

給電腦大腦配塊記憶板

本課講 2014 年的一篇文章。DeepMind 公司的 Graves、Wayne 和 Danihelka 給神經網絡接上了一塊「記憶格子板」。他們管它叫神經圖靈機。

適合 10 歲左右

約 40 分鐘

需要：方格紙、便利貼、12 張卡片、2 位同伴

1 熱身：記在腦子裡，還是記在紙上？

請同伴讀一遍第 1 串詞，只讀一遍。讀完後，你把記住的詞按順序寫進格子。

第 1 串（只能記在腦子裡）：蘋果、火車、月亮、小貓、雨傘、鉛筆、西瓜、足球

第 2 串（可以邊聽邊寫）：香蕉、飛機、太陽、小狗、帽子、橡皮、葡萄、籃球

第 1 串								
第 2 串								

第 1 串記對了 _____ 個，第 2 串記對了 _____ 個。第 2 串容易得多，因為你有了一塊腦子外面的記憶板。

2 普通神經網絡：只能記在腦子裡

神經網絡就是電腦裡的「大腦」，由很多小零件連成。前面幾課的神經網絡，都把要記的東西揉在零件裡。第 1 課的小本子、第 2 課的傳送帶，都長在網絡裡面。零件既要算，又要記。

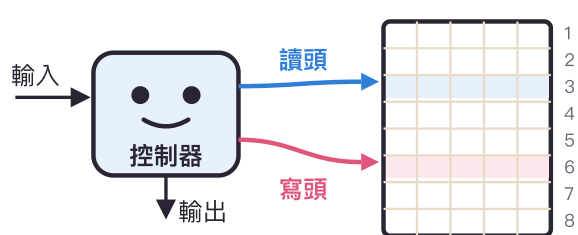
麻煩是：小本子的大小是固定的，東西一多就擠不下。想多記一點，只能把整個大腦做大。這就像為了記住更長的購物清單，先換一個更大的腦袋。能不能給神經網絡一張「草稿紙」？

3 電腦的辦法：一塊單獨的內存

你用的電腦，早就把「算」和「記」分開了。處理器（CPU）只管算。內存只管存，像走廊裡一排儲物櫃。每個櫃子都有號碼，號碼叫地址。電腦說「打開 37 號櫃子」，就能直接讀或存。這叫按地址讀寫。

1936 年，數學家圖靈設想了一臺機器：一條長紙帶分成很多格子，一個「頭」能讀、寫、往左右挪一格。它叫圖靈機。神經圖靈機的名字就從這裡來。

4 神經圖靈機 = 大腦 + 記憶格子板



控制器：一個神經網絡，就是大腦。它決定讀哪裡、寫什麼、往哪挪。

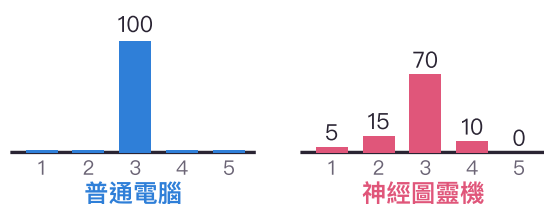
讀頭和寫頭：大腦伸出去的兩隻「手」。讀頭負責讀。寫頭負責寫，分兩步：先擦掉一點舊的，再寫上新的。

記憶格子板：大腦外面的一塊板子，像一張方格紙。論文實驗裡，板子有 128 行，每行 20 個小格子，每格放一個數。科學家管它叫**記憶矩陣**。想多記一點，換一塊更大的板子就行，不用把大腦做大。

5 難點：讀寫頭必須是「模糊」的

普通電腦一次只讀一個格子：要麼讀第 3 格，要麼讀第 4 格，沒有中間狀態。可是神經網絡學本事，靠的是第 1 課說的「猜錯了就悄悄調整自己」，也就是把很多小旋鈕各擰一點點（第 4 課講過旋鈕）。

只讀一個格子的讀頭，像只有「開」和「關」的電燈開關。旋鈕擰一點點，讀到的還是同一格，電腦得不到任何提示。擰過了頭，又一下子跳到別的格子。所以它沒法一點點學。



神經圖靈機的讀寫頭像**調光燈**。它把 100 份注意力分給所有格子。大部分集中在一兩個格子，其他格子只分到一點點。

讀出來的東西，是各格內容按份數混在一起。寫的時候，份數多的格子改得多。這樣，旋鈕擰一點點，讀到的也變一點點，電腦就知道往哪邊調。這種「分份數」的辦法，就是第 13 課和第 15 課講過的**注意力**。那兩課的注意力只能讀，這裡的格子還能寫。

6 找格子的兩種辦法

找法	怎麼找	生活裡的例子
按內容找	說出「我要找長這樣的」，跟每一行比一比。越像的行，分到的注意力越多。	在操場上找「戴紅帽子的人」
按位置找	從上一步的位置出發，往前或往後挪一格，或者不動。	讀完這一行，就讀下一行

兩種找法可以搭配：先按內容找到一個東西，再按位置往後讀。每一步多用哪種，是大腦自己學會的。挪格子時注意力會散開一點，像手電筒的光變淡，所以它挪完還會「收一收」。

7 實驗：它自己學會了小程序

科學家沒給它寫程序，只給它看很多題目和正確答案。論文測了 5 個任務，這裡講 4 個：

- **複製**：給它一串東西，讓它原樣說一遍。每個東西是一排 8 個小燈，亮滅隨便排。
- **重複複製**：給它一串東西和一個次數，讓它連著複製幾遍。
- **聯想回憶**：給它好幾組東西，再拿出一組問：「這組後面是哪一組？」
- **排序**：每個東西帶一個分數，讓它把分數高的那些從高到低排出來。

複製任務：練短的，考長的

練習時，每串只有 1 到 20 個東西。考試時，串加長到 120 個。

串有多長	神經圖靈機	LSTM（第 2 課）
1 到 20 個（練過）	學得更快，錯得更少	也能複製，錯得多一些
超過 20 個（沒練過）	到 120 個還大體正確	很快就出錯

科學家把讀寫頭每一步的注意力畫出來，發現它學會的做法很像一段程序：



只要格子板還有空格，這段程序不管串多長都管用。LSTM 把東西揉在零件裡，串一長就擠不下了。

8 這和今天有什麼關係

說實話，今天幾乎沒人直接用神經圖靈機。它練起來不穩定，速度慢，很難做大。可是它的想法活了下來：

- **按內容找**：ChatGPT、Claude 這些聊天 AI 的核心零件裡，用的就是同樣的找法，叫注意力。
- **外接記憶板**：有的 AI 先去資料庫裡查東西，再回答。有的 AI 幹長活時，把中間結果寫進文件，再讀回來。
- **練短的，考長的**：今天的 AI 還不擅長。比如學過短的加法題，換成位數很多的加法，常常算錯。

9 動手：記憶格子板複製遊戲

玩法（3 個人，約 15 分鐘）

準備：方格紙上畫一系列 10 個格子，第 1 格畫 ★。粉色便利貼當寫頭，黃色當讀頭，都貼在第 1 格旁。A 出題，B 當控制器，C 檢查。B 只照程序卡做，**不許用腦子記**。

程序卡

寫：聽到一個符號，畫進寫頭指的格子，寫頭往下挪一格。

聽到「完」：讀頭找到畫著 ★ 的格子（按內容找）。

讀：說出讀頭指的格子裡的符號，讀頭往下挪一格。遇到空格就停。

1. A 慢慢說 4 個符號，比如「圓、三角、星、方」，再說「完」。B 照卡寫、照卡讀，C 檢查。
2. 擦乾淨，A 換 8 個符號再來一遍。程序卡不改。
3. 最後一輪不用格子板：A 說 8 個符號，B 用腦子記，再說出來。

記一記：B 說錯了幾個？第 1 輪 _____ 個，第 2 輪 _____ 個，第 3 輪 _____ 個。

你會發現：同一張程序卡，4 個和 8 個符號都能用。只靠腦子記，8 個就容易錯。

10 動手：按內容找 vs 按位置找

玩法（2 到 4 個人，約 10 分鐘）

準備：12 張卡片，每張畫一種動物，比如貓、狗、兔子、大象、魚。正面朝上排成一排，**不寫號碼**。一人念題，其他人搶答，一人計時。比完再自己出 3 道新題。

題目	用的找法（圈一個）	用了幾秒
從左數第 9 張是什麼動物？	內容 / 位置 / 兩種	
有長鼻子的動物在哪張？用手指出來。	內容 / 位置 / 兩種	
找到貓，再往右挪 2 張，是什麼動物？	內容 / 位置 / 兩種	

你會發現：找「長鼻子」一眼就看到，這是按內容找。找「第 9 張」得一張張數，這是按位置找。有的題要兩種合起來。神經圖靈機每一步也要選用哪種。

11 詞語小卡片

控制器

神經圖靈機的大腦，決定讀、寫和挪動。

記憶矩陣

大腦外面的記憶格子板。

讀頭和寫頭

大腦的兩隻「手」，一隻讀，一隻寫。

模糊讀寫（注意力）

注意力按份數分給很多格子。

12 想一想

1. 你的書包、鉛筆盒、記事本，哪個最像神經圖靈機的格子板？為什麼？

2. 讀頭如果只能「要麼這格要麼那格」，電腦為什麼很難練會？用電燈開關和調光燈說一說。

3. 在班裡找同學時，什麼時候你按內容找？什麼時候按位置找？各舉一個例子。

一句話記住：神經圖靈機給電腦大腦接上一塊記憶格子板，用模糊的讀寫頭一點點學會讀寫。

給家長和老師：本課講的是 Alex Graves、Greg Wayne、Ivo Danihelka 2014 年的論文 [Neural Turing Machines](#)。論文給控制器（前饋網絡或 LSTM）接上外部記憶矩陣（實驗用 128×20 ），讀寫都用對所有行的連續權重分佈完成，整個系統可用梯度下降端到端訓練。尋址分 4 步：內容尋址、插值、循環卷積移位、銳化。5 個任務是複製、重複複製、聯想回憶、動態 N 元語法和優先級排序。複製任務訓練長度 1 到 20，測到 120 時神經圖靈機仍大體正確，LSTM 超過 20 後很快出錯。

講義做了簡化：權重之和為 1，講義說成「100 份注意力」；擦除和添加說成「先擦掉一點舊的，再寫上新的」；銳化說成「收一收」；排序任務實際是輸出優先級最高的若干個向量。講義沒有講公式。熱身詞表、柱狀圖裡的份數（5、15、70、10、0）、8 行格子板圖、電燈開關和儲物櫃比喻、兩個動手遊戲都是我們編的示意，不是原文內容。真實的格子板裡存的是數，不是文字或符號。想深入可看本站的 [第 17 天中文導讀](#)，代碼見開源項目 [30u30](#) [第 17 天目錄](#)（英文）。