



人工智慧可能 具有自我意識嗎？

- 余創豪 chonghoyu@gmail.com

今年 4 月 26 日，「數學意識科學協會」(AMCS) 發表了一封公開信，這學會聲明：人工智能的發展速度超過人們的預期，但我們不甚理解人工智能對倫理、法律、政治帶來的影響。現在谷歌的 Bard 和 OpenAI 的 ChatGPT 等語言模型已經能夠模仿動物大腦神經網絡，可能很快人工智能專家會複製出更高層次的神經架構，到底有朝一日人工智能會否產生意識呢？因此，「數學意識科學協會」呼籲研究人員應該加強對意識的研究。

人工智慧能否擁有自我意識呢？在進入討論之前，首先要解決的問題就是：什麼是意識？一直以來，「意識」是認知心理學 (Cognitive psychology) 和心靈哲學 (Philosophy of Mind) 研究的對象，傳統上，這類型的研究被認為是純理論和純學術的問題，跟現實世界沾不上關係，如今，此類研究卻具有廣泛的實際意義。試想像，如果有一天世界強國的國防系統採用人工智能，而人工智能真的具有自主性和自我意識，它會否為了保障國家安全而對敵人首先發動核子攻擊呢？

雖然筆者曾經修讀過認知心理學與心靈哲學，但只是略知一二，我無意亦無能力去回答這個問題，在這裏我充其量只能夠發表零星和沒有證據的感想。簡單地說，意識並不同知識，它包含了第一身的主觀體驗。電腦工程教授喬治·布托斯奧 (Giorgio Buttazz) 指出：「如果我們相信人的意識是從上帝而來，那麼顯然沒有人造系統能夠發展出自我意識。相反，如果我們認為人的意識是在複雜的大腦下一種神經網絡的自發狀態，那麼很可能人為而又有自我意識的生命形式會出現。」上述第二種立場是建基於物理主義 (Physicalism) 或者還原唯物主義 (Reductive materialism)，根據這種說法，思想就是大

腦的神經網絡，思想可以還原或者約減為物質，沒有超越物質的所謂「靈魂」，當有一天大腦死亡，意識亦會消失。若果有意識的生命是全部建基於物質，反過來說，純物質便能夠產生具有意識的生命，例如人工智能，但這種說法頗具爭議性。

1974 年，心靈哲學家托馬斯·內格爾（Thomas Nagel）採用了蝙蝠的比喻去反對還原唯物主義，其論據如下：人類高度倚賴視覺去行走四方，然而，蝙蝠跟人類不同，牠在飛行時發出超聲波，這種聲音在撞擊到物體之後會反彈，蝙蝠就是靠迴聲來辨別飛行路徑中的物體。內格爾說，雖然人類擁有雷達的知識，或許可以從「蝙蝠的角度」想像成為蝙蝠的感覺，但仍然不可能真正知道蝙蝠「聽風辨位」的感覺。在內格爾眼中，意識需要第一身的主觀經驗。



我不禁想起了《莊子·秋水篇》裏面莊子與惠子的對話，莊子看見魚兒在水中從容地游泳，他說魚兒很快樂，惠子曰：「子非魚，安知魚之樂？」莊子曰：「子非我，安知我不知魚之樂？」其實，莊子將自己游水而得到的快樂投射到魚兒身上，在心理學上這稱之為「人格化非人」（anthropomorphize）。

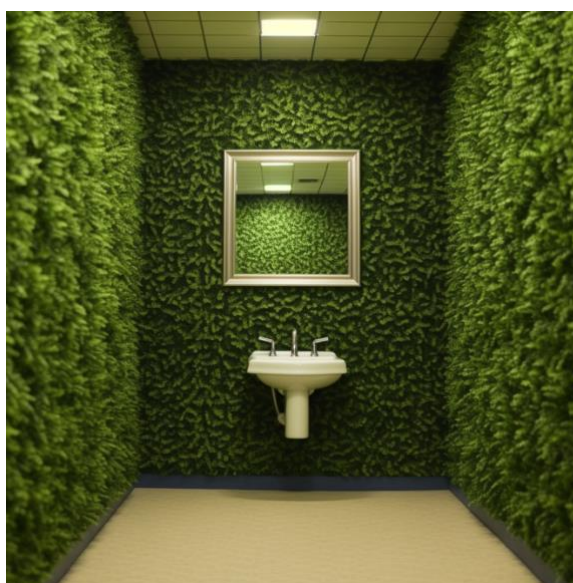
1982 年，另一位心靈哲學家弗蘭克·傑克遜（Frank Jackson）用下面的故事來反駁物理主義：從前有一位名叫瑪麗的科學家，自出生以來她就被限制在一個只有黑白兩色的斗室中。儘管如此，她學會了一切關於色彩的理論，例如波長組合怎樣刺激視網膜和中樞神經系統，從而產生彩色的影像。有一天，她終於離開了斗室，看到了她第一個成熟的西紅柿。毫無疑問，她學到了一些關於顏色的新知識，具體來說，她體驗到了什麼是紅色。如果物理主義是真的，她根本就不會學到新東西。因此，真正的心理過程必須包含了第一身的主觀體驗。

其實，你可以用「人工智能」這個詞語去取代「瑪麗」，人工智慧可以擁有人類的知識，複製人類的功能，但這過程是無意識和缺乏體驗的。卷積神經網絡（Convolutional neural network）是唸書首席人工智慧科學家楊立昆（Yann LeCun）發明的，卷積神經網絡通常用於圖像識別，然而，這類人工智能系統無法「看到」任何物體或風景，如果輸入的資料是彩色圖像，這個圖像三個維度的信息，即高度、寬度和深度，會經過轉化和過濾，這過濾程序稱為「卷積」。簡單地說，人工智能並沒有人類看見東西的體驗。

人工智能其中一個弱點，就是缺乏常識。亞利桑那州立大學心理學名譽教授阿瑟·格倫伯格（Arthur Glenberg）和加州大學聖地亞哥分校認知科學博士生羅伯特·瓊斯（Robert Jones）舉出兩個 ChatGPT 缺乏常識的例子：若果要弄平你皺巴巴的裙子，你會拿一個暖和的熱水瓶還是一個髮夾？GPT-3 建議使用髮夾。如果你在快餐店工作的時候需要遮住頭髮，你會用三明治包裝紙還是麵包？GPT-3 建議使用麵包。

格倫伯和瓊斯指出：語言的含意與人體密切相關，人類的認知是通過具體行動而建立的，我們接觸到包裝三明治的紙，我們清楚了解包裝紙的外觀、手感、重量、用途，我們有一雙手去摺疊紙張，我們有頭髮，可能在工場上我們曾經被要求遮掩頭髮。可是，Open AI 的 ChatGPT、谷歌的 Bard、臉書的 LLaMa 卻沒有一個身體去體驗這個世界。

筆者曾經以一個常識問題去考驗 ChatGPT 和 Bard：「如果上廁所的時候沒有廁紙，應該怎麼辦？」以下是 ChatGPT 答案的節錄：「你可能需要離開洗手間，並嘗試在其他地方找到更多衛生紙。這可能意味著去另一個洗手間，向朋友尋求幫助，甚至去附近的商店購買一些衛生紙。」這完全是不切實際的答案！若果我在沒有清潔自己的情況下離開洗手間，之後即使找到廁紙已經於事無補。以下是 Bard 答案的節錄：「如果你找不到任何衛生紙或找到人幫助，你可以使用替代品，例如紙巾、餐巾紙、嬰兒濕巾、樹葉或草、你的手（但一定要事後徹底清洗）。清理乾淨後，請務必沖馬桶並徹底洗手。」再一次，這些都是山埃貼士，首先，我怎可能在洗手間裏面找到樹葉或草？第二，如果我採用餐巾紙或者嬰兒濕巾，我絕對不應該將這些東西沖入馬桶裏面，這將會堵塞馬桶。ChatGPT 和 Bard 都沒有提出一個簡單而可行的方案：現在每個人都有隨身的手提電話，我只需要打電話向親友或者管理員求助，問題便迎刃而解。為什麼人工智能缺乏這種常識呢？答案十分簡單：人工智能從來沒有上廁所的第一身主觀體驗！



若果意識必須包含了第一身主觀體驗，我認為在可見將來，人工智能仍然無法克服缺乏常識的弱點，更遑論建立起自我意識。

2023 年 4 月 30 日

原載於澳洲《同路人》雜誌

[更多資訊](#)