



讀心不再是夢想

電腦科學遇上腦科學，交迸出智慧的火花！

科幻小說裡經常出現可以窺視人類心靈的儀器，比如艾西莫夫《基地》系列小說提到的心靈探測器，就能夠感知人類的想法。而在21世紀的今天，藉著人工智慧的研究成果，這類「讀心器」或許真能實現。

美國卡內基美倫大學密契爾（Tom Mitchell）率領的團隊，花了將近10年研究如何透過機器學習（machine learning），讓電腦經由判讀功能性磁共振造影（fMRI）的影像來了解人腦的思維。腦科學的研究早已證實，大腦在思考不同類型的概念時，會觸發不同區域的神經元，活化程度不一的區域在fMRI上會以不同的顏色顯示出來。密契爾的想法是：「能否將一些fMRI影像以及相配對的概念輸入電腦，讓電腦自動學習怎樣的觀念會產生怎樣的影像，進而學會透過這些影像來判讀人們心中所想的東西？」

於是他們從一個簡單的實驗開始。首先，找出25個不同的動詞，並且分成三類（第一類為感知，如吃或聽；第二類為動作，如跑或推；第三類較抽象，如開或填），他們將每個概念對應的英文詞彙呈現於受試者面前，然後拍下受試者腦部的fMRI影像。每個詞彙所對應的影像有數張，分別是从不同角度的三維影像。把這些影像及其相對應的概念輸入電腦後，即可利用機器學習的方法自動產生許多「分類器」（classifier），這些分類器有能力學習辨識不同類別的特性，於是電腦就可以學會自動分辨不同概念。

當然，電腦能辨識25個概念還稱不上是多大的成果。密契爾團隊再進一步利用這25個概念為基礎，讓電腦可以學習並分辨更多新詞彙。他們將這25個基本詞彙逐一與每個新的受測詞彙組合成新的詞組，接著利用Google之類的搜尋引擎來查詢新詞組，並且把符合的結果數記錄下來。不

難想像，詞組中的兩個詞彙如果同時出現的頻率越高，表示它們之間的相關程度越高。接著就可以利用此關聯性當做權重，將這25個基本詞彙的fMRI組合起來，做為代表此新詞彙的fMRI影像，繼而運用這個fMRI影像，訓練出能分辨新詞彙的分類器。

密契爾的團隊還利用實驗來驗證許多假說。例如：被活化的腦部區域所對應的是「意義」或是「外觀」？於是他們讓受試者觀看某個詞彙代表意義的圖片，記錄其fMRI

影像，接著利用之前辨識英文詞彙影像所學到的分類器，來分辨這些經由圖片得到的影像，結果準確度不變。也就是說，會活化人類神經元的是詞彙的意義，所以不管用文字或是圖片來刺激受試者，腦部有反應的區域是類似的。

另一個實驗則著重在驗證是否不同個人的大腦受到概念刺激的模式也不同。他們利用了透過某位受試者的fMRI影像所訓練出來的分類器，來判讀其他人的

腦中在想哪類概念，結果發現，準確度依然很高。這也代表了腦部被活化的區域大致上不會因人而異。

藉由機器學習技術，腦科學家可以大量分類肉眼難以快速分辨的fMRI圖片，進而預測人腦在思考不同概念時腦部活動的可能狀態。密契爾團隊目前的研究成果著重在具體且單一的物體，例如飛機、芹菜。但是諸如民主、焦慮等抽象名詞，或是複合片語（可愛的兔子、飢餓的兔子）上，效果還有待加強。

儘管電腦打敗人腦已不再是無法想像的事，人工智慧學家對這些成就仍不滿足，他們希望電腦不僅能打敗人腦，甚至能做到像是讀心術般連人腦也做不到的事。也許不久的將來，人類的思想在電腦之前也終將無所遁形。 SA

林守德是台灣大學資訊工程系助理教授

