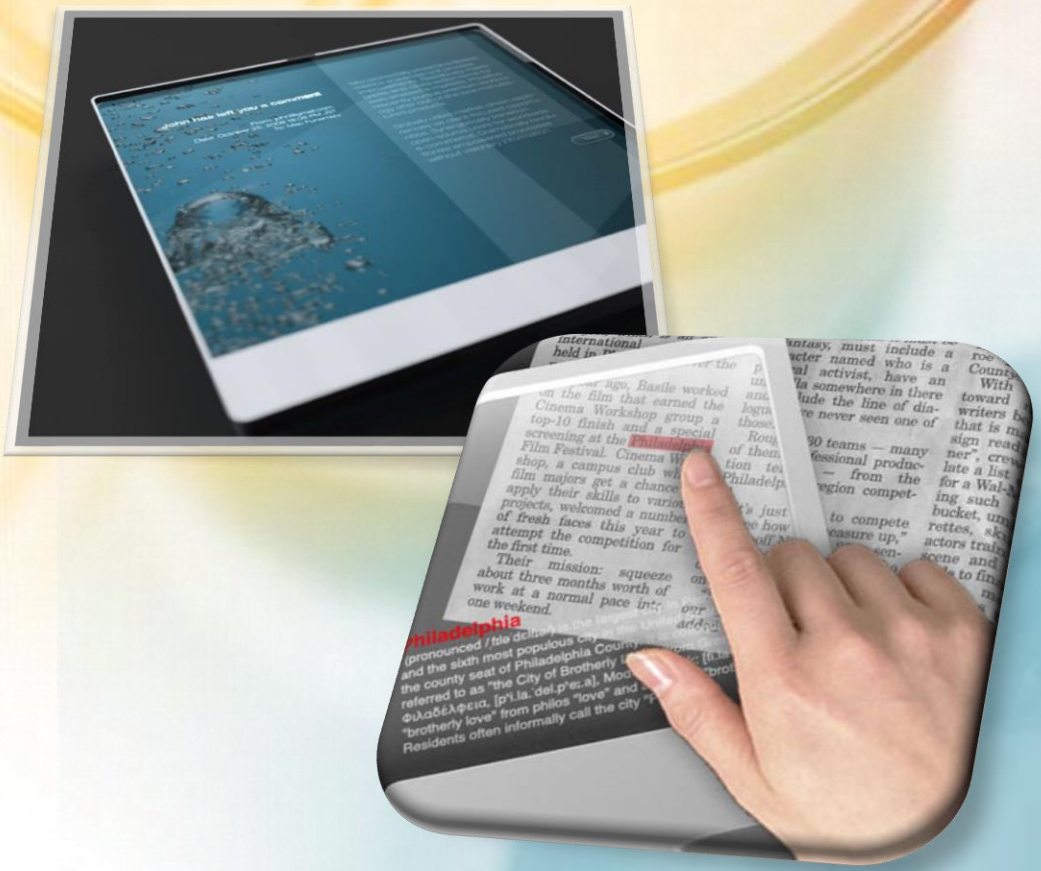
AI大未來



樹德科技大學
研發長
陳毓璋 教授

姓名 ： 陳毓璋

現職職稱 ：

樹德科技大學 資訊工程系教授

兼研究發展處研發長

兼產學營運總中心執行長

兼育成中心主任

學經歷：

- 國立中山大學 資訊工程研究所 博士

  1996-09-01 ~ 1999-06-30

- 中華電信研究所（研究院）

  1999-10 ~ 2004-01

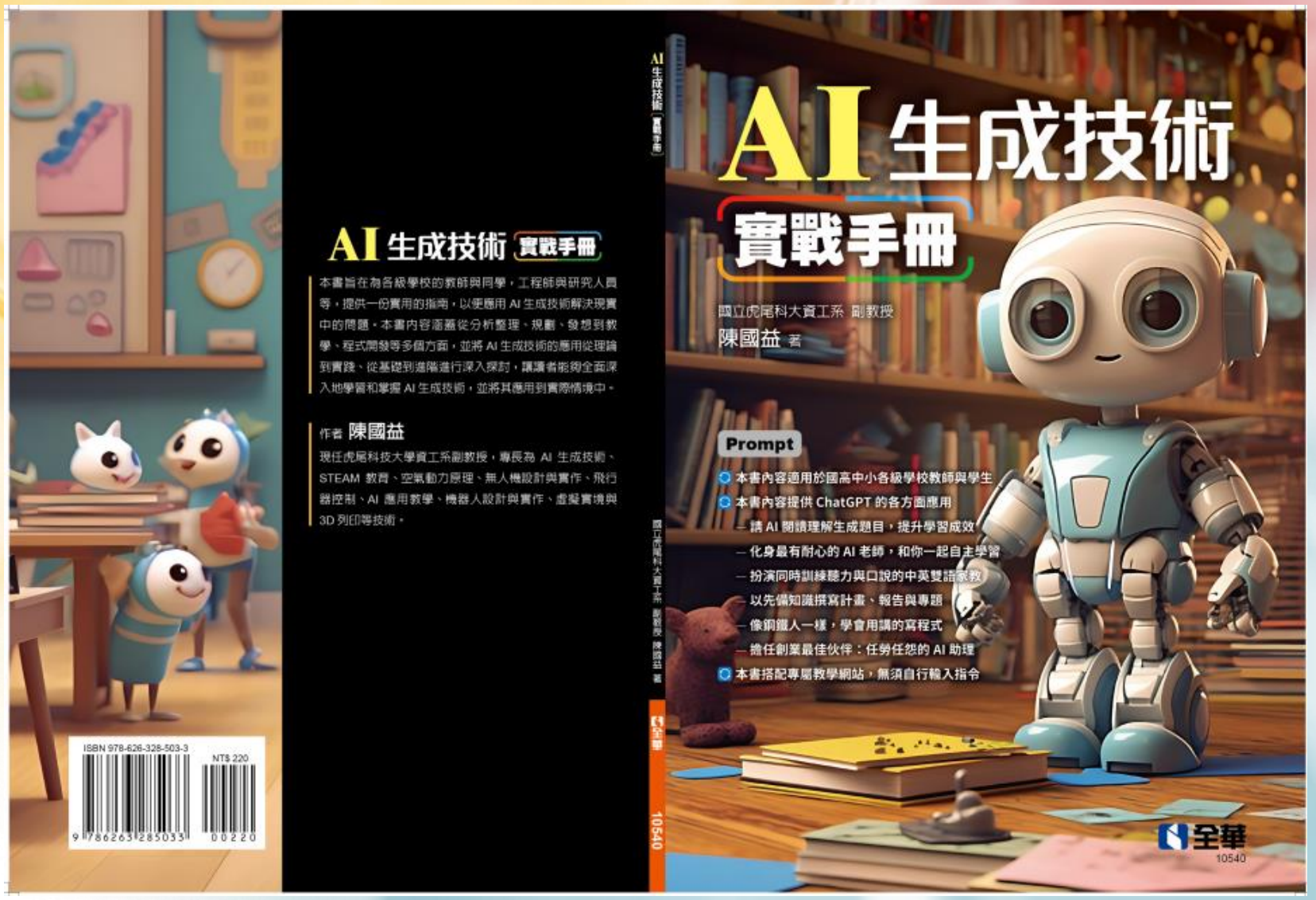
- 樹德科技大學教職

  2004-02 ~ 迄今

引用聲明

- 本報告資料非原創，引用自以下資訊:
- AI生成技術(虎尾科技大學資訊工程系 陳國益老師) <<全華圖書>>
- ChatGPT
- <Mastering the game of Go ***without human knowledge***> (NATURE | VOL 550 | 19 OCTOBER 2017)
- Internet 網路資訊

AI生成技術(虎尾科技大學資訊工程系陳國益老師)<<全華圖書>>



大綱

- 生成式AI
- 人工智慧教學倫理
- **AI奇異點** <Mastering the game of Go *without human knowledge*> **(NATURE | VOL 550 | 19 OCTOBER 2017)**
- 突破框架的思考教育
- AI & 華人之光
- 學生作品

生成式 AI

- 生成式AI在各個領域都有廣泛的應用，其能力包括生成文字、圖像、音頻、視頻等。以下是一些生成式AI的主要應用領域：

自然語言處理（NATURAL LANGUAGE PROCESSING, NLP）

- 文本生成和自動摘要：可以生成文章、新聞、故事，或者創建文本的總結和摘要。
- 對話系統：用於製作聊天機器人，客服自動回覆，或者其他對話型應用。

自然語言處理聊天機器人

圖像生成與編輯

- 圖像生成：能創建逼真的圖像，包括人臉、風景、藝術作品等。
- 圖像編輯：可以幫助修復圖像、改變顏色、創建特殊效果等。
- https://scribblediffusion.com/
- https://firefly.adobe.com/

scribblediffusion.com

開始使用 輪圈與輪胎 - 請推... 從 IE 匯入 部門預算查詢 IHG® Rewards Cl... 新增分頁 課程討論板 加入會員 - 新光人...

a super man

Go

Results

音樂與音頻生成

- 音樂創作：能生成新的音樂、樂曲，或者改編現有音樂。
- 語音合成：可以生成自然的語音，用於語音助手、有聲書等。




音樂與音頻生成

視頻生成與處理

- **視頻生成**： 能創建動畫、短片、影片等。
- **視頻編輯**： 用於視頻修剪、特效添加等。



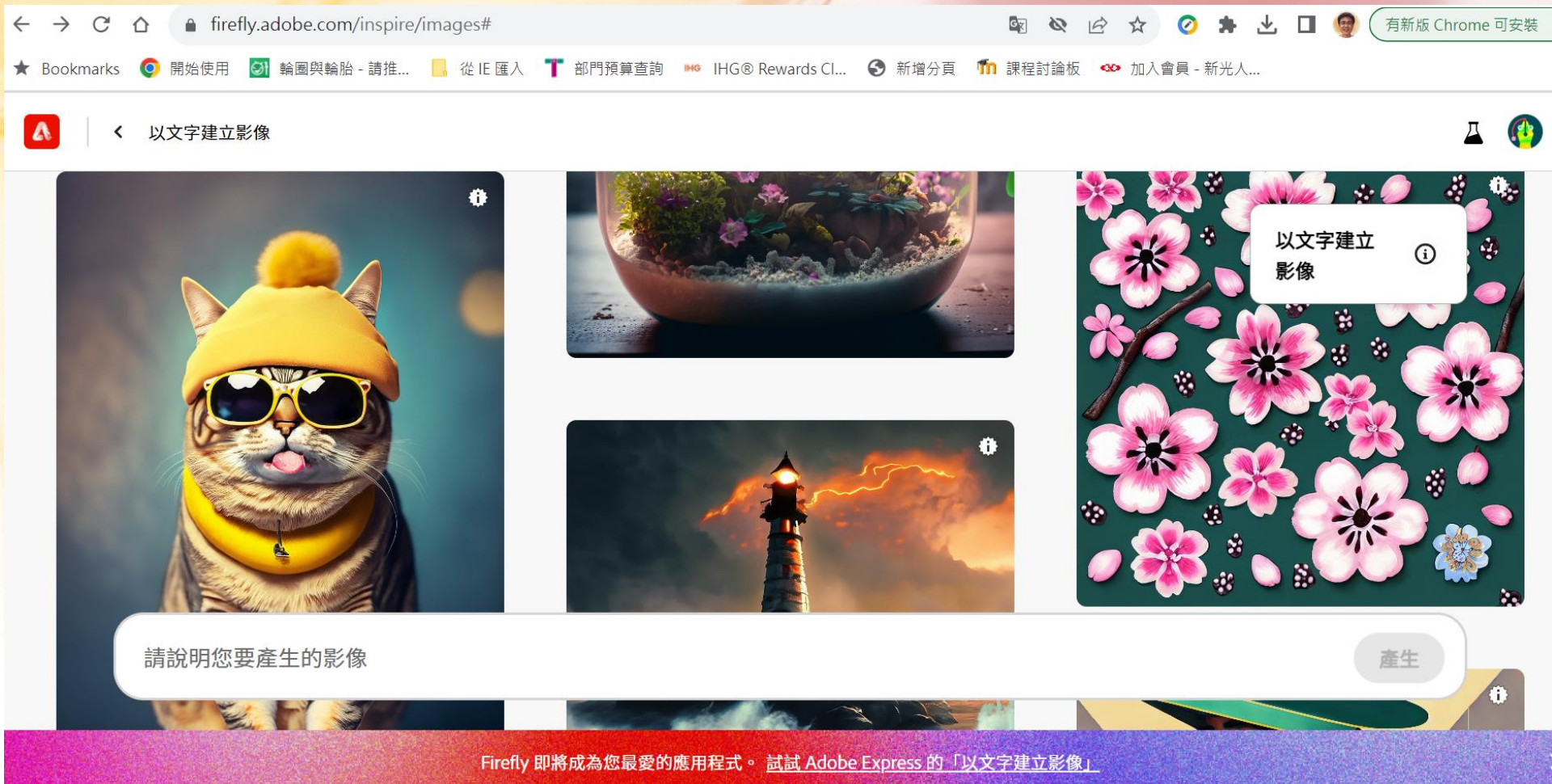

遊戲開發

- **遊戲場景、角色生成：** 可以自動創建遊戲中的場景、角色設計等。

宇宙探險遊戲開發

藝術創作

- **藝術領域：** 生成式AI能夠創作畫作、雕塑、設計等藝術品。



醫療領域

- **疾病預測：** 使用生成模型進行疾病預測，幫助醫生做出更準確的診斷。
- **分子設計：** 在藥物開發方面應用生成模型，設計新的藥物。

人工智慧醫療領域

科學研究

- **科學數據分析與預測**：使用生成模型進行科學數據的分析與預測，幫助研究人員做出更精準的預測。

科學數據分析與預測

生成式人工智慧倫理聲明

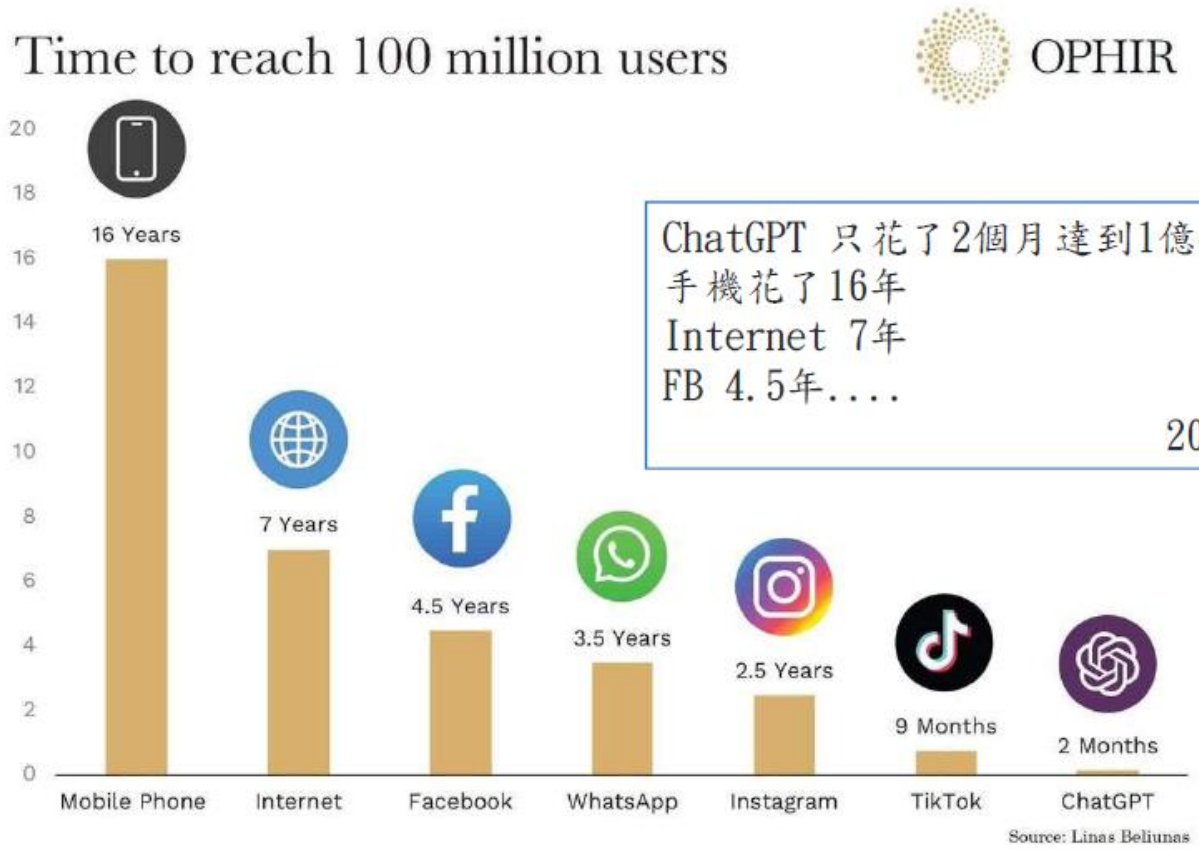
- 國立台灣大學
- 國立清華大學
- 國立陽明交通大學

程式設計

- ChatGPT閱讀過2021年前所有的GitHub程式庫（微軟為OpenAI大股東，並擁有GitHub）
- GitHub是世界上最大的代碼代管網站

- ChatGPT最強的部分，其實是程式設計
- 即使完全沒有程式基礎，也能夠開發自己的程式，只要有想法

ChatGPT 的驚人成長速度



ChatGPT 只花了2個月達到1億使用者
手機花了16年
Internet 7年
FB 4.5年....

2023/02/07

- 類神經網路的運作方式，跟以前傳統的電腦程式很不一樣。傳統的程式，靠的是加減乘除以及邏輯運算，一切都是程式設計師「設計」出來的，所以對於傳統程式的運作原理，程式設計師們自己當然是一清二楚。
- 但類神經網路是模仿人類大腦而建造的，他的能力是「訓練」出來的，而不是「設計」出來的。所以就像身為老師的人，往往搞不懂自己的學生是怎麼想的一樣，身為類神經網路訓練者的工程師們，往往也搞不懂自己做出來的類神經網路究竟是怎麼運作的。

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

- 人腦的思考能力，來自於一些透過突觸（Synapses）互相連結的腦神經元細胞（Neurons）；而人工智慧類神經網路的思考能力，也是來自於一些透過參數（Parameters）互相連結的人造類神經元。

- 人類在學習的時候，腦神經元的突觸連結方式會發生改變；而人造的類神經網路在學習的時候，參數的數值也會發生改變。

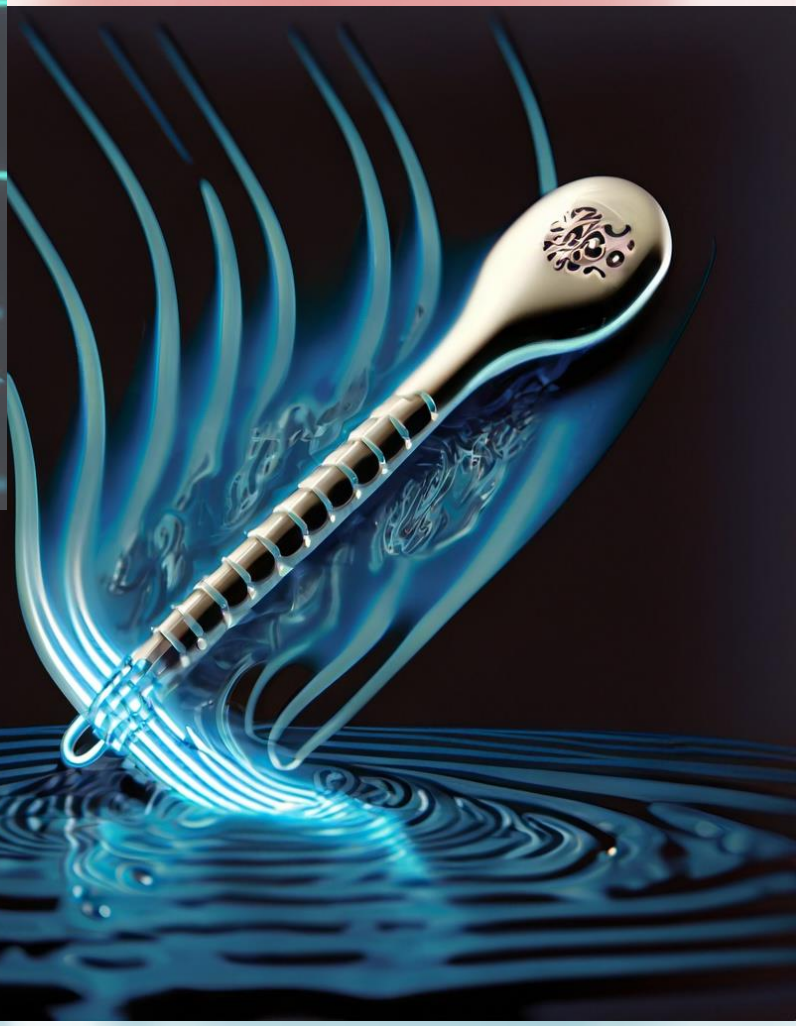
Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

ChatGPT的不同之處

- 目前的Foundation Models超大型類神經網路。像是ChatGPT以及DALL·E背後的GPT-3，就是一個擁有1750億個參數的超大型的類神經網路 。
- 而科學家們估計，人腦大約有100兆個突觸。
- GPT-3.5的1750億個參數跟人腦的100兆個突觸相比，只有五百分之一左右，所以GPT-3的思考記憶能力，終究還是比不上人腦

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx






ChatGPT的不同之處

- 目前訓練 GPT-3 的方式，是讓GPT-3大量閱讀各種從網路上蒐集來的文章，然後叫 GPT-3 寫短文給真人老師批改。

- 然後科學家們為了省事，又去設計訓練出一個「人工智慧家教」（Reward Model），用他來替代真人，專門幫 GPT-3 批改作業，進行訓練。

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

ChatGPT的不同之處

- GPT-3是個好學生，他不會死讀書，他會消化吸收，他會把它學到的知識內化（Internalize），轉化成類神經元上面的參數數值。

- ChatGPT 很老實，當我們問 ChatGPT 問題的時候，它並不會臨時抱佛腳，偷偷上網去找資料，它也不會偷看小抄，把預先寫好的答案整段背給我們聽。ChatGPT 的回答真的都是「GPT-3 自己想的」，「GPT-3 自己記得的」。

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

ChatGPT的不同之處

- 下一代的 GPT-4 已經在 2023/03/14 推出，1.76兆個參數，大約是人腦的 1/50。
- 由 GPT-3、GPT-4 發展脈絡可知，未來 AI 會越來越接近人腦腦容量，成為一個名副其實的「電腦」。
- 然後GPT-4絕對不會是人工智慧研究的終點站，接下來很可能會有人腦 500 倍容量的 GPT-5、人腦 25 萬倍容量的 GPT-6。

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

ChatGPT的不同之處

- 一年之後，研究所的學生都不用再自己寫論文了，因為論文可以交給 GPT-4 寫，而且寫的絕對比大部分的研究生都好；

- 同樣的，大學的教授們也不用再浪費時間上課、指導論文了，因為反正學生的作業跟論文都是 GPT-4 寫的，教授們不如把 GPT-4 複製一份，直接去教 GPT-4 做研究、寫論文；

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

ChatGPT的不同之處

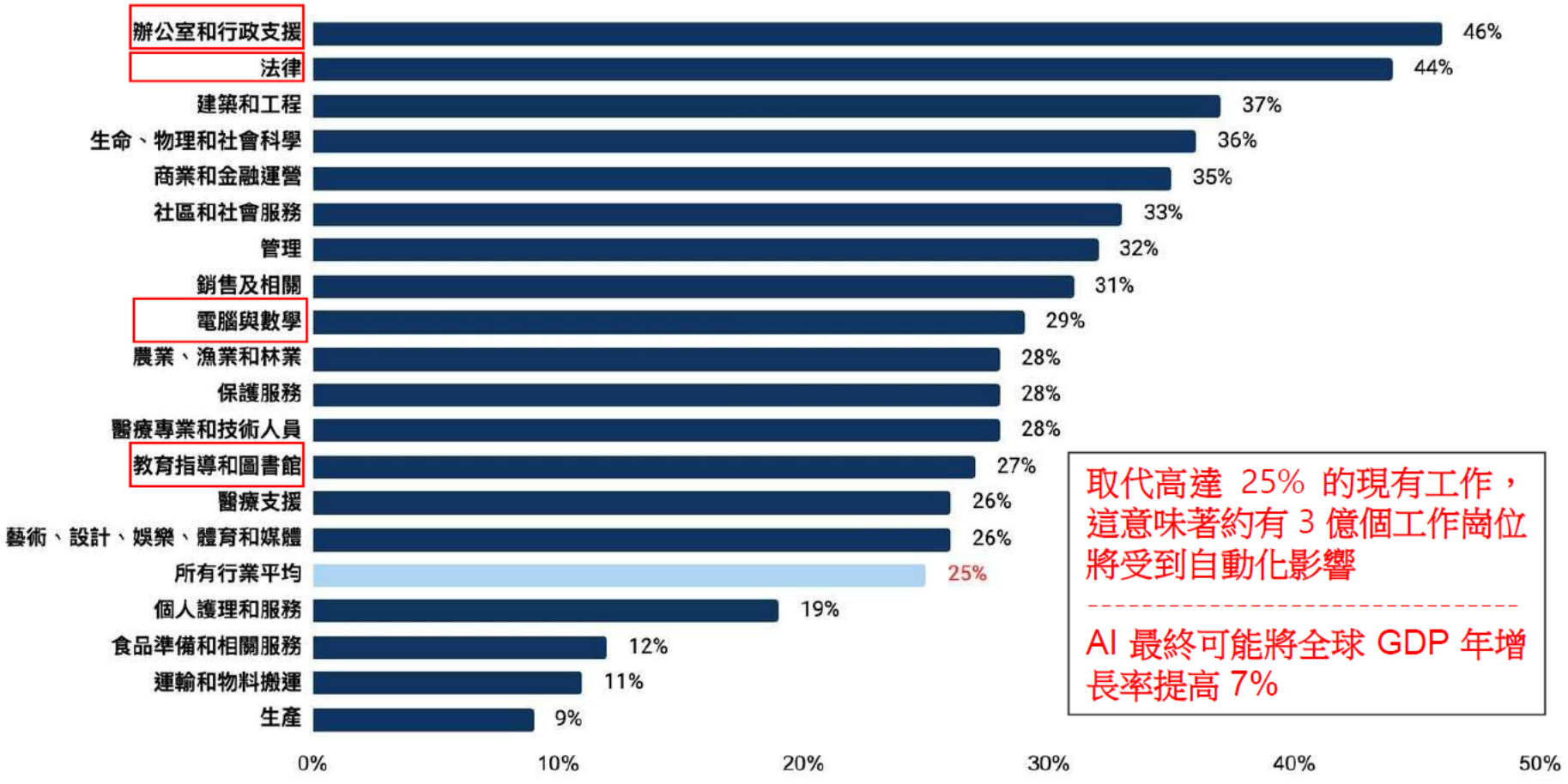
- 然後教授們訓練出來的GPT-4、GPT-5、GPT-6會越來越厲害，厲害到他們青出於藍，總有一天會比當初訓練出他們的教授們還厲害，可以自己做研究、自己訓練下一代的GPT-7、而且去跟Google、百度訓練出來的類神經網路進行學術討論了。

- 所以我們很可能是最後一代。我們可能是最後一代的真人研究生、最後一代的真人指導教授、最後一代的真人學者。

Ref:作者：林宜敬【我們是最後一代】https://reurl.cc/rZzqjx

預估受到 AI 影響的就業比例（%）

地點：美國（資料來源：Goldman Sachs）



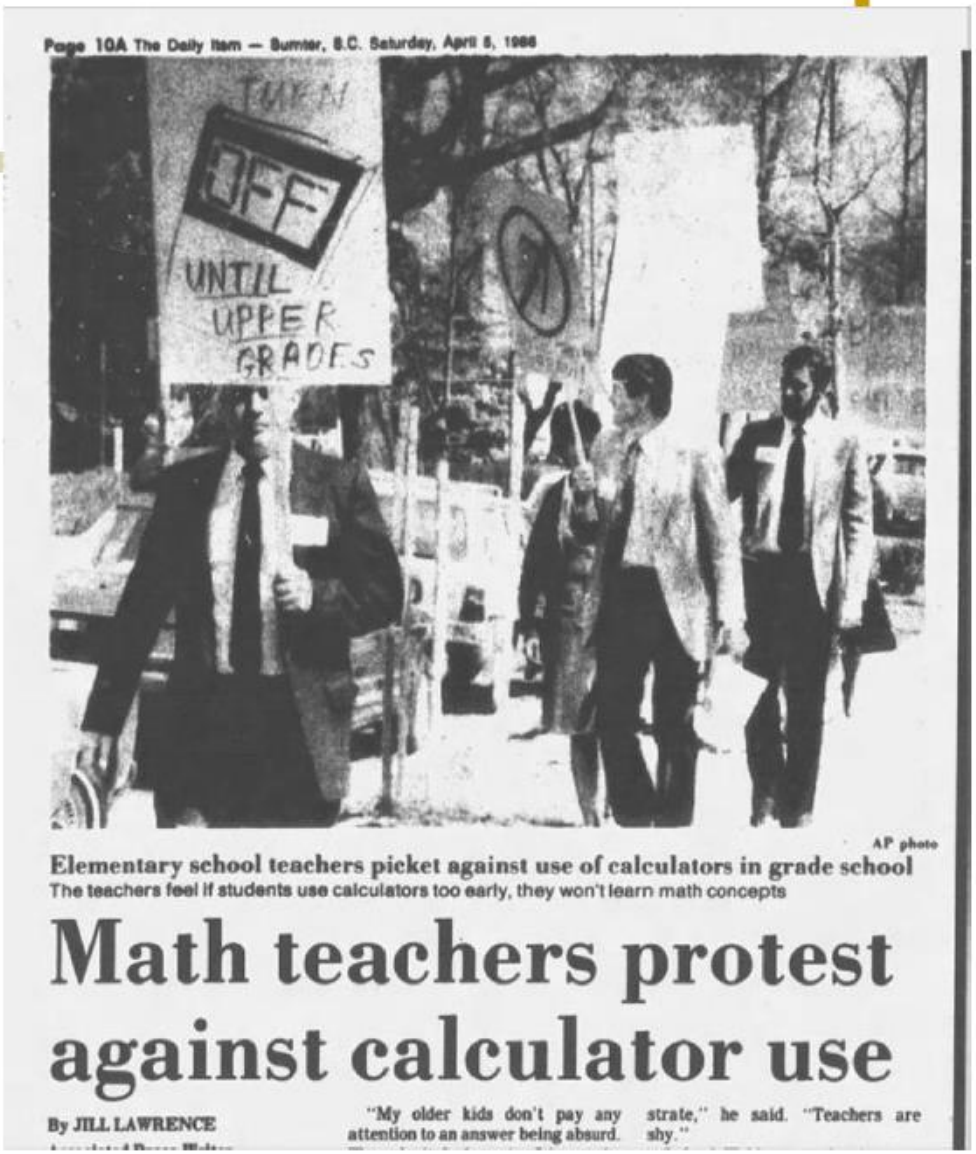
取代高達 25% 的現有工作，這意味著約有 3 億個工作崗位將受到自動化影響

AI 最終可能將全球 GDP 年增長率提高 7%

科技奇異點

- 1988 年：數學老師上街抗議、反對使用計算機
- 2023 年：（部分）老師反對使用 ChatGPT

- 擁抱科技變革，這對你來說就是最好的時代
- 忽視科技進步，這對你來說就是最壞的時代



Page 10A The Daily Item — Sumter, S.C. Saturday, April 5, 1986

AP photo

Elementary school teachers picket against use of calculators in grade school

The teachers feel if students use calculators too early, they won't learn math concepts

Math teachers protest against calculator use

By JILL LAWRENCE

"My older kids don't pay any attention to an answer being absurd. strate," he said. "Teachers are shy."


海峰棋院

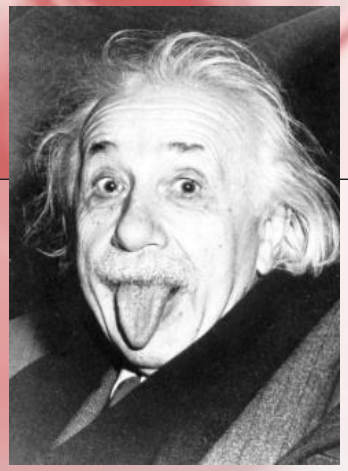
台灣棋王許皓鋐的奇蹟

- ~~4歲學棋、15歲打敗世界冠軍 台灣棋王許皓鋐的奇蹟~~
- .一夜之間，全台灣都知道了許皓鋐這個名字，他奪下亞運金牌，成為台灣圍棋史上第一人。
- 年紀幼小的他，有著紅通通的雙頰，在棋院內被暱稱為「蘋果」，小一的蘋果便達到五段實力，12歲晉升職業棋士，被視為神童。
- 柯潔在中國是明星棋手，1997年出生於浙江麗水的他，個性外放、自信，活躍於社交媒體，常起爭議，如2016年，AlphaGo對弈韓國棋王李世乭，引起世界關注，柯潔則在微博上寫道「就算阿法狗戰勝了李世石，但它贏不了我」（但10個月後柯潔在非正式網路快棋中三次敗於AlphaGo升級版）。
- 春蘭盃敗給柯潔後，許皓鋐持續努力，2022年，他連續拿下海峰盃、名人賽、十段賽，快棋賽、棋王賽冠軍，一年總獎金就高達850萬元，成為名符其實的棋王。
- 2023年，許皓鋐終於成功打敗「柯潔大帝」，成為世界冠軍，「柯潔銀牌」衝上微博熱搜第一，許皓鋐則被大批網友封為「許仙」，這個總維持一號表情、氣質沈穩的男孩，成為了台灣圍棋史上的一大驚奇。

許皓鋐 (2023/09 亞運圍棋金牌)

- 棋靈王真的存在，他是一個台灣人。
- 他叫許皓鋐。
- 他22歲，來自台灣，他是世界排名第35名的圍棋手。
- 他剛獲得亞運金牌。
- 媒體形容許皓鋐的這個金牌「大爆冷門」。
- 許皓鋐不是靠賽制、槓桿而奪冠，而是與世界頂尖三大高手直球對決，過關斬將，一一擊退，所以這個「金牌純度極高」。
- 在此之前，他在台灣的九大冠已經八冠在手，原本就已經是台灣最強。
- 但圍棋在台灣關注的人不多，這是個很專業，但很小眾的圈子，新聞報導，台灣的棋士，只有108人。
- 在全世界的Go Ratings圍棋排名裡，世界棋手共有2461人，中、韓、日，佔最大比例。

教育出具創造力的下一代



- 剛剛的題目告訴我們：

  要跳出限制思考的框架…

  - 擺脫人類經驗限制：AlphaGo Zero 三天學習打敗了被人類誤導的AlphaGo，AI人工智慧利用自我學習突破棋局的限制，這篇英文論文 <Mastering the game of Go ***without human knowledge***>（登上 NATURE 論文發表 | VOL 550 | 19 OCTOBER 2017）

    圍棋複雜度:

    （考量黑、白、空白)3^361 = 1.74089650659031927907188238070 56 e+172

  - （只考量黑、白）2^361 = 4.697085165547666455778961193578 7e+108

  - 愛因斯坦曾經被認為是自閉低能，也是聯考的失敗者… 但是他的思考邏輯卻遠超過凡人！

MASTERING THE GAME OF GO WITHOUT HUMAN KNOWLEDGE

- A long-standing goal of artificial intelligence is an algorithm that learns, tabula rasa, superhuman proficiency in challenging domains. Recently, AlphaGo became the first program to defeat a world champion in the game of Go. The tree search in AlphaGo evaluated positions and selected moves using deep neural networks. These neural networks were trained by supervised learning from human expert moves, and by reinforcement learning from self-play. Here we introduce an algorithm based solely on reinforcement learning, without human data, guidance or domain knowledge beyond game rules. AlphaGo becomes its own teacher: a neural network is trained to predict AlphaGo's own move selections and also the winner of AlphaGo's games. This neural network improves the strength of the tree search, resulting in higher quality move selection and stronger self-play in the next iteration. Starting tabula rasa, our new program AlphaGo Zero achieved superhuman performance, winning 100–0 against the previously published, champion-defeating AlphaGo.

- 人工智慧長久以來的目標之一，是能夠在具有挑戰性的領域中，透過空白學習的演算法，達到超越人類的高度能力。最近，AlphaGo 成為第一個在圍棋比賽中擊敗世界冠軍的程式。AlphaGo 中的樹狀搜索評估不同棋局的局面，並選擇著步著，這是透過深度神經網路實現的。這些神經網路透過從人類專家的棋著中進行監督式學習，以及從自我對弈中進行強化學習而進行訓練。
- 在這裡，我們介紹一種完全基於強化學習的演算法，不倚賴人類的資料、指導或領域知識，僅僅根據遊戲規則運作。AlphaGo 成為了自己的老師：一個神經網路被訓練來預測 AlphaGo 自身的棋步選擇，以及 AlphaGo 比賽的勝者。這個神經網路提升了樹狀搜索的能力，進而在下一次迭代中實現了更高品質的棋步選擇以及更強大的自我對弈能力。從空白開始，我們的**新程式 AlphaGo Zero 達到了超越人類水準的表現**，以 100 比 0 的成績擊敗了之前發表的、曾經擊敗過冠軍的 AlphaGo。

- Much progress towards artificial intelligence has been made using supervised learning systems that are trained to replicate the decisions of human experts1–4. However, expert data sets are often expensive, unreliable or simply unavailable. Even when reliable data sets are available, they may impose a ceiling on the performance of systems trained in this manner5. By contrast, reinforcement learning systems are trained from their own experience, in principle allowing them to exceed human capabilities, and to operate in domains where human expertise is lacking. Recently, there has been rapid progress towards this goal, using deep neural networks trained by reinforcement learning. These systems have outperformed humans in computer games, such as Atari6,7 and 3D virtual environments8–10. However, the most chal lenging domains in terms of human intellect—such as the game of Go, widely viewed as a grand challenge for artificial intelligence11—require a precise and sophisticated lookahead in vast search spaces. Fully gene ral methods have not previously achieved human-level performance in these domains.

- 透過受過監督學習訓練的系統，已取得了許多人工智慧方面的進展，這些系統旨在模仿人類專家的決策。然而，專家資料集通常昂貴、不可靠，或者根本無法獲得。即使可靠的資料集可用，它們可能限制了以此方式訓練的系統的表現上限。相比之下，強化學習系統從自身的經驗中進行訓練，原則上允許它們超越人類的能力，並在缺乏人類專業知識的領域中運作。最近，在使用強化學習進行訓練的深度神經網路方面已取得了快速進展。這些系統在計算機遊戲（例如 Atari6、7 和 3D 虛擬環境）方面優於人類。然而，在以人類智慧為基準的最具挑戰性領域，如被廣泛視為人工智慧重大挑戰的圍棋遊戲，需要在龐大的搜索空間中進行精確而複雜的前瞻。完全通用的方法以前在這些領域中並未達到人類水準的表現。

- AlphaGo was the first program to achieve superhuman performance in Go. The published version12, which we refer to as AlphaGo Fan, defeated the European champion Fan Hui in October 2015. AlphaGo Fan used two deep neural networks: a policy network that outputs move probabilities and a value network that outputs a position eval uation. The policy network was trained initially by supervised learn ing to accurately predict human expert moves, and was subsequently refined by policy-gradient reinforcement learning. The value network was trained to predict the winner of games played by the policy net work against itself. Once trained, these networks were combined with a Monte Carlo tree search (MCTS)13–15 to provide a lookahead search, using the policy network to narrow down the search to high-probability moves, and using the value network (in conjunction with Monte Carlo rollouts using a fast rollout policy) to evaluate positions in the tree. A subsequent version, which we refer to as AlphaGo Lee, used a similar approach (see Methods), and defeated Lee Sedol, the winner of 18 inter national titles, in March 2016.

- AlphaGo 是第一個在圍棋遊戲中達到超越人類表現的程式。我們稱之為 AlphaGo Fan 的已發表版本於2015年10月擊敗了歐洲冠軍范輝。AlphaGo Fan 使用了兩個深度神經網路：一個策略網路，輸出棋步概率；以及一個價值網路，輸出局面評估。策略網路最初通過監督學習來準確預測人類專家的棋步，然後通過策略梯度強化學習進行了進一步的優化。價值網路則被訓練來預測策略網路對弈自己時的比賽勝者。一旦訓練完畢，這些網路被結合到蒙特卡洛樹搜索（Monte Carlo tree search，MCTS)中，提供了一個前瞻搜索，使用策略網路來縮小搜索範圍到高概率的棋步，並使用價值網路（與使用快速推演策略的蒙特卡洛推演相結合）來評估樹中的局面。後來的版本，我們稱之為 AlphaGo Lee，採用了類似的方法（詳見方法一節），於2016年3月擊敗了贏得18個國際冠軍頭銜的李世石

- Our program, AlphaGo Zero, differs from AlphaGo Fan and AlphaGo Lee12 in several important aspects. First and foremost, it is trained solely by self-play reinforcement learning, starting from ran dom play, without any supervision or use of human data. Second, it uses only the black and white stones from the board as input features. Third, it uses a single neural network, rather than separate policy and value networks. Finally, it uses a simpler tree search that relies upon this single neural network to evaluate positions and sample moves, without performing any Monte Carlo rollouts. To achieve these results, we introduce a new reinforcement learning algorithm that incorporates lookahead search inside the training loop, resulting in rapid improve ment and precise and stable learning. Further technical differences in the search algorithm, training procedure and network architecture are described in Methods.

- 我們的程式 AlphaGo Zero 在幾個重要的方面與 AlphaGo Fan 和 AlphaGo Lee有所不同。首先，它僅通過自我對弈的強化學習進行訓練，從隨機對弈開始，不需要任何監督或使用人類數據。其次，它僅使用棋盤上的黑白子作為輸入特徵。第三，它使用單個神經網路，而不是獨立的策略和價值網路。最後，它使用一個更簡單的樹狀搜索，依賴這個單個神經網路來評估局面並選擇棋步，而不執行任何傳統蒙特卡洛推演。為了實現這些結果，我們引入了一種新的強化學習演算法，在訓練循環內部結合前瞻搜索，從而實現快速改進和精確穩定的學習。在搜尋演算法、訓練過程和網路架構方面，還有進一步的技術差異，詳情請參考方法一節。

- Reinforcement learning in AlphaGo ZeroOur new method uses a deep neural network fθ with parameters θ. This neural network takes as an input the raw board representation s of the position and its history, and outputs both move probabilities and a value, (p, v) = fθ(s). The vector of move probabilities p represents the probability of selecting each move a (including pass), pa = Pr(a | s). The value v is a scalar evaluation, estimating the probability of the current player winning from position s. This neural network combines the roles of both policy network and value network12 into a single architecture. The neural network consists of many residual blocks4 of convolutional layers16,17 with batch normalization18 and rectifier nonlinearities19 (see Methods).

- AlphaGo Zero 中的強化學習方法：我們的新方法使用具有參數 θ 的深度神經網路 $f\theta$。這個神經網路以局面的原始棋盤表示 s 及其歷史作為輸入，並輸出移動概率和一個值，即 $(p, v) = f\theta(s)$。移動概率向量 p 表示選擇每個棋步 a（包括過）的概率，$pa = Pr(a| s)$。值 v 是一個標量評估，估計了當前玩家從局面 s 出發的勝利概率。這個神經網路將策略網路和價值網路的角色合併成單一架構。神經網路由多個帶有卷積層、批量正規化和整流非線性的殘差塊組成（詳見方法一節）。

- The neural network in AlphaGo Zero is trained from games of selfplay by a novel reinforcement learning algorithm. In each position s, an MCTS search is executed, guided by the neural network fΘ. The MCTS search outputs probabilities π of playing each move. These search probabilities usually select much stronger moves than the raw move probabilities p of the neural network fΘ(s); MCTS may therefore be viewed as a powerful policy improvement operator20,21. Self-play with search—using the improved MCTS-based policy to select each move, then using the game winner z as a sample of the value—may be viewed as a powerful policy evaluation operator.The main idea of our reinforcement learning algorithm is to use these search operatorsrepeatedly in a policy iteration procedure22,23: the neural network's parameters are updated to make the move probabilities and value $(p, v) = f_\Theta(s)$ more closely match the improved search probabilities and self-play winner (π, z); these new parameters are used in the next iteration of self-play to make the search even stronger. Figure 1 illustrates the self-play training pipeline.

- AlphaGo Zero 中的神經網路是通過一種新穎的強化學習演算法從自我對弈的遊戲中進行訓練的。在每個局面 s 中，執行一次由神經網路 $f\theta$ 引導的 MCTS 搜索。MCTS 搜索輸出了每個棋步的播放概率 π。這些搜索概率通常選擇比神經網路 $f\theta(s)$ 的原始棋步概率 p 更強大的棋步；因此，MCTS 可以被看作是一種強大的策略改進運算子。使用改進的基於 MCTS 的策略來選擇每個棋步，然後使用遊戲勝者 z 作為值的樣本，可以被看作是一種強大的策略評估運算子。
- 我們強化學習演算法的主要思想是在策略迭代過程中重複使用這些搜索運算子：更新神經網路的參數，使棋步概率和值 $(p, v) = f\theta(s)$ 與改進的搜索概率和自我對弈的勝者 (π, z) 更加吻合；這些新參數在下一次自我對弈的迭代中被用來使搜索變得更加強大。圖 1 展示了自我對弈訓練流程。

- MCTS may be viewed as a self-play algorithm that, given neural network parameters θ and a root position s, computes a vector of search probabilities recommending moves to play, π = αθ(s), proportional to the exponentiated visit count for each move, πa ∝ N(s, a)1/τ, where τ is a temperature parameter.The neural network is trained by a self-play reinforcement learning algorithm that uses MCTS to play each move. First, the neural network is initialized to random weights θ0. At each subsequent iteration i ≥ 1, games of self-play are generated (Fig. 1a). At each time-step t, an MCTS search π α=θ−s()tti 1 is executed using the previous iteration of neural network θ−fi 1 and a move is played by sampling the search probabilities πt. A game terminates at step T when both players pass, when the search value drops below a resignation threshold or when the game exceeds a maximum length; the game is then scored to give a final reward of rT ∈ {− 1,+ 1} (see Methods for details). The data for each time-step t is stored as (st, πt, zt), where zt = ± rT is the game winner from the perspective of the current player at step t. In parallel (Fig. 1b), new network parameters θi are trained from data (s, π, z) sampled uniformly among all time-steps of the last iteration(s) of self-play. The neural network =θp vfs(,)()i is adjusted to minimize the error between the predicted value v and the self-play winner z, and to maximize the similarity of the neural network move probabilities p to the search probabilities π. Specifically, the parameters θ are adjusted by gradient descent on a loss function l that sums over the mean-squared error and cross-entropy losses, where c is a parameter controlling the level of L2 weight regularization (to prevent overfitting).

$$(\boldsymbol{p}, v) = f_\theta(s) \text{ and } l = (z - v)^2 - \boldsymbol{\pi}^{\mathrm{T}} \log \boldsymbol{p} + c\|\theta\|^2 \qquad (1)$$

- 進化式 MCTS 可被視為一個自我對弈演算法，根據神經網路參數 θ 和根局面 s，計算一個建議下棋步的搜索概率向量，π = αθ(s)，與每個棋步的指數化訪問計數成正比，πa ∝ N(s, a)1/τ，其中 τ 是一個溫度參數。神經網路通過一個自我對弈強化學習演算法進行訓練，該演算法使用 MCTS 來下每一步棋。首先，神經網路初始化為隨機權重 θ0。在每個後續的迭代 i ≥ 1 中，生成自我對弈遊戲（圖 1a）。在每個時間步 t，使用上一次迭代的神經網路 θ−fi 1 進行 MCTS 搜索 π α=θ−s()tti 1，並通過採樣搜索概率 πt 來下一步棋。當遊戲在步數 T 終止時，當兩名玩家通過，當搜索值低於放棄閾值或者當遊戲超過最大長度時，遊戲終止，然後對遊戲進行評分以給出最終獎勵 rT ∈ {− 1,+ 1}（詳見方法中的細節）。每個時間步 t 的數據存儲為（st, πt, zt），其中 zt = ± rT 是從當前玩家角度在步 t 上的遊戲勝者。同時（圖 1b），從上次自我對弈迭代的所有時間步中均勻抽樣的數據（s, π, z）來訓練新的網路參數 θi。神經網路 =θp vfs(,)()i 通過調整來最小化預測值 v 與自我對弈勝者 z 之間的誤差，並最大化神經網路棋步概率 p 與搜索概率 π 之間的相似性。具體而言，參數 θ 通過對損失函數 l 進行梯度下降來進行調整，該函數分別對均方誤差損失和交叉熵損失求和：，其中 c 是控制 L2 權重正則化程度的參數（以防止過度擬合）。

$$(\boldsymbol{p}, v) = f_\theta(s) \text{ and } l = (z - v)^2 - \boldsymbol{\pi}^{\mathrm{T}} \log \boldsymbol{p} + c\|\theta\|^2 \qquad (1)$$

- Empirical analysis of AlphaGo Zero training We applied our reinforcement learning pipeline to train our program AlphaGo Zero. Training started from completely random behaviour and continued without human intervention for approximately three days. Over the course of training, 4.9 million games of self-play were gen erated, using 1,600 simulations for each MCTS, which corresponds to approximately 0.4 s thinking time per move. Parameters were updated from 700,000 mini-batches of 2,048 positions. The neural network contained 20 residual blocks (see Methods for further details).

- AlphaGo Zero 訓練的實證分析：我們將強化學習流程應用於我們的程式 AlphaGo Zero 的訓練。訓練從完全隨機的行為開始，並在約三天的時間內持續進行，無需人為干預。在訓練過程中，共生成了 490 萬場自我對弈遊戲，平均每秒56場以上訓練。

- To assess the merits of self-play reinforcement learning, compared to learning from human data, we trained a second neural network (using the same architecture) to predict expert moves in the KGS Server data set; this achieved state-of-the-art prediction accuracy compared to pre vious work12,30–33 (see Extended Data Tables 1 and 2 for current and previous results, respectively). Supervised learning achieved a better initial performance, and was better at predicting human professional moves (Fig. 3). Notably, although supervised learning achieved higher move prediction accuracy, the self-learned player performed much better overall, defeating the human-trained player within the first 24 h of training. This suggests that AlphaGo Zero may be learning a strategy that is qualitatively different to human play.

- 為了評估自我對弈強化學習相對於從人類數據學習的優勢，我們使用相同的結構訓練了第二個神經網路，來預測 KGS 伺服器數據集中的專家棋步；這在預測準確性方面達到了與之前工作相比的最新狀態（請參考擴展數據表格 1 和 2，分別為當前和之前的結果）。監督學習在初始性能方面表現得更好，並且在預測人類專業棋步方面表現更好（圖 3）。值得注意的是，儘管監督學習實現了更高的棋步預測準確性，但自我學習的程式在整體上表現要好得多，在訓練的前 24 小時內就擊敗了從人類數據學習的程式。這表明 AlphaGo Zero 可能正在學習一種與人類下棋不同質的策略。

- To separate the contributions of architecture and algorithm, we compared the performance of the neural network architecture in AlphaGo Zero with the previous neural network architecture used in AlphaGo Lee (see Fig. 4). Four neural networks were created, using either separate policy and value networks, as were used in AlphaGo Lee, or combined policy and value networks, as used in AlphaGo Zero; and using either the convolutional network architecture from AlphaGo Lee or the residual network architecture from AlphaGo Zero. Each network was trained to minimize the same loss function (equation (1)), using a fixed dataset of self-play games generated by AlphaGo Zero after 72 h of self-play training. Using a residual network was more accurate, achieved lower error and improved performance in AlphaGo by over 600 Elo. Combining policy and value together into a single network slightly reduced the move prediction accuracy, but reduced the value error and boosted playing performance in AlphaGo by around another 600 Elo. This is partly due to improved computational effi ciency, but more importantly the dual objective regularizes the network to a common representation that supports multiple use cases.

- 為了分離架構和演算法的貢獻，我們將 AlphaGo Zero 中的神經網路架構與之前在 AlphaGo Lee 中使用的神經網路架構進行了比較（參見圖 4）。我們創建了四個神經網路，分別使用了與 AlphaGo Lee 中使用的獨立策略和價值網路相同的結構，或者使用了 AlphaGo Zero 中使用的結合策略和價值的網路結構；同時，我們使用了 AlphaGo Lee 中的卷積網路結構或 AlphaGo Zero 中的殘差網路結構。每個神經網路都被訓練來最小化相同的損失函數（方程（1）），使用 AlphaGo Zero 在進行了 72 小時的自我對弈訓練後生成的固定數據集的自我對弈遊戲。
- 使用殘差網路更準確，錯誤率更低，將 AlphaGo 的性能提高了超過 600 Elo。將策略和價值結合到一個單一網路中稍微降低了棋步預測的準確性，但降低了值的誤差，並且使 AlphaGo 的遊戲表現提高了另外約 600 Elo。這部分是由於改進的計算效率，但更重要的是雙重目標使網路規範化為一個支援多種用例的共同表示。(PS: Elo 是下棋能力等級)

- Knowledge learned by AlphaGo Zero AlphaGo Zero discovered a remarkable level of Go knowledge dur ing its self-play training process. This included not only fundamental elements of human Go knowledge, but also non-standard strategies beyond the scope of traditional Go knowledge. Figure 5 shows a timeline indicating when professional joseki (corner sequences) were discovered (Fig. 5a and Extended Data Fig. 2); ultimately AlphaGo Zero preferred new joseki variants that were previously unknown (Fig. 5b and Extended Data Fig. 3). Figure 5c shows several fast self-play games played at different stages of train ing (see Supplementary Information). Tournament length games played at regular intervals throughout training are shown in Extended Data Fig. 4 and in the Supplementary Information. AlphaGo Zero rapidly progressed from entirely random moves towards a sophisti cated understanding of Go concepts, including fuseki (opening), tesuji (tactics), life-and-death, ko (repeated board situations), yose (endgame), capturing races, sente (initiative), shape, influence and territory, all discovered from first principles. Surprisingly, shicho (‘ladder’ capture sequences that may span the whole board)—one of the first elements of Go knowledge learned by humans—were only understood by AlphaGo Zero much later in training.

- AlphaGo Zero 在自我對弈的訓練過程中學到了一個非常出色的圍棋知識水平。這不僅包括人類圍棋知識的基本要素，還包括了傳統圍棋知識範圍之外的非標準策略。圖 5 顯示了專業角式（棋盤角落的一系列走法）的發現時間軸（見圖 5a 和擴展數據圖 2）；最終，AlphaGo Zero 更偏好之前未知的新角式變體（見圖 5b 和擴展數據圖 3）。圖 5c 顯示了在不同訓練階段進行的幾場快速自我對弈遊戲（詳見補充信息）。整個訓練過程中定期進行的錦標賽長局對局在擴展數據圖 4 和補充信息中展示。
- AlphaGo Zero 從完全隨機的棋步迅速進展到了對圍棋概念的精湛理解，包括布局（開局）、手筋（戰術）、活死（生死）、打吃（反複的局面）、余生（終局）、吃子比賽、先手（主動性）、形狀、影響和領地等，全部都是從基本原理中發現的。令人驚訝的是，對於人類最早學到的圍棋知識要素之一 ——“天梯”（可能橫跨整個棋盤的捉子系列）——AlphaGo Zero 在訓練的較晚階段才理解。

- Final performance of AlphaGo Zero We subsequently applied our reinforcement learning pipeline to a second instance of AlphaGo Zero using a larger neural network and over a longer duration. Training again started from completely random behaviour and continued for approximately 40 days. Over the course of training, 29 million games of self-play were gener ated. Parameters were updated from 3.1 million mini-batches of 2,048 positions each. The neural network contained 40 residual blocks. The learning curve is shown in Fig. 6a. Games played at regular intervals throughout training are shown in Extended Data Fig. 5 and in the Supplementary Information. We evaluated the fully trained AlphaGo Zero using an internal tournament against AlphaGo Fan, AlphaGo Lee and several previous Go programs. We also played games against the strongest existing program, AlphaGo Master—a program based on the algorithm and architecture presented in this paper but using human data and fea tures (see Methods)—which defeated the strongest human professional players 60–0 in online games in January 201734. In our evaluation, all programs were allowed 5 s of thinking time per move; AlphaGo Zero and AlphaGo Master each played on a single machine with 4 TPUs; AlphaGo Fan and AlphaGo Lee were distributed over 176 GPUs and 48 TPUs, respectively. We also included a player based solely on the raw neural network of AlphaGo Zero; this player simply selected the move with maximum probability.

- AlphaGo Zero 的最終表現：我們隨後將我們的強化學習流程應用於 AlphaGo Zero 的第二個實例，使用了更大的神經網路並且訓練時間更長。訓練再次從完全隨機的行為開始，並持續約 40 天。在訓練過程中，共生成了 2900 萬場自我對弈遊戲。參數來自於 310 萬個大小為 2,048 的小批次的局面。神經網路包含 40 個殘差塊。學習曲線如圖 6a 所示。整個訓練過程中定期進行的對局在擴展數據圖 5 和補充信息中展示。我們對完全訓練好的 AlphaGo Zero 進行了內部錦標賽，與 AlphaGo Fan、AlphaGo Lee 以及幾個先前的圍棋程式進行了對局。我們還與最強的現有程式 AlphaGo Master 進行了對局，該程式基於本文中介紹的演算法和結構，但使用了人類數據和特徵（詳見方法），在 2017 年 1 月的在線對局中以 60-0 擊敗了最強的人類專業選手。
- 在我們的評估中，所有程式每步被允許 5 秒的思考時間；AlphaGo Zero 和 AlphaGo Master 每個在一台機器上使用 4 個 TPU 進行對局；AlphaGo Fan 和 AlphaGo Lee 則分別使用了 176 個 GPU 和 48 個 TPU 進行了對局。我們還包括了一個僅基於 AlphaGo Zero 的原始神經網路的選手；該選手僅選擇具有最大概率的棋步。

- Fig.6 shows the performance of each program on an Elo scale. The raw neural network, without using any lookahead, achieved an Elo rating of 3,055. AlphaGo Zero achieved a rating of 5,185, compared to 4,858 for AlphaGo Master, 3,739 for AlphaGo Lee and 3,144 for AlphaGo Fan. Finally, we evaluated AlphaGo Zero head to head against AlphaGo Master in a 100-game match with 2-h time controls. AlphaGo Zero won by 89 games to 11 (see Extended Data Fig. 6 and Supplementary Information).

- 圖 6 顯示了每個程式在 Elo 評分上的表現。僅使用原始神經網路而不進行任何預先搜索的程式達到了 3,055 的 Elo 評分。AlphaGo Zero 的評分為 5,185，而 AlphaGo Master 為 4,858，AlphaGo Lee 為 3,739，AlphaGo Fan 為 3,144。最後，我們將 AlphaGo Zero 與 AlphaGo Master 進行了一場 100 局的對局。AlphaGo Zero 以 89 勝 11 負取得勝利（詳見擴展數據圖 6 和補充信息）。

- Conclusion Our results comprehensively demonstrate that a pure reinforcement learning approach is fully feasible, even in the most challenging of domains: it is possible to train to superhuman level, without human examples or guidance, given no knowledge of the domain beyond basic rules. Furthermore, a pure reinforcement learning approach requires just a few more hours to train, and achieves much better asymptotic performance, compared to training on human expert data. Using this approach, AlphaGo Zero defeated the strongest previous versions of AlphaGo, which were trained from human data using handcrafted fea tures, by a large margin. Humankind has accumulated Go knowledge from millions of games played over thousands of years, collectively distilled into patterns, prov erbs and books. In the space of a few days, starting tabula rasa, AlphaGo Zero was able to rediscover much of this Go knowledge, as well as novel strategies that provide new insights into the oldest of games.

- 結論：我們的結果全面地展示了純粹的強化學習方法在最具挑戰性的領域中是完全可行的：在沒有人類示例或指導的情況下，即使對領域的知識僅限於基本規則，也有可能訓練到超越人類的水平。此外，與使用人類專家數據進行訓練相比，純粹的強化學習方法只需要多幾個小時的訓練時間，並且在漸進性能方面取得了更好的成績。使用這種方法，AlphaGo Zero 以極大的優勢擊敗了之前由人類數據訓練並使用手工特徵的 AlphaGo 最強版本。數千年來，人類從數百萬場對局中積累了圍棋知識，這些知識被集體地提煉成了模式、格言和書籍。在短短幾天內，從零開始，AlphaGo Zero 能夠重新發現這部分圍棋知識，並提出了新的策略，為這個古老的遊戲帶來新的見解。

未來的希望在 … 教育

- 學思達教育「share start」訓練學生自「學」、閱讀、「思」考、討論、分析、歸納、表「達」、寫作等等能力。

教育出具創造力的下一代

- 我是否具創造力也能接受創造力 ？
  - 未來的希望在下一代 -> 下一代的希望在教育

    How can you encourage（stimulate）a child ？

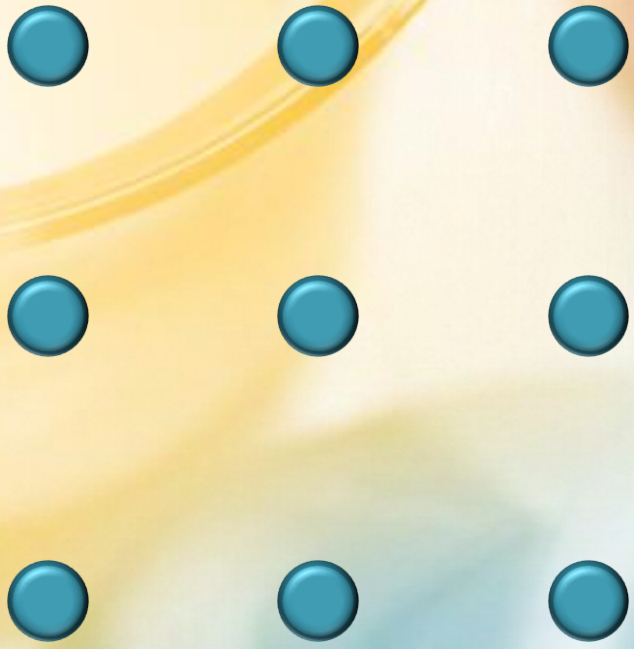
教育出具創造力的下一代

短片：Use your imagination.

教育出具創造力的下一代

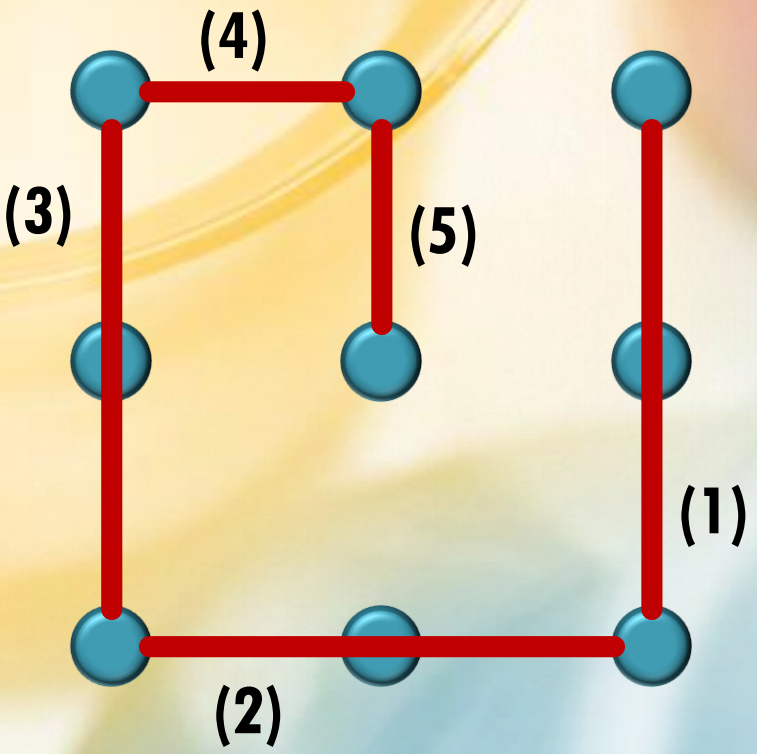
- 接著有兩個題目，大家一起來動動腦吧！

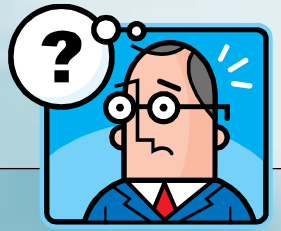
有九個點，請用一筆劃通過那九個點，直線可以交叉，但限定只能用四條直線。



有九個點，請用一筆劃通過那九個點，直線可以交叉，但限定只能用四條直線。

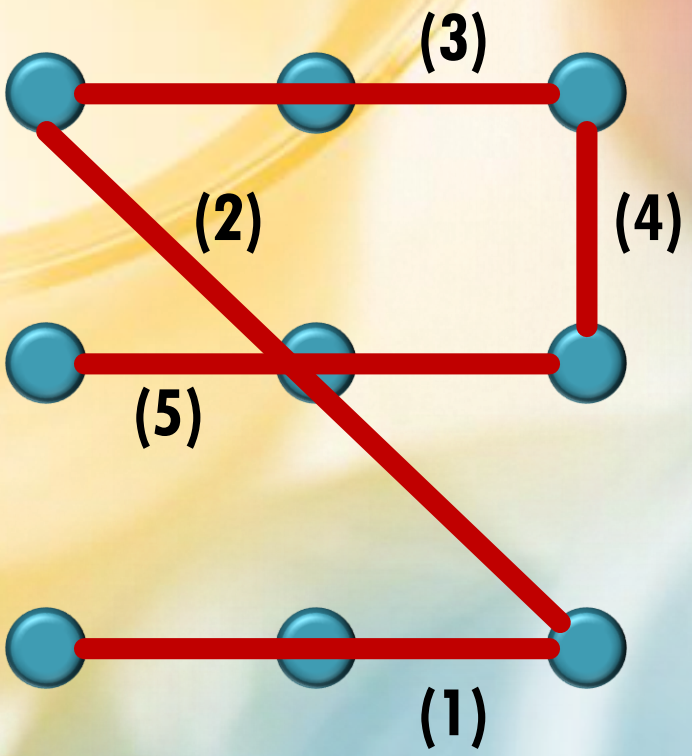
試試看

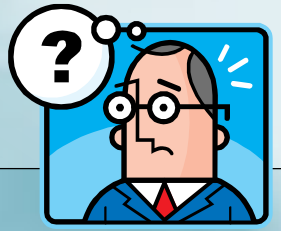




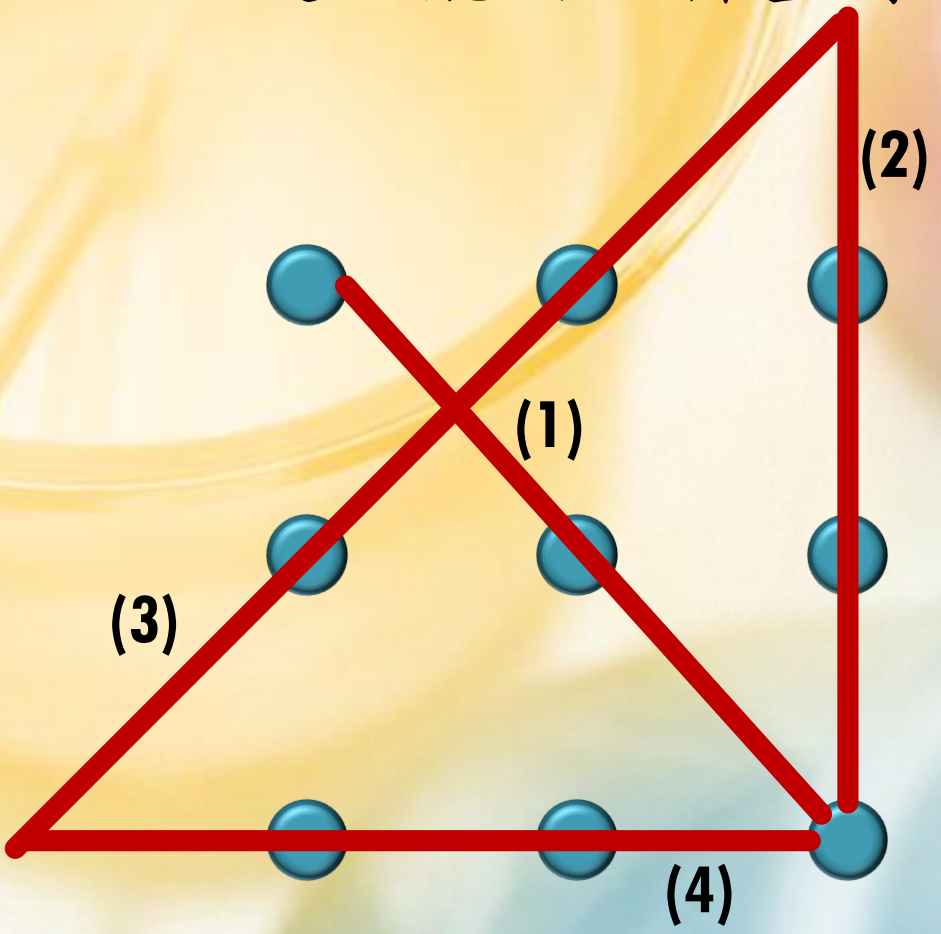
有九個點，請用一筆劃通過那九個點，直線可以交叉，但限定只能用四條直線。

試試看





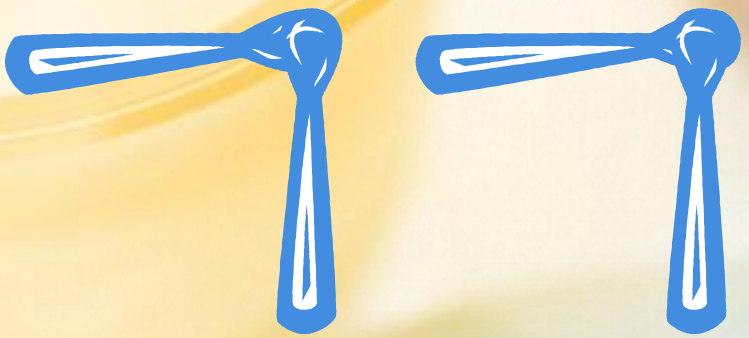
有九個點，請用一筆劃通過那九個點，直線可以交叉，但限定只能用四條直線。



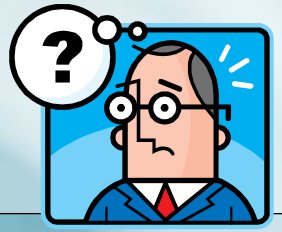
這裡有四根火柴棒，你可以任意擺放它，請排出一個最大的數字。

這裡有四根火柴棒，你可以任意擺放它，請排出一個最大的數字。

試試看



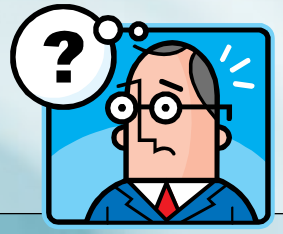
77



這裡有四根火柴棒，你可以任意擺放它，請排出一個最大的數字。

試試看

1111



這裡有四根火柴棒，你可以任意擺放它，請排出一個最大的數字。

正解

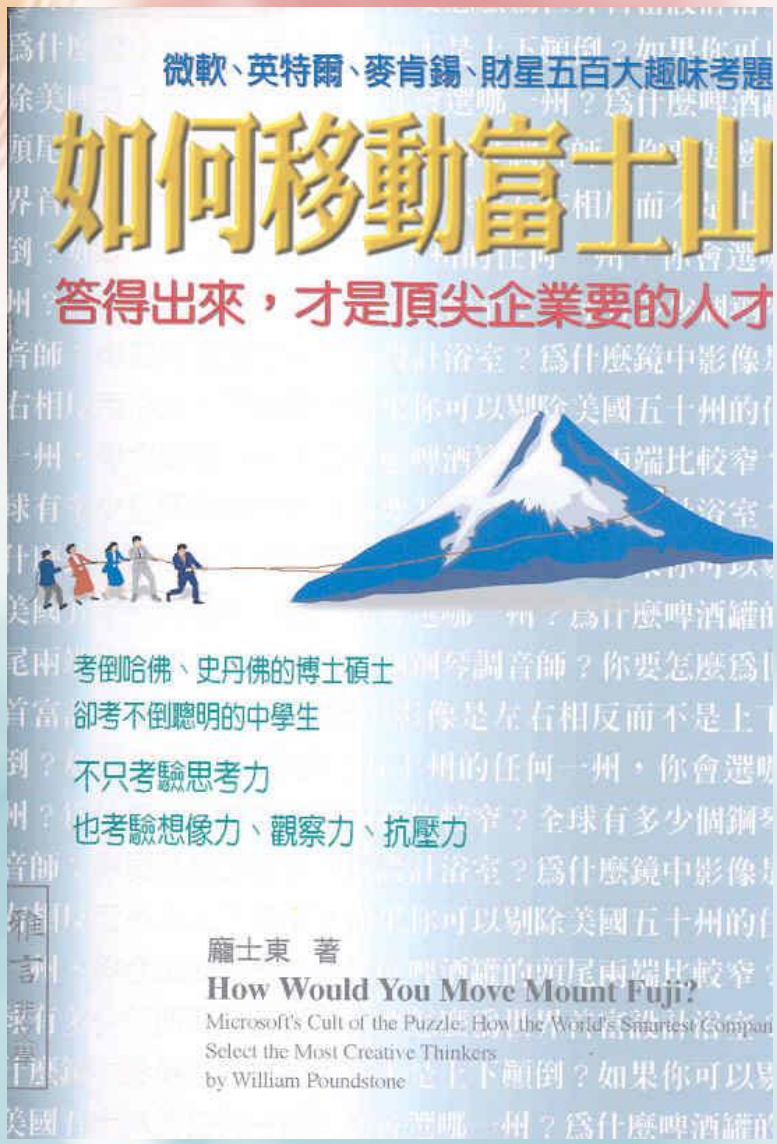
11^{11}

11的11次方：11

這個值是285,311,670,611

教育出具創造力的下一代

- 如何移動富士山？



完美的答案 --- 跳躍式思考

- 狂風暴雨的夜晚，您開車經過一條黑暗的道路，路上有三個人在等車，一位是您夢中情人、一位是風燭殘年病弱的老婦人、一位是您多年好友且是救命恩人，但是您的車子只能再多載一個人，請問您會如何選擇 ？

- 企業徵才 --- 完美的答案

資訊教育的未來（2006~2049）

- 2006年，位在科羅拉多州的一所默默無聞的公立高中Arapahoe剛結束暑假；新學期開始的時候，校長請該校科技中心的負責人Karl Fisch為老師們解說一下目前教育界的資訊技術發展和趨勢。

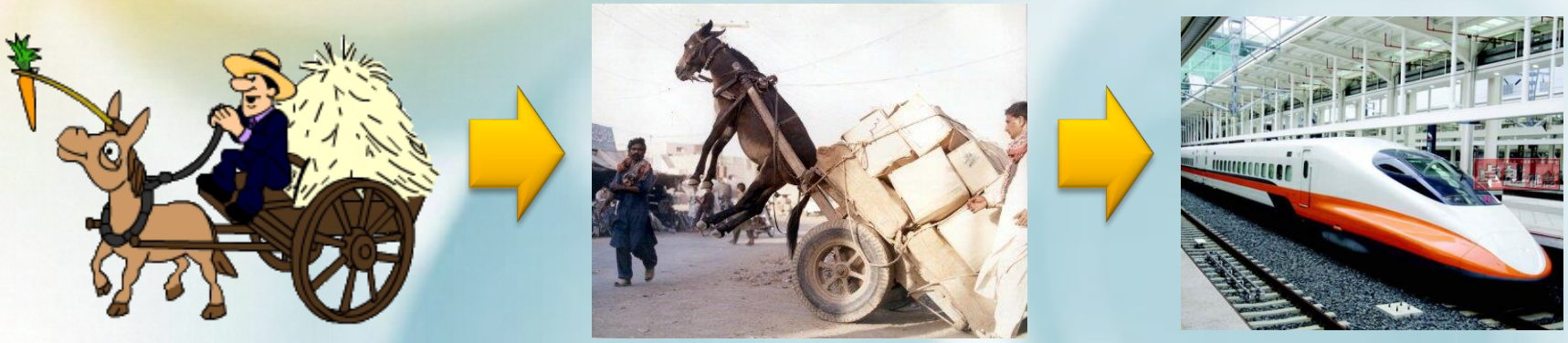
  Karl從書籍、網路、政府資料中整理出了一些數據，並且用淺顯易懂的比喻加入投影片中；在做好影片，配上音樂之後，他把這個影片取名為「Did you know?」。

資訊教育的未來（2006~2049）

- 這些數據來自於「世界是平的」作者、教育界的知名人士、美國前教育部長、美國勞工部、麻省理工學院等等……
- 這份報告後來被翻譯十種以上語言，至今總瀏覽次數突破一億次。

DID YOU KNOW?

- 我們必須教導現在的學生，畢業後投入目前還不存在的工作...
  使用根本還沒發明的科技...
  解決我們從未想像過的問題。
- 美國前教育部長Richard Riley認為...
  2010年最迫切需要的十種工作，在2004年時根本不存在。新的科技知識，每2年增加一倍，大學生前2年所學的知識，到大三就已經過時，這些資訊到2010年時，甚至每72小時就增加一倍。所以，我們必須教導現在的學生，畢業後投入目前還不存在的工作（而非當前的熱門工作）…
- 根據估計，《紐約時報》一週所產生的資訊量...
  比十八世紀一個人一生可能接觸到的資訊量還要多。



DID YOU KNOW?

- 2013年已發展出超越人類大腦運算能力的電腦！
- 2017年AlphaGo打敗人類圍棋棋王
- 2023年售價1000美金的電腦運算能力就可以超越人腦！
- 2049年售價1000美金的電腦運算能力就可以超越全人類大腦運算速度的總合！
- 當我們已經知道這些之後…，**然後呢？**

未來生活

- 未來的生活 1
- 未來的生活 2
- 未來的生活 3

突破框架的阿里巴巴

- 阿里巴巴創辦人 – 馬雲
  - 從一個月入不過89元人民幣的英文老師，變成全球最大電子商務網站阿里巴巴的經營者
  - 如果阿里巴巴從世界上消失，那可能至少會有數百萬個中小企業因而破產、倒閉，數千萬人失業

突破框架的阿里巴巴

- 馬雲自認成功的關鍵在：**缺錢**、**不懂技術**、**不做計劃**、**夢想不設限**（精彩募資過程）
  - **因為缺錢，所以花每一分錢都必須非常小心**，小器沒什麼不好，公司的錢就是要很小器，一點一點地去花才對，節儉就會成為公司文化。
  - **因為不懂技術，所以尊重專業**。
  - 至於不做計劃，這可能讓人感到不可思議。計劃寫得再漂亮，遇到環境變動便失去意義，**有沒有即時應變的能力才是重點**。西元兩千年網路泡沫，那些計畫做的很好，但是應變能力差的公司都消失了。

– **今天很痛苦，明天更痛苦，後天很美好。但絕大多數的人會死在明天晚上。**

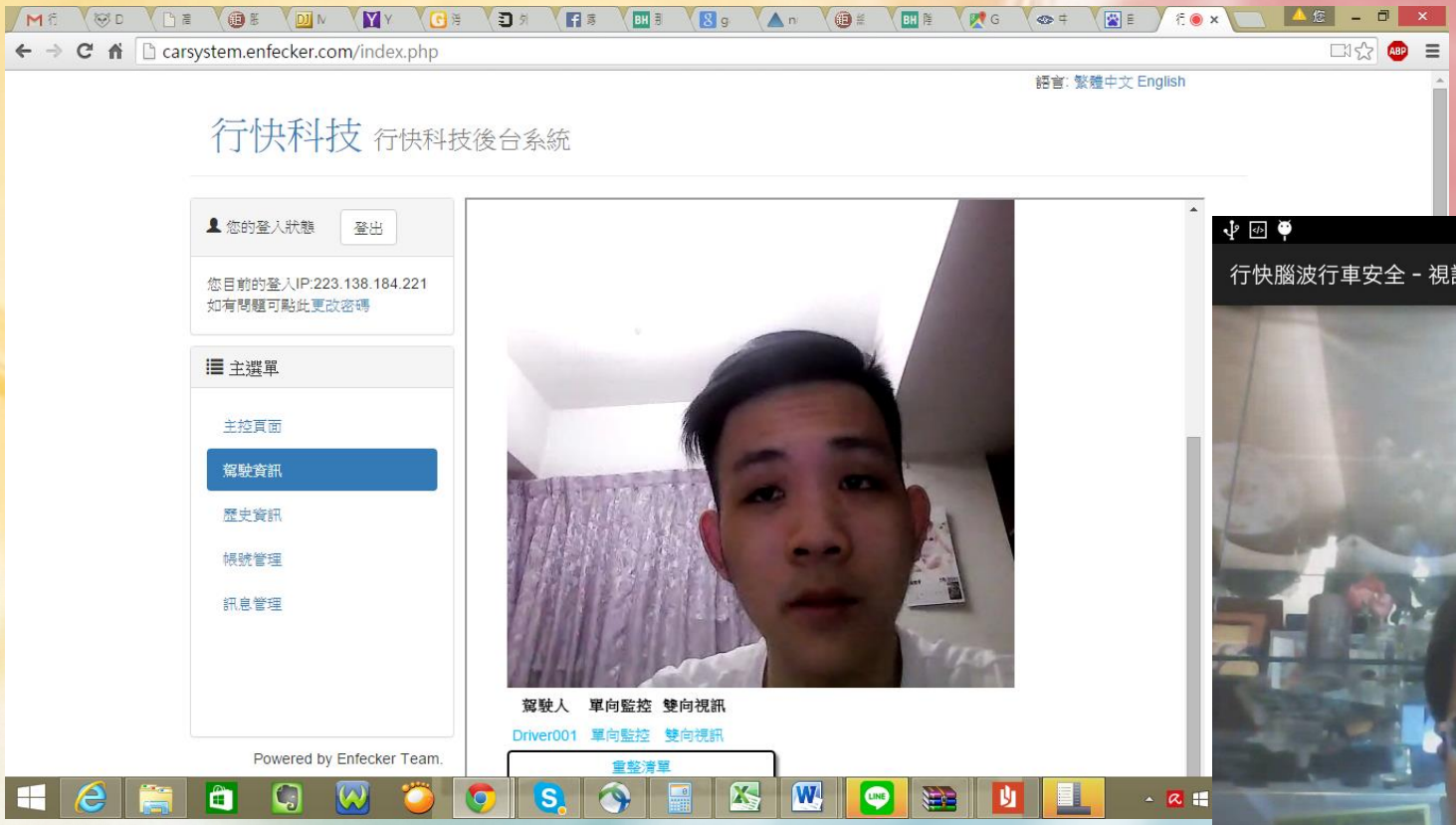
為什麼華人總是在美國發光

- NVIDIA、AMD、YOUTUBE、YAHOO

葉丙成：「台灣教育一直塞東西，養成一堆虛胖的學生」越讓孩子餓，讓他們去“黑白舞”，學得就越多……

- 學生作品(一)雲端腦波偵測安全駕駛系統
- 完成專利技術移轉上市汽車集團

腦波手機 APP 開發


carsystem.enfecker.com/index.php
語言: 繁體中文 English
行快科技 行快科技後台系統
您的登入狀態
登出
您目前的登入IP:223.138.184.221
如有問題可點此更改密碼
主選單
主控頁面
駕駛資訊
歷史資訊
帳號管理
訊息管理
Powered by Enfecker Team.
駕駛人 單向監控 雙向視訊
Driver001 單向監控 雙向視訊
重整清單


97% 20:25
▼35.0KB/S
行快腦波行車安全 - 視訊擴充元件
Connecting.

行車中控系統開發

語言: 繁體中文 English

行車中控系統開發


語言: 繁體中文 English
行快科技 行快科技後台系統
您的登入狀態
登出
您目前的登入IP:
如有問題可點此更改密碼
主選單
主控頁面
駕駛資訊
歷史資訊
帳號管理
訊息管理
駕駛資訊
駕駛人帳號 行車危險指數 行車不專注指數 電池 連線狀態
Driver001 67 57 72% 斷訊
Driver002 59 51 85% 斷訊
Driver003 無資料 無資料 無電量資料 離線
Driver004 無資料 無資料 無電量資料 離線
Driver005 無資料 無資料 無電量資料 離線
Driver006 無資料 無資料 無電量資料 離線
Driver007 無資料 無資料 無電量資料 離線
Driver008 70 59 58% 斷訊
Driver009 無資料 無資料 無電量資料 離線
Driver010 離線

行車中控系統開發


語言: 繁體中文 English
行快科技 行快科技後台系統
您的登入狀態
登出
您目前的登入IP.
如有問題可點此更改密碼
主選單
主控頁面
駕駛資訊
歷史資訊
帳號管理
訊息管理
駕駛資訊
修改靈敏度
修改資料
視訊通話
發送訊息
駕駛人：Driver002
姓名：
性別：
生日：
年資：
車牌：
電話：
住址：
血型：
59
51
駕駛目前位置
地圖
衛星檢視
台中市
花蓮縣
台灣
台南市
高雄市
Google
地圖資料©2015 AutoNavi, Google, ZENRIN
使用條款

創業成功案例

- BigGo 比價網

勉勵同學的話

做你不敢做的事叫突破

做你不願意做的事叫改變

做你沒做過的事叫成長

凡事不要說“我不會”
或“不可能”
因爲你根本還没有去做