



人工智慧何必模仿人類智慧

用來判斷機器有沒有智慧的「涂林測試」，其實存有盲點。

人工智慧之父、英國數學家涂林 (Alan M. Turing) 1950年提出了一個測試電腦能否思考的「涂林測試」：讓人類裁判與受測機器對話，若機器能在問答過程中讓裁判無法分辨其為人類或機器，就算通過涂林測試，表示機器具有某種程度的智慧。

接下來的半世紀裡，出現了不少「宣稱」通過涂林測試的人工智慧程式。1966年，美國的維森班 (Joseph Weizenbaum) 開發的聊天程式 ELIZA 首先宣稱通過涂林測試。此程式先從人類的問題中找出預載的關鍵字，若匹配成功就回覆相對設定好的句子；若無，則從預設的通用答案中擇一回覆。之後還有很多類似的「聊天機器人」(chatbot) 程式，例如 ALICE、Elbot，都宣稱可騙過人類裁判。

但是後人重新檢驗這些程式，卻認為爭議重重。例如有些實驗在測試過程中，並沒有先告知裁判對話的另一方可能是機器。實驗也證實，有沒有這種「先入為主」的概念，會影響裁判的對話選擇以及判斷。今年媒體爭相報導的聊天機器人 Goostman，也號稱通過涂林測試，爭議點就在於一開始就說自己是正在學英文的13歲烏克蘭小孩，使得裁判把它一些不合邏輯的回答解釋成對英文的認知不足。

上述結果讓學者開始質疑：涂林測試是否真正適合用來測試電腦能不能

思考？最有名的當屬於席艾勒 (John Searle) 的「中國室」(Chinese room) 思想實驗。在這個實驗中，中國室裡充滿了寫好的規則，讓不懂中文的人把一個中文句子轉換成另一個相關的中文句子。根據這些規則，例如把「你好嗎？」轉換成「我很好」，就可讓身在中國室裡不懂中文的人假裝成懂中文的人。同理，一個記載很多規則的電腦程式就好像中國室一樣，可讓人誤認為它懂得對話的含意。

要測試電腦有沒有智慧，為何不直接測試它解決問題的能力？

公認的涂林測試盲點，在於它利用「模仿人類」的成功與否來判斷機器有沒有智慧。然而人類會犯錯，也會做出不合常理的行為，使得機器可藉由模仿人類的愚昧來通過涂林測試。1992年某個涂林測試競賽的冠軍，就是藉由故意打錯字來迷惑評審。此外，有些所謂「智慧行為」是人類做不到的，例如快速的四則運算，所以電腦必須要「限制」自己的能力以通過涂林測試。於是有人提出，要測試電腦有沒有智慧，為何不直接測試它的智商或解決問題的能力？就如同要測試飛機的性能，我們可以直接測試它能不能飛、能飛多久，而不是測試它像不像一隻鳥。

通過涂林測試的程式是不是真的沒有應用價值？其實不然。有一些自動導覽程式便試著藉由跟使用者簡單互動以增進滿意度，人工智慧程式「虛擬愛情」(Cyberlover) 甚至不幸被當成惡意軟體使用：它欺騙網路上希望尋求情感慰藉的使用者，進而竊取個人資料或是引導對方造訪一些有害的網站。

近年來，更多人工智慧學者往「弱人工智慧」(weak AI) 的方向靠攏。也就是不強求電腦能夠有「像人類一樣」思考的智慧，而是希望電腦能夠用它所擅長的快速運算與記憶能力來達成「看似有智慧」的任務。例如 IBM 的西洋棋程式「深藍」達成「電腦在棋賽中勝過人腦」的測試；IBM 的「華生」也在益智問答競賽中勝過人類參賽者。在21世紀不可或缺的搜尋引擎，更可以看成電腦在「搜尋測試」中遠遠超越人類智慧的體現。還有另一些測試，例如自動車駕駛、自動翻譯、機器人與真人PK足球、預測金融股市的程式等，目前雖然沒有完全實現，但普遍認為在不久的將來都有很大的成功機會。

如果涂林復生，看到這些無法通過涂林測試的智慧程式或服務，漸漸在一般人日常生活中扮演不可或缺的角色，心中必定是五味雜陳吧。 SA

林守德是台灣大學資訊工程系副教授。