

歸檔記憶的概念神經元

每個概念，包括每個人或每件事以及每天的經驗，都可能有一組相對應的神經元。

撰文／基洛加（Rodrigo Quian Quiroga）、弗里德（Itzhak Fried）、柯霍（Christof Koch）

翻譯／謝伯讓

重點提要

■過去數十年，神經科學家一直在爭論記憶究竟如何儲存。目前許多相互競爭的理論中，有一個理論認為單一神經元可以儲存關於祖母或電影明星等的記憶。

■另一個理論認為，每個記憶都是由數百萬個細胞分散儲存。最近有些與腦部手術同時進行的實驗發現，某些特定腦區中的相對小群神經元跟記憶儲存有關。

■此外，這一些小群的細胞可以記憶同一件事物的各種不同面向，例如祖母的臉或身形，或美國明星珍妮佛安妮斯頓的正臉、側臉或聲音。

阿卡希維奇（Akakhi Akakhievitch）是一位傑出的俄國神經外科醫師，他有一位病人想要忘掉令人難以忍受的母親。阿卡希維奇熱心地打開了病人頭殼，並一顆顆清除掉上千個和他母親記憶有關的神經元。當病人從全身麻醉中甦醒時，失去了所有關於他母親的記憶。成功之餘，醫生欣喜地轉向下一個目標：尋找與「祖母」記憶有關的細胞。

當然，這個故事是虛構的。已故的神經科學家雷特溫（Jerry Lettvin，真實人物）於 1969 年在美國麻省理工學院對一群學生講述了這個故事，用來說明一個充滿爭議的可能性：只需大約 1 萬 8000 個神經元就可以產生關於事物、親友或其他人的各種意識經驗、想法或記憶。雷特溫沒能夠證實他的假說，而過去 40 年來，科學家也大多戲謔地爭論著「祖母細胞」這個點子。

關於神經元以極度特定的方式儲存記憶的想法，可以追溯到詹姆斯（William James），他在 19 世紀提出了「教皇細胞」（pontifical cell），認為這些細胞是意識所在之處。不過這個概念和當時的主流想法背道而馳，當時大家認為有關任何人或物的知覺，都是由數十億或至少數百萬個神經元共同負責，也就是諾貝爾獎得主薛林頓（Charles Sherrington）在 1940 年提出的「百萬民主」概念：單一神經元的活動是無意義的，只有一大群神經元的集體活動才有意義。

神經科學家一直在爭論是否只需要少數的神經元（數千或更少），就足以呈現單一特定概念，或是需要廣佈腦中的上億個神經元才夠。研究這個問題的過程和結果，讓我們對記憶與意識的運作有了更新的理解，好萊塢對此也有些小幫助。

發現珍妮佛安妮斯頓神經元

幾年前，我們與克雷曼(Gabriel Kreiman，現任教於美國哈佛醫學院)和雷迪(Leila Reddy，現為法國腦與認知研究中心研究員)進行了一些實驗，發現有位病人的海馬回(與記憶有關的腦區)中有一個神經元會對因電視影集「六人行」而走紅的美國女演員珍妮佛安妮斯頓(Jennifer Aniston)的照片有強烈反應，但卻對其他數十位演員、名人及各種景點和動物毫無反應。另一位病人的海馬回中則有一個細胞對演員荷莉貝瑞(Halle Berry)的照片甚至英文名字(在電腦螢幕上)有反應。另有一個神經元則只對美國脫口秀主持人歐普拉(Oprah Winfrey)的照片和名字(在電腦上或電腦合成發音)有反應。還有一個神經元對電影「星際大戰」中絕地武士天行者路克的照片，以及他的名字(無論被寫出或唸出)有反應。

這些結果都是透過直接測量單一神經元的活動才發現的。其他如功能性腦造影(fMRI)等較常見的技術，雖然可以找出受試者在進行某項活動時特定腦區的反應，但卻只能追蹤數百萬個細胞共同消耗的能量，無法定位更小群神經元的活動，更不用說單一細胞了。要記錄單一細胞的神經脈衝，必須在腦中植入比頭髮還細的微電極。這種技術不像fMRI一般普遍，而且只有在某些特殊的醫療狀況下才可能允許在人類腦中植入這些電極。

其中一種狀況就是治療癲癇。當癲癇無法用藥物控制時，病人可能得接受手術治療。醫療團隊必須先找到癲癇的起始位置(癲癇焦點)，然後才可能移除該部位來治療。一開始的評估方法是非侵入式的，例如透過腦造影、臨床證據，以及電生理的病理訊號(從頭殼上量到的腦電圖中有密集的癲癇電位變化)。但是當這些方法都不能精確判定癲癇焦點時，在醫院中，神經外科醫師可以在顱內植入電極，持續數天監測分析癲癇的症狀。

科學家有時候會請病人在這段期間參與實驗，並記錄下各種認知活動運作時的腦部變化。在加州大學洛杉磯分校，我們採用了一種獨特的技術，以柔軟的電極來記錄顱內神經元的活動。這項技術是本文作者中的弗里德發展的，參與合作的還包括了加州理工學院的柯霍團隊，以及英國萊斯特大學基洛加的實驗室。這項技術提供了難得且獨特的機會，可以直接記錄清醒病人的單一神經元活動長達數天，藉此研究各種實驗下神經的活動變化，例如觀看筆記型電腦上的圖片、記憶回想等。這就是我們發現珍妮佛安妮斯頓神經元，以及無意間重新引發雷特溫寓言論戰的過程。

重新檢視祖母細胞理論

珍妮佛安妮斯頓神經元是否就是備受爭議的祖母細胞？要回答這個問題，我們必

須先精確定義祖母細胞。有一種極端的說法認為：單一神經元對應單一概念。但是如果我們可以找到對珍妮佛安妮斯頓有反應的神經元，那就強烈暗示還有更多這種細胞，畢竟要在幾十億個神經元中恰巧找到這唯一的一個，機率實在太低了。此外，如果只有一個神經元負責整個關於珍妮佛安妮斯頓的概念，那當這個細胞因疾病或意外而損毀或死亡時，所有關於珍妮佛安妮斯頓的記憶就會消失，這種情況的可能性似乎很低。

另一種較不極端的說法認為，有不只一個神經元對應到一個概念。這個假說可能正確，但卻很難證實（如果可以證實的話）。我們不能無窮無盡去測試所有的概念以驗證該神經元是否真的只對珍妮佛安妮斯頓有反應。事實上，真實情況通常剛好相反：神經元大多不只對一種概念有反應。因此，如果神經元在某個實驗中只對某個人有反應，我們仍無法排除它可能也會對其他某些從未測試過的事物有反應。

例如，在發現珍妮佛安妮斯頓神經元的隔天，我們拿更多與她有關的圖片來重複實驗，結果發現該細胞也會對麗莎庫卓（「六人行」中另一位知名演員）有反應。此外，對天行者路克有反應的神經元，對尤達（另一位絕地武士）也有反應；另外有一個細胞對兩位籃球選手有反應，還有一個細胞則對本文作者之一（基洛加）以及一些在加州大學洛杉磯分校中和病人有互動的同事有反應。不過在比較寬鬆的條件下，你仍可以主張這些細胞是廣義的祖母細胞，換言之，它們可以代表「『六人行』中的金髮女子」、「絕地武士」、「籃球選手」或者「與病人互動的科學家」。這種延伸的定義，讓祖母細胞的討論陷入了語義問題之中。

讓我們先把語義的問題放一邊，專注在這些珍妮佛安妮斯頓神經元的一些重要特徵上。第一，我們發現每個細胞的反應都非常有選擇性：每個細胞只會對少數的名人、政治人物、親屬、地標等事物有反應。第二，每個細胞都會對某個特定人物或地點的各種面向產生反應，這些細胞在遇到同一個人的照片、名字的文字或語音時，都有類似的反應，宛如藉由反應來說：「我知道這是珍妮佛安妮斯頓，無論她穿著紅禮服、只顯示側身、只讓我看到或聽到名字時都一樣。」換言之，這個細胞似乎對該事物的概念和表徵有反應。因此，這些細胞或許更適合被稱為「概念細胞」。概念細胞有時候會不只對一種概念產生反應，但通常這些概念彼此密切相關。

少數神經元就能建立概念

為了理解少數細胞和某些特定概念（例如珍妮佛安妮斯頓）產生連結的機制，我們要先知道大腦如何捕捉並儲存周遭各種事物和人物影像的複雜過程。進入眼睛的資訊（從眼球經由視神經）會先傳送到頭部後方的主要視覺皮質。該處的神經

元會對影像中某個小部位的細節產生反應，這些細節就宛如數位影像中的畫素，或是秀拉（Georges Seurat）畫作中的彩點一般。

一個神經元不足以辨識出它所見的一個小點是來自於一張臉、一個杯子或是艾菲爾鐵塔。每個細胞都是群體的一部份，群體整合起來才能產生一張複合的圖像，就像是秀拉的畫作「傑特島的星期日下午」。當影像稍微改變時，某些細節也會出現變化，相對應的神經元反應也會因此有所不同。

大腦必須捕捉許多影像來處理知覺訊息，它要能夠認出物體並和已知的知識做整合。影像所引發的神經活動，會從主要視覺皮質開始，經過一連串的皮質區域後到達較接近前額的腦區，這些高階視覺區域中的神經元會對整張臉或整個物體、而非局部的細節有反應。光是一個這種高階神經元就可以指出某個影像是臉而不是艾菲爾鐵塔。當我們稍微改變圖像，例如移動它或改變光照的方式時，圖像的某些特徵就會改變，但這些細胞並不太在乎這些細節變化，還是會做出大致相同的反應，這種特質稱為「視覺不變性」（visual invariance）。

高階視覺腦區的神經元會把訊息傳送到內顱葉（海馬回與周遭腦區），該處負責記憶功能，也是我們發現珍妮佛安妮斯頓神經元的地方。海馬回中的神經元反應比高階視覺腦區中的神經元更具有專一性，這裡的每個細胞只對某個特定人物（更準確的說法是該人物的概念）有反應：不只對其臉孔和外表的各個面，也對密切相關的各種特質（例如名字）有反應。

在我們的研究中，我們試圖探究需要多少神經元的活動才能呈現某個特定概念。我們想知道究竟是只需要一個、數十、數千、還是要數百萬個細胞才夠。換言之，呈現概念時所需神經元的「稀疏」（sparse）程度如何？很明顯，這個數字無法直接測量，因為我們無法同時記錄一個區域中所有的細胞活動。當時本文作者之一柯霍在加州理工學院的博士班研究生偉多（Stephen Waydo），以統計方法估算出單一概念所激發的神經元不會超過 100 萬個（整個內側顱葉約有 10 億個神經元）。不過，我們採用的圖片大多是病人非常熟悉的事物，會激發較多的反應，因此這個數字應該被視為上限。普通狀況下，呈現一個概念所需的神經元是這個數字的 1/10 到 1/100，可能非常接近雷特溫的預測：每個概念由 1 萬 8000 個神經元所呈現。

相對於上述的主張，有一個理由讓有些人認為大腦並非用稀疏的神經元來呈現概念，而是利用非常大量神經元以分散（distributed）的方式來呈現概念，也就是說我們可能沒有足夠的神經元可以呈現所有可能的概念和產生的變化。我們有足夠的腦細胞來記錄微笑的祖母、編毛衣的祖母、喝茶與等巴士的祖母，以及向群眾致意的英國女皇、沙漠星球上的年幼天行者路克和對抗黑武士的天行者路克等

諸多概念嗎？

要回答這個問題前，我們必須先考慮到一個事實：一般人通常不會有一萬種以上的概念。而這個數字和內顳葉中 10 億個神經元比較起來，並不多。此外，我們有很好的理由相信，概念是以非常有效率的稀疏方式儲存並呈現的。內顳葉的神經元不在乎概念的細微變化，它們不管路克是坐或站，而只注意該事物是否和路克有關。它們會對某概念產生反應，而不在乎其呈現的方式。如果概念比較抽象（對路克的所有形象都產生反應），就可以降低神經元所需要存錄的訊息，也能夠讓它變得比較有選擇性（只對路克而不對珍妮佛安妮斯頓產生反應）。

偉多的模擬研究更強調了這項可能性，他根據視覺訊息處理的細部模型建立了一套可以學習辨識許多無標示圖片（飛機、汽車、摩托車和人臉）的軟體模擬神經網絡。這套軟體可以在沒有指導者的情況下學會辨識。在沒有人告訴它「這是飛機、那是汽車」的情況下，它必須假設現實中各種圖片都只是來自少量的人或物，而且每個人或物都只由一小撮神經元所負責（就跟我們在內顳葉中發現的情況一樣），並藉此自己學會辨識。當這樣的稀疏式呈現方式植入軟體模擬後，該網絡學會了辨識，即使圖片中的人和物出現變化也沒有問題，就像我們在人腦細胞中發現的反應。

為什麼有概念細胞？

我們的研究與「大腦如何詮釋外在世界並把知覺轉變成記憶」這個問題密切相關。想想 1953 年的著名病例 HM，他飽受無法治療的癲癇之苦。為了停止癲癇，神經外科醫師移除他的兩側海馬回及周邊腦區。手術後他仍可認得手術前認識的人事物，卻無法形成新的長期記憶。沒了海馬回，他所經歷的一切很快就消逝無蹤。2000 年的電影「記憶拼圖」就是在描述類似的神經學症狀。

HM 的案例顯示海馬回以及內顳葉大致上並不是知覺所需的腦區，卻是把短期記憶（可以暫時記住事物）轉變成長期記憶（可以記住長達數小時、數天或數年）的關鍵。同樣的，我們認為此腦區中的概念細胞是把意識內容（由感官刺激或內在回想所激起的任何內容）轉變成長期記憶的關鍵（這些記憶產生後會儲存在大腦其他皮質）。我們相信珍妮佛安妮斯頓神經元並非辨識或回憶起她所必需的細胞，而是把她置入意識以便形成新的連結和記憶（例如記得自己見過她的照片）的關鍵。

大腦可以利用少數概念細胞將一件事物的多種變化，歸類成一個獨特的概念（稀疏且不變的表徵方式）。概念細胞的運作方法甚至可以用來解釋我們的回憶機制：我們可以回憶起珍妮佛安妮斯頓和路克的樣貌，但不會記得他們臉上所有的

毛孔。我們既不需要也不想要記得周遭所有的細節。

重要的是抓住各種情境中與我們有關之人或概念的主旨，而不是去記住多如牛毛且無意義的細節。當你在咖啡店遇到熟人，重要的是要記住會面中一些顯著的事件，而不是對方的穿著與說出的每字每句，或是店裡其他陌生顧客的長相。概念細胞通常會對與自己相關的事物有反應，因為我們一般會記得的事件都涉及了自己熟悉的人或事物，對於不甚相關的東西，我們是不會花心力去記住的。

記憶絕對不只是單一獨立的概念。與珍妮佛安妮斯頓有關的記憶是一連串她（或她在「六人行」中所飾演角色）所參與的事件。一個完整的記憶片段涉及了許多不同但相關的概念，像是把珍妮佛安妮斯頓連結到你坐在沙發上大吃冰淇淋以及看「六人行」等概念。

當兩個概念連結時，有些原本只存錄其中一個概念的神經元也會開始對另一個概念產生反應。這個假說替大腦存錄關聯性的方式提供了一個生理方面的解釋。神經元會對相關概念產生反應的傾向或許正是產生情節記憶（例如咖啡店中發生的一連串事件）或意識流動（從一個概念自然的移動到另一個概念）的基礎。我們看到珍妮佛安妮斯頓，該知覺就會引發關於電視、沙發和冰淇淋的記憶（和其他諸多觀看「六人行」的記憶有關的概念）。同樣的過程也會讓同一個概念但存在於不同大腦皮質中的諸多面向產生連結，例如玫瑰的氣味、外形、顏色以及質感，或是珍妮佛安妮斯頓的外表和聲音。

既然把高階記憶儲存成抽象概念有許多明顯的優點，我們就可以接著問，為什麼這些概念的表徵必須散佈在內顳葉中少數的神經元中。其中一種可能的答案來自模擬實驗，結果顯示，符合稀疏式的呈現方式乃是建立快速連結所必需的。

這其中的技術細節非常複雜，但基本的想法很簡單。假想你有一個關於咖啡店中那位熟人的分散式（相對於稀疏式）記憶，這種方式運用到許多神經元來存錄這位熟人的所有特徵。再假想你有另一個關於咖啡店的分散式記憶。當你要連結兩者時，必須先把兩個概念中所有不同細節的連結建立起來，但又不能把兩者與其他概念相混淆（因為咖啡店看起來像是舒服的書店，那位熟人跟其他人看起來也有點像）。

要在分散式網絡上建立這樣的連結很花時間，而且容易產生記憶混淆。相形之下，要在稀疏式網絡上建立這種連結，則只需要在負責兩個概念的兩群細胞間產生些許連結即可（僅需讓幾個細胞同時對兩個概念產生反應就可達成），快又容易。另一個稀疏式的優點在於添加新概念並不會大幅改變網絡其他部份。這種區隔在分散式網絡中很難達成，因為新概念加入會讓整個網絡的邊界產生變化。

概念細胞把知覺和記憶連結在一起，它們透過抽象且稀疏的方式來呈現語意知識（生活周遭的人、物、地點等各種有意義的概念）。它們是我們生活中關於各種事件與事實記憶的基石，其巧妙的記錄架構讓心智可以忽視許多不重要的細節，並且萃取出其中意義，進一步產生新的關聯和記憶，也存錄了我們經驗中值得保存的關鍵要素。

概念細胞並不像是雷特溫心中所想的祖母細胞，但它們可能是人類認知能力的重要神經基礎，以及思想與記憶的要件。