

電腦有意識嗎？

電腦的智能極限在哪裡？它能和人類對答如流，但真的能夠感知外界嗎？用幾張圖考考它吧。

撰文／柯霍（Christof Koch）、托諾尼（Giulio Tononi）

翻譯／王怡文

電腦發展得更接近有智慧的人類了，瞧瞧IBM的華生是怎麼打敗電視益智問答節目「危險邊緣」（Jeopardy）的人類常勝軍。然而，儘管電腦擁有以超人速度消化資料的優異能力，但是大多數人並不認為電腦真能「看到」攝影機前充滿形狀與色彩的視覺景象、真能從麥克風「聽見」問題，真能像人類一樣有感覺、擁有意識經驗。

要如何知道機器已獲得「意識」這種似乎難以言喻的特質呢？我們的策略是利用下述知識：有意識的機器才能主觀了解某張普通照片裡的景象是「對」是「錯」。這種用一堆事實組合出合理的現實樣貌的能力（例如知道大象不應高掛在巴黎鐵塔頂端），是意識的決定性特質。相較之下，佔滿整個房間的IBM超級電腦，仍然無法看出影像合不合理。

釐清有意識的機器具有哪些特性，不只能讓人類了解自己的大腦如何運作，也能讓我們預做準備，當科幻小說所預言的日子來臨時，學習如何與我們所創造的另一種意識體共同生活。了解這些，或許還能

幫我們解答一個困擾世世代代哲學家的根本問題：什麼是意識？

是人類還是魔像？

哲學家有個思索已久的問題：不管是傳說裡的魔像或裝在盒子裡的機器，這些人造的擬人產物真的能擁有感覺嗎？後來，在第二次世界大戰協助破解納粹潛艦軍隊「謎碼」（Enigma code）的英國數學家涂林（Alan Turing），於1950年發表一篇論文，開啟了人工智慧的研究。涂林在《心智》（*Mind*）期刊發表的文章中提議，用一個比較可行的問題「是否能做出一部機器，以電報打字機和它進行問答時，其表現與真人無異？」取代模糊而無解的問題「機器能不能思考？」，能滿足這個問題的機器，就等於通過涂林測試。

現代版的涂林測試方法是，找個人來判斷：在電腦螢幕的另一端，以日常的「自然語言」與其互動的對象是人或軟體程

右頁的圖有何古怪？要機器判斷這張圖有何不對勁，除非內建特別為這類照片寫的程式，要不然就是機器必須能「意識」到這世界的許多事情。

GEORGE KEERN



關於作者

柯霍是美國加州理工學院認知與行為生物學特羅思多講座教授，也是艾倫腦科學研究中心科學總監。他上一次為《科學人》撰寫的文章是2007年11月號的〈名家論戰：意識如何產生？〉。

托諾尼是威斯康辛大學麥迪遜分校意識科學特聘教授兼睡眠醫學懷特講座教授。



重點提要

- 電腦的智慧不斷進步，但這些機器要能「意識」周遭世界，路途還遠得很。
- 計算機科學家和神經生物學家喜歡思索一個兼具技術性與形上學意味的問題：有辦法分辨機器是否真正具有意識嗎？
- 在家拿把剪刀和幾張照片做個簡單測試，你就能確認人類是否已做出有意識的電腦。

式？判斷者和互動對象可以聊任何事，若經過一段合理的時間後，判斷者仍無法確定對方是不是真人，對方起碼就稱得上具有和人一樣的智慧，通過了涂林測試。這麼多年來，人們設計的聊天程式偶爾也能騙倒人，但瞞不了多久。

本文兩位作者不是以計算機科學家的角度，而是基於神經生物學家對大腦如何產生主觀感受的興趣，來探討機器意識的問題。我們檢測志願者或神經疾病患者的大腦，以核磁共振掃描或腦電圖儀記錄腦波；我們也用類似方式研究齧齒類及其他動物的腦。我們和許多同事就這樣鑽研意識的神經關聯性，這是腦內機制的最小單位，任何意識知覺（例如橙色夕陽餘輝的視覺）都是它們聯手造就的。然而直到不久前，這個領域還一直缺少一個能讓我們遵循、用以評估腦傷病人、胎兒、老鼠或矽晶片能否感受意識知覺的通用理論。

我們提出一套「意識的整合資訊理論」，來應付這項挑戰，其中涉及一項意識判斷的關鍵。許多人憑直覺知道，每個人對氣味、景物、思維或回憶等日常感受，都有自己獨特的感覺，而這些主觀現象狀態的形成，必定與大腦如何將它接收到的感覺訊號以及來自記憶的資訊，整合成世界的完整面貌，有某種關聯。只是要怎麼把這種直覺講得更精確？

整合資訊理論提出兩個公設來解決這個問題。第一，意識具有高度資訊價值。這是因為當某個意識狀態發生時，就排除了無數的其他狀態，而它與其他狀態的不同處也是它特有的。想想你在所有電影裡看過的所有畫面，每張畫面的視覺都是獨特的意識知覺，當你知覺到某個畫面時，大腦就排除了億萬個其他可能的畫面。即使是「在一個黑暗房間裡醒來」這種最簡單的視覺經驗，漆黑一片的知覺也代表你所見的不是明亮的客廳、茂密的叢林，或心裡可能出現的無數其他景色。

其次，意識資訊是整合的。當你意識到朋友的臉，就不可能沒看到她在哭或戴著眼鏡。無論你多努力去試，都不能把左視野和右視野分割開來，或把視覺切換成黑白模式。無論什麼景色進入意識，都會保持完整，不能切割為獨立而不相關的成份，讓人分別感受。意識的統合來自腦中相關腦區的大量互動，如果腦區彼此不相連結，就像麻醉或熟睡時的狀況，意識就會衰減，甚至消失。



糾錯：要測試機器是否有意識，只要問它「這兩張圖哪一張錯了？」當今任何自動化科技都會被難倒。

也就是說，你必須是一個擁有大量可區辨狀態（亦即資訊）的整合單一個體，才會有意識。系統整合資訊的能力（亦即意識）可以這樣評估：該系統所含的資訊，比組成它的各部份所擁有資訊的總和，多出了多少？這個數值就稱為 Φ (phi)。原則上，無論是人腦、機器人或手動調整的控溫器，任何系統都能算出 Φ 值。 Φ 代表系統無法再化約為更小組成元件的部份，衡量的單位是位元。 Φ 和意識程度高的系統，必定由各有所司且整合良好的成份所組成，這些組成份整合後的能力超越個別能力的總和。

如果系統的組成份大多互相獨立，就像數位相機的感光元件或電腦記憶體裡的位元， Φ 值就低。如果組成份沒有分工、角色重複，全都做同樣的事， Φ 值也低。如果系統組成份是以隨機方式互相連結， Φ 值還是低。但在腦內特定區域（例如大腦皮質）內，神經元都擁有豐富而獨特的連結方式， Φ 值就高。這種系統整合性的評估也適用於金屬盒子裡的積體電路，只要電晶體和記憶單元間的連結夠複雜，電腦就能像人腦一樣，達到高度的資訊整合。

除了用機器內的線路來評估 Φ （這可是苦差事），還能如何知道機器有沒有感覺？可行的測試是什麼？有個辦法能測試資訊整合度，就是問它六歲小孩也能輕易回答的問題：「這張圖有什麼不對？」要解答這個簡單問題，需要大量背景知識，遠超過先進電腦用來辨識人臉或偵測信用卡盜刷的演算法。

物體或自然景物的圖片中，畫素與物體彼此有著錯綜複雜的大量關聯性，俗話說得好：「一張圖勝過千言萬語。」視覺系統的演化、兒時的神經發育及生活的經驗，讓我們能即時認出圖中各組成加在一起是否適當：那些紋路、深度、色彩、各部份的空間關係等，是否合理？

電腦分析影像、檢查圖片中資訊是否協調，所需處理的內容，遠比用語言查詢資料庫更多。電腦或許能在複雜的比賽中打敗人類，但拿一張照片隨便問它裡面的內容，電腦還是沒有能力回答，箇中原因可以用資訊整合程度來解釋。雖然現代電腦硬碟的容量超過人類一生的記憶量，但裡面的資訊是未經整合的，系統裡的每個元素大多彼此不相連結。

就以你電腦相簿裡的一張書桌照片為例，電腦無法看出

在亂糟糟的桌面中，左邊有 iMac 而右邊有 iPad 是否合理。更糟的是，電腦不但不知道 iMac 和 iPad 放在一起很搭，更不知道在放鍵盤的地方放個盆栽很怪；它不知道 iPad 不可能飄浮在桌面上方，或這張照片的左半邊與右半邊很合，但若搭上其他照片的右半邊就不對勁。對電腦而言，所有畫素只是一大堆三個數字的組合（分別代表三原色），彼此間互無連結，不具什麼意義。但對你而言，影像是有意義的，因為裡面各層次（畫素、物體，乃至於景色）的各部份彼此充滿連結，而這些關聯性決定了影像哪些部份合起來搭調或不搭調。根據我們的理論，這張由關聯性知識組成的整合網，讓每張影像不同於其他無數影像、具有獨特性，也讓你對這世界充滿意識。

就連六歲小孩也能藉由這種整合性，看出溜冰選手站在客廳地毯上、透明的牛、貓追狗，這些不合理畫面的荒謬之處，而其中潛藏的秘密就是判斷電腦有無意識的關鍵。這些明顯違反我們預期的事物，證明我們擁有非凡的知識，知道特定事件與物體會一起出現，而其他難以計數的組合則不會。

影像版的涂林測試

測試電腦看不看得懂圖片，不是像傳統涂林測試那樣，輸入一個問題給機器；而是從網路上隨機選幾張影像，垂直分三等份，把中間部份塗黑，再將剩下的左右兩部份隨機搭配，要求電腦選出正確的組合。除了原本就來自同一張影像的左側和右側，其他組合都不會正確。由於中間塗黑，目前電腦用的影像分析策略（例如尋找能組合在一起的影像紋路或顏色）太過簡單，無法正確回答。要通過這種分割影像測試，需要很強的視覺理解以及篩選影像可能組合的能力。

另一個測試是在幾張影像中插入物體，使其中只有一張看起來不合理，再讓電腦挑出奇怪的那張。工作檯和鐵鎚很搭，但鐵鎚一定不會懸浮在半空中；iMac 前面放個鍵盤是正確的，放盆栽就不對。

電腦有很多運用影像特徵（例如色彩、邊界或紋路）的低階統計資料來尋找適當配對的策略，它們或許能應付其中一種測試，卻無法通過多種不同的影像測試。要做出讓這些測試發揮實用價值的詳細設計，還要更多努力。然而這些測試能凸顯一件事：人類意識知覺到的大量整合知識，與目前機器視覺系統所擁有極度狹隘又專門的知識，形成強烈對比。是的，今日電腦能從有 100 萬張臉孔的資料庫中，挑出可疑的恐怖份子，但無法辨認他的年紀、性

別或種族，不知道他是否看著鏡頭、是笑還是皺眉，也看不出這個人和華盛頓握手的照片很可能是數位合成的。任何有意識的人類只要一眼就能看出這些，甚至更多的事。

知道這些之後，我們對不久的未來可以有怎樣的期待？只要能找出某個任務的某些特徵與其他任務無關，就能交給機器去處理。高速演算法能迅速搜尋龐大的資料庫，打敗人類西洋棋王和「危險邊緣」冠軍。人類可以挑選大量相關實例輸入電腦，訓練複雜的機器學習演算法學會辨識人臉或偵測行人，電腦就能表現得比人類更快更好。不難想像，將有越來越多的專門任務由機器接手。高階電腦視覺系統即將成熟，不出 10 年，就會有可靠而幾乎全自動化的駕駛模式可供選用。

不過，我們預料這類機器視覺系統並無法回答有關車子前方景物的簡單問題：開車從高速公路駛近芝加哥，遠方地平線上的景物像不像從一團煙霧中浮起的燒焦樹叢？它也看不出加油站旁邊出現一根巨無霸香蕉很詭異（或許在洛杉磯就不會很怪）。要回答上百萬個這類問題，或指出那根香蕉有何不對，需要無數個專為問題所設計的軟體模組，但沒人能有先見之明去設計那些模組。沒錯的話，儘管高階機器視覺系統能根據一組專業分工的平行模組，實現近乎全自動化的駕駛，而其他許多日常工作也能比照辦理，這些系統仍無法意識到它們前方的景物。

不過或許也會有另一種不同的機器，把世間事物彼此數不清的關聯性化為知識，納入高度整合的單一系統。這種機器面對「這張圖有何不對」的問題，依舊會跳出答案，因為不管哪裡出錯，都無法符合它內建系統裡，隨資訊整合方式所決定的某些內在限制。

這種機器會擅長處理不易分割為許多獨立任務的事情，它有整合資訊的能力，會對景物產生意識知覺。我們推測，為了要達到高層次的整合，機器也許會善加仿效哺乳類大腦的結構原則。這些機器將輕易通過我們提出的測試，果真如此，就代表它們和我們一樣，擁有宇宙中最神秘的特質——意識的天賦。 **SA**

王怡文是台灣大學資訊工程學研究所碩士，現就讀於台灣大學心理學研究所。

延伸閱讀

Can Machines Be Conscious? Christof Koch and Giulio Tononi in *IEEE Spectrum*, Vol. 45, No. 6, pages 54-59; June 2008.

Consciousness as Integrated Information: A Provisional Manifesto. Giulio Tononi in *Biological Bulletin*, Vol. 215, No. 3, pages 216-242; December 2008.

[Scientific American Online](http://ScientificAmerican.com)

歡迎參加欺騙電腦大賽，請上：ScientificAmerican.com/jun2011/koch-contest