

科學講堂

逢星期三見報

借鑑大自然 唯「製作」需時 出錯率高

研DNA 儲存海量數據

我們經常說現在是「大數據」的時代，因為我們每天的活動都在不停地製造新的數據：在YouTube收看及上載短片、在各社交平台分享的高清照片、到網上購物平台搜購心愛的物品、或是每天使用八達通卡四處遊歷的記錄。

2014年總數據佔百億張藍光碟

《自然材料科學》雜誌(Nature Materials)去年3月刊載的一篇文章指出，2014年全世界共累積儲存了10億TB

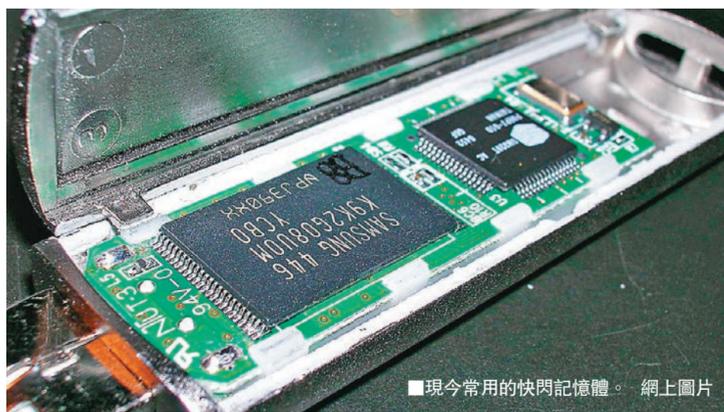
(terabyte)的數據，大約等於100億張藍光光碟的容量(一張3層BDXL格式藍光光碟的容量大約為100GB)。倘若依照過去增長的趨勢去預測，到了2040年，全球累積的數據數量將會再增多100倍。大家可有想過，如何能夠儲存如此大量的資訊？

除了使用光學原理運作的各種光碟外，現今我們很多數據都是儲存於電腦硬盤(hard disc drive)和各種快閃記憶體(flash memory)中。硬盤主要是運用磁力來記取數據，令硬盤上的不同位置產生磁

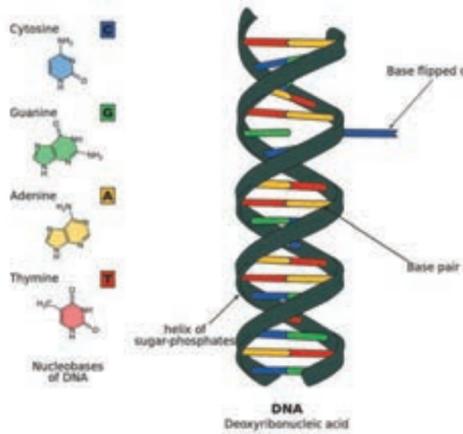
性就代表了不同的數據；快閃記憶體則利用了半導體(semiconductor)物料的特性，因此矽(silicon)這種半導體就成為了近年來需求殷切的材料。

上一段提過的文章甚至指出，到了2040年，矽的全球累積產量並不能用來製造足夠的快閃記憶體去儲存如此大量的數據。再者，要收藏好這多的光碟、記憶體或硬盤，也很是令人頭痛。

(*編者註：標題為「Nucleic acid memory」，第一作者為Victor Zhirnov。)



■現今常用的快閃記憶體。網上圖片



■DNA呈雙螺旋狀，由A、C、G、T4種「英文字母」串連而成。網上圖片

《自然》文章：存全世界數據 僅需一公斤DNA

世界各地的一些科研人士(包括之前提過的Victor Zhirnov、英國的Nick Goldman、哈佛大學的George Church)卻有一個很有趣的靈感：大自然是如何儲存數據的？操控各種生命如何成長的遺傳密碼，不就是有效率地以DNA儲存於每個細胞核中？

人體的基因圖譜，可以看成是30億個串連在一起的「英文字母」，這樣大量的資訊，現正收藏於我們每顆連肉眼也看不見的細胞之中。

去年8月《自然》雜誌的另一篇報道也提到，倘若利用大腸桿菌基因儲存數據的密度來估算，要儲存

全世界的數據，只需一公斤的DNA就足夠了。

那麼這個主意是否可行？其實科研人員早已掌握了按照自己心意製作DNA的技術，不過成本太高、需時太久——製造一個DNA「英文字母」需時大約400秒，而且出錯率

還未到令人放心的地步。我們還需要繼續想一些減低錯誤或自行修正錯誤的方法。



■硬盤運用磁力來記取數據。網上圖片

小結

正如船隻流線型的設計取材於魚類身體的形狀，照相機的設計跟眼睛有幾分相像，看來人類又再一次在向自然學習。

當然用DNA儲存數據的意念還在起步階段，還有許多問題需要解決。不過，我們不時聽到，數十萬年前古生物的DNA還能被現在的我們檢測到。試想想：用DNA來做我們電腦的記憶裝置，那會是多麼的耐用啊！

■張文彥博士

作者簡介：香港大學土木及結構工程學士。短暫任職見習土木工程師後，決定追隨對科學的興趣，在加拿大多倫多大學取得理學士及哲學博士學位，修讀理論粒子物理。現於香港大學理學院任職講師，教授基礎科學及通識課程，不時參與科學普及與知識交流活動。

奧數揭秘

逢星期三見報

談未知數代換

天氣炎熱的日子，被太陽曬得發滾，若果一杯冷冷的水果沙律，給自己拿着涼快一下，感覺倒是爽得很。想起買水果，買的時候當然知道價錢，不過事後只記得自己花了多少錢，忘了每個水果價錢的話，那麼有沒有方法計出來？

比如陳太和周太各買了一些水果，陳太買了5個梨和4個蘋果，共29元；周太買了3

個梨和2個蘋果，共18元。那樣設梨為x元，蘋果為y元，就得到以下方程組：

$$\begin{cases} 5x+3y=29 \dots(1) \\ 3x+2y=18 \dots(2) \end{cases}$$

由(1)×2-(2)×3，可解得x=4，然後代入(1)，可得y=3。

這個是中二程度的課內數學，奧數程度當然不止是這樣。

問題

解方程組：
$$\begin{cases} \frac{5}{u+v} + \frac{3}{u-v} = 29 \dots(1) \\ \frac{3}{u+v} + \frac{2}{u-v} = 18 \dots(2) \end{cases}$$

答案

若是把等式左邊通分母加起來，出來的結果可以很煩的，有2次方。留意方程組的各個數字，比較一下上邊買梨和蘋果的方程組，就會看到類似的東西。原來只要設 $x = \frac{1}{u+v}$ 和 $y = \frac{1}{u-v}$ ，問題就跟之前的一樣，於是得 $\frac{1}{u+v} = 4$ 和

$\frac{1}{u-v} = 3$ 。整理後得如下方程組：
$$\begin{cases} u+v = \frac{1}{4} \dots(3) \\ u-v = \frac{1}{3} \dots(4) \end{cases}$$
由(3)+(4)，得 $u = \frac{7}{24}$ 。由(3)-(4)，得 $v = -\frac{1}{24}$ 。

接通其他代數技巧 增解難能力

以上的解難的關鍵，用到了代換的方法，即是把 $\frac{1}{u+v}$ 和 $\frac{1}{u-v}$ 整個兒當成了x和y。這個技巧和課內其實也有多少涉獵的，事實上較難的中二題目，也有代換的部分，只是多數是一道較長的題目的第二部分，而第一部分就好像上邊買水果的情景的那組方程差不多。奧數問題中，用到代換的時刻是很多的，而難處之一就是看不出要用代換，而且即使覺得用着可能有幫助，也想不到把哪一堆當成是未知數好。

用代換的想法，可以延伸出很多問題的，比如買水果情景中的方程組，設 $x = a + b$, $y = ab$ ，又會出了道看來複雜的難題：

$$\begin{cases} 5(a+b)+3ab=29 \\ 3(a+b)+2ab=18 \end{cases}$$

這樣 $a+b=4$, $ab=3$ ，可用一元二次方程，解出 $\{a,b\} = \{1,3\}$ ，注意a和b可以互換。

現在說穿了技巧，看剛剛這一道題，也沒怎樣的。不過初初接觸的話，大概有不少學生都會做不來。

未知數的代換，一方面能夠把較複雜的問題，分拆成較簡單的表達形式，使計算變得簡便；另一方面，同時亦是出題的方法，使自己能夠在代換之中，接通其他代數技巧，增加自己解決問題的能力。良好的代換，對解難的效果，跟沒有做代換可以差別極大。

代換的原則，其實就是觀察到算式之中，有類似的部分，好像問題中的 $\frac{1}{u+v}$ ，在兩道算式中都出現，然後就聯想到把它當成一個未知數x，是一個不錯的嘗試。

平常解難之中，若看到算式不同部分都有類似的項，也可以嘗試代換，這就要自己多觀察了。

■張志基

科技暢想

逢星期三見報

影像處理系統的閾值可由系統使用者自行設定。所謂閾值同時是指系統可接受的準確度。當系統計算的估算結果與輸入相片之小於系統的閾值，系統便會採納此估算結果；相反，此相片便會是「無法辨識」。

閾值可調整 以拓展應用

在保安角度，英國倫敦警方用了CCTV輔以人面辨識技術逮捕疑犯。這樣系統的閾值便需要很少，即系統準確度要很高，高於90%以上。

相反在社交媒體，例如facebook的人面相片標籤功能(facial recognition photo tagging)，用戶的相片會出現建議標籤的朋友名單，所需要系統準確度就不需要太高。正正因為系統的閾值可由系統開發者自行設定，把人面辨識技術由保安擴展至社交及其他用途。

辨識VIP 活動後share相片

無可否認，CCTV加上人面辨識技術大大增加了警方捉拿疑犯的機會，同時此技術正向其他方向發展。維珍航空(Virgin Atlantic)於2014年就把此技術放至Google Glass，讓他們的空中服務員向VIP客人打招呼。在未來，人面辨識

「認人樣」不再限保安



■fb的人面辨識技術，替用戶快速標籤相片。網上圖片

技術有望擴展至一些PR公司或活動主辦單位(event organizers)。

試想像我們去到一些大型活動，主辦單位在活動門口只需放置鏡頭且使用人面辨識技術，除了可以知道哪一個客人/VIP已到現場；同時在活動過後，更可以把活動所拍攝的影像和相片自動分類且傳送到相關活動參與者。

馬拉松比賽現時就使用了RFID(Radio Frequency Identification) Race

Timing System記錄參賽者到達終點的時間。主辦單位在比賽後，會把每位參賽選手到達終點前後的相關影像和相片傳送給此選手。

未來，可否使用人面辨識技術呢？主辦單位只需放置鏡頭，辨識選手且於比賽結束後把相關相片影像傳送至相關選手，又是否可行呢？

(二之二)

■香港新興科技教育協會 姚敏婷

簡介：本會培育科普人才，提高各界對科技創意應用的認識，為香港年輕人提供更多機會參與國際性及大中華地區的科技創意活動，詳情可瀏覽www.hknetea.org。



氣象萬千

隔星期三見報

青藏高原千里「呼風喚雨」

青藏高原是世界上最高的高原，平均海拔高度4,500米，面積250萬平方公里，有「世界屋脊」及「第三極」之稱。高原的邊界，東面是橫斷山脈，南面和西面是喜馬拉雅山脈，北面是崑崙山脈。雖然青藏高原離香港很遠，但原來它對我們的天氣有很大的影響。

加熱作用 令華南降雨

夏季高原的日照強烈，對近地面的大氣加熱作用較強，近地面空氣受熱膨脹，氣壓降低，夏季季候風由氣壓較高的海洋吹向氣壓較低的陸地。青藏高原的加熱作用進一步降低了陸地氣壓，加強了夏季季候風的威力。

冬季高原的散熱較快，出現冷高壓，冬季季候風由氣壓較高的陸地，吹往氣壓較低的海洋。青藏高原的冷高壓，加強了陸地的氣壓，增強冬季季候風的威力。

青藏高原本身是一道牆，會阻擋氣流的流動。在冬季亞洲大陸的上空，通常都是受到西風橫流的控制，青藏高原會阻擋西風的去路，高空的西風遇上高原時，分為南北兩支。北部形成高壓脊，南部形成低

壓槽。南支槽經過華南的時候，若其他天氣系統配合，可能引發降雨，有時在春天甚至引發暴雨。

它亦會阻擋西伯利亞的冷空氣，冬天時，令南面的印度不受冬季寒潮侵擾。南下的冷空氣被迫在高原的東面向南移動，導致寒潮繞道侵襲中國東面偏南的範圍。

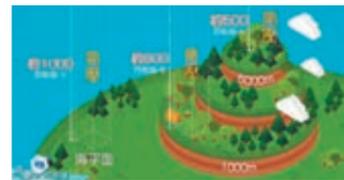
高原氣壓低 礙呼吸煮食

大家如果有興趣到訪青藏高原，看看它的真面目。記住要小心高山反應，距離海平面越高，空氣越稀薄，氣壓越低。

一般海平面的氣壓約1,000百帕斯卡，在海拔高度大約1,000米的大帽山，氣壓是海平面的90%，我們吸入的氧氣量也是海平面的90%；海拔高度5,000米，氣壓只是海平面的一半，相信在這個高度，大部分人都會有高山反應，高山反應有可能

致命。大家出發去高原之前，記得做足資料搜集，徵詢專業人士的意見。

大家在高原煮食也要注意，大家知道水的沸點是100℃，但這是在海平面的情況下。水的沸點溫度和大氣壓力的高低是成正比，即是說，氣壓越高水的沸點越高。反之亦然，在海拔高度5,000米，水的沸點下降至83℃，意即開水的溫度可能只有83℃，如果在高原煮食，記得要確保食物熟透，方才進食呀！



■氣壓隨高度改變。視頻截圖

簡介：本欄以天文台的網上氣象節目《氣象冷知識》向讀者簡介有趣的氣象現象。詳情可瀏覽天文台YouTube專頁：https://www.youtube.com/user/hkweather。

《青藏高原》

