

## 專欄



發布日期：113-05-14

## 我們真有可能因AI更了解人性？康納曼的啟示



本文作者為神經科學研究所 吳仕煒教授

機器學習與人類心智的研究存在一個兩難。一方面，在電腦視覺，演算法一直要到深度學習發展應用後，才逐漸達到近似人類物體辨識能力的表現。但另一方面，機器學習在決策能力上，卻能很輕易超越人類的表現。主要是因為他們不會犯下許多人類會犯的決策錯誤。所以，機器學習的目的究竟是要肖似人類，還是要追求能力表現的極致？如果要像人，就要有產生人類決策偏誤的演算法。這其實不難，因為以康納曼（Daniel Kahneman, 代號K）和特沃斯基（Amos Tversky, 代號T）為首的心理學家和經濟學家，在1970到1990年代就已經把這個問題研究的很透徹了。其中的精髓在K&T提出的「展望理論」（prospect theory）（Kahneman & Tversky, 1979），內容不在此贅述，相信翻翻康納曼的「快思慢想」一書即可一窺其要。康納曼也因此於2002年得到諾貝爾經濟學獎。一個心Michael Lewis（同時也是魔球Moneyball一書作者）所寫的Undoing project（「橡皮擦計畫：兩位天才心理學家，一段改變世界的情誼」）。

康納曼於今年三月過世，享壽九十。他曾表示，「如果我現在要開始研究生涯，我會在人工智能和神經科學中擇一。它們是當今探究人性特別令人感到期待的方法。」（“If I were starting my career now, I would be choosing between artificial intelligence and neuroscience, because those are now particularly exciting ways of looking at human nature.”）。居然，透過AI我們可以了解人性，而且是非常令人興奮的方法！其實，康納曼不用擇一，因為神經科學和人工智能正走在一起。我們只要看看在2020年由Deep Mind和哈佛大學神經科學家Nao Uchida團隊合作，發表在「自然」期刊（Nature）的論文，就可略知一二（Dabney et al., 2020）。Deep Mind利用在AI強化學習發展出的新演算法，即「分佈式時間差學習」（distributional temporal difference learning），重新分析了Nao Uchida紀錄小鼠在酬賞學習作業過程中，大腦多巴胺神經元（dopamine neurons）的活動訊號，發現多巴胺神經細胞群體，像個執行分佈式時間差學習的機構，透過分析不同神經元的活動，研究者可重建小鼠依據經驗所習得環境中的酬賞分佈訊息。這項發現不僅改變了我們對於多巴胺神經元的了解，也是機器學習演算法被用來作為人類心智運作模型（Algorithms as models for the mind）的經典例子。

我總覺得康納曼那句話的重點更在他對於了解人性（human nature）的渴望。相信這也是許多人的渴望。如果是這樣，那他是不是點出機器學習走到最後，其實不在最佳化，而是人類為了要更了解自己所做的嘗試？我們真有可能因AI更了解人性？讓我們拭目以待。

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263-291.

Dabney W, Kurth-Nelson Z, Uchida N, Starweather CK, Hassabis D, Munos R, Botvinick M (2020)

A distributional code for value in dopamine-based reinforcement learning. *Nature* 577, 671-675.

照片擷取自：

[www.vanityfair.com/news/2016/11/decision-science-daniel-kahneman-amos-tversky](http://www.vanityfair.com/news/2016/11/decision-science-daniel-kahneman-amos-tversky)  
[willdabney.com/publication/distributions\\_from\\_dopamine/](http://willdabney.com/publication/distributions_from_dopamine/)

A distributional code for value in dopamine-based reinforcement learning

Will Dabney, Dan Kurth-Nelson, Sebastian Lee, Clark Aaron Starkweather, Eyal Seitzberg, Tom Schuster, Charles Botvinick  
January 2020



回上一頁 &gt;

▼ 展開

NYCU 國立陽明交通大學

☑ 校址：300093 新竹市東區大學路1001號 ☑

☎ 電話：+886-3-571-2121

從美國免費撥打：+1-833-220-6426

陽明校區

☑ 地址：112304 臺北市北投區立農街2段155號 ☑

☎ 電話：+886-2-2826-7000

交大校區

☑ 地址：300093 新竹市東區大學路1001號 ☑

☎ 電話：+886-3-571-2121