

化合物千億種 電腦幫選新藥

科學講堂

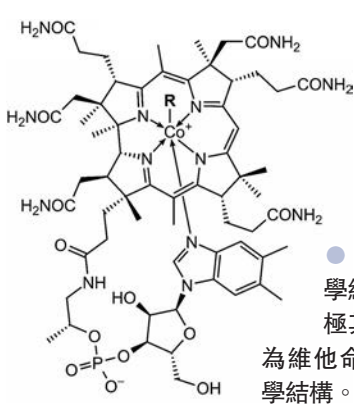
面對新冠疫情，各大藥廠都加緊開發及優化相對應的疫苗與藥物，不知有否令大家回想起研發藥物和疫苗，其實是一個漫長而困難的過程？撇開臨床測試不談，單是尋找適合測試的藥物就是一份苦差：根據化學家的估算，有藥用功效的化合物種類，比太陽系裏所有原子加起來的總數還要多，因此單單找出並優化拿來測試的藥物，往往就需要四五年。近年電腦能力愈來愈強，機器學習的技術愈來愈成熟，研究人員又如何利用它們去加速藥物的開發呢？今次就和各位分享一下這方面的發展。

收窄「關注」範圍 挑選止痛「候選人」

如前所述，有可能成為藥品的化合物實在是太多了，因此不少研究人員都希望電腦能夠幫忙挑選，或者起碼收窄「關注」的範圍。要順利進行這個「任務」，一個完善的化學品數據庫自然不可少。瑞士伯爾尼大學 (University of Bern) 的化學家 Jean-Louis Reymond 早在 2001 年，就已開始收集化學品的資料。

到了 2017 年，這個數據庫已儲存了 1,660 億種化合物的資料。如此龐大的數據庫，收藏了最多由 17 顆原子構成的有機化合物的數據；僅是整理出數據庫中化學品的名單，電腦就可能需要超過 10 小時。

為了更方便利用這個數據庫，Reymond 將其中的化學品依照 42 個特徵來分類 (例如化合物中有多少個碳原子)。須知不少化合物都有相類的化學與藥物特性，因此縱然我們已認定需要的特徵，但還得慢慢地尋找最合適的化學結構。Reymond 經過分類的數據庫，正正可以用於這個方面：



● 藥品的化學結構，可以極其複雜。圖為維他命 B12 的化學結構。網上圖片

只要輸入一個化合物，電腦就可在 3 分鐘從 1,660 億種化合物中找出相類的化合物，好讓我們更容易從中挑選，然後再進行測試。加州大學的化學家 Brian Shoichet 在 2016 年就利用電腦檢視三百萬種市面上有出售的化合物，從中找出 23 個適合用來發展新的、更少副作用的止痛劑的「候選人」。



● 開發藥物時，除了要考虑效用，還要考虑生產難度等問題。

資料圖片

改善數據資料 考慮製造難度

除了運用電腦來儲存化學品的特性，近年發展十分成熟的機器學習也已被應用於藥品開發之上。在利用我們已有的知識去訓練以後，電腦可以尋找當中的規律，再向我們提議可作深入研究的化學分子；電腦也可以預先警告我們不適合的化學結構 (比如說毒性太猛、不太可能通過毒理測試的化學品)，好讓我們少走點冤枉路，

以免浪費精力。

不過，利用機器學習，暫時還有它的不足之處。例如說電腦在提出建議的時候，不一定會把合成化學品的難度也一併考慮，以致部分提議的化合物，可能極難製造。為此 Reymond 已整理出另一個版本的數據庫：當中只有一千萬種化學物，但都是容易在實驗室中合成的。除此之外，

電腦建議的精確度，自然亦有改善的空間。這就需要大家努力地提供更多更好的資料和數據，好讓電腦能學得更精準。

運用電腦來開發新藥物，雖然還未達至完美，但已經可以省卻人們不少工夫。在不久的將來，更好的發展方向可能是人機合作，利用相互的長處來研發出最好的藥品。

● 杜子航 教育工作者

早年學習理工科目，一直致力推動科學教育與科普工作，近年開始關注電腦發展對社會的影響。

巧算

奧數揭秘

這次談一道巧算題，數字看來有夠複雜，全是四位數，加減乘除兼平方，想過題簡潔一點，就得費點功夫。

問題：計算 $\frac{(1986^2 - 1992)(1986^2 + 3972 - 3)(1987)}{(1983)(1985)(1988)(1989)}$

答案：設 $x = 1986$ 。

$$N = \frac{(x^2 - x - 6)(x^2 + 2x - 3)(x + 1)}{(x - 3)(x - 1)(x + 2)(x + 3)}$$

$$= \frac{(x + 2)(x - 3)(x + 3)(x - 1)(x + 1)}{(x - 3)(x - 1)(x + 2)(x + 3)}$$

$$= x + 1$$

$$= 1987$$

解題的關鍵，大概就是把幾個差不多大的數字，設其中一個為 x ，那其他數字都可以用 x 來表示，之後在代數層面上化簡，就容易看到有因式分解的部分，然後就可以約簡。而設哪個數為 x ，選擇也挺明顯的，就是算式當中最麻煩的數字，即是有平方的，剛好平方下的都是 1986，於是選取 1986 為 x 。

做完之後，固然發現設了 x 為 1986 會比較方便，但做的時候難免也會把數字拆開看看，有時看到數字太複雜，未必立即就想到能用十字相乘做因式分解。雖然題目本身未算太複雜，但也不是直接到一眼就看得通所有過程。

數字代數可互通

這種把數字設成 x 的問題，在教學上有個好處，就是能夠讓學生明白數字和代數之間，不時可以互通。至於哪一種的表達方式會令人更易看得清當中的數理，這就不一定了，還得兩邊都看得通，再比較才行。

在高班的時候，看代數和數字之間互換，可能還自然一點，但初中的時代，還是有不少人未懂得運用恒等式來解的。

這些巧算的問題，在奧數裏算是比較淺，多少用來鼓勵一些基礎不錯的學生，令他們有點滿足感，拔尖或者篩選的功能比較少。

有時與數班的氣氛過於沉寂，原因大概是因為問題太難了。如果在學生的眼中，題目的程度多少是可以開始嘗試的，有點想法的，那會比較容易起動，氣氛也會好一點。

過於高深的內容，學生學了固然是好，但水平未到的時候，內心對求學的熱誠還不夠，面對難題時，願意挑戰的人少，知難而退的人多，氣氛就會差了一點。

淺的問題混合着深的，學生學起來也舒服一點，久不久又有點成就感，那樣士氣又好些，快樂也更多。教數學時，令學生感覺到做數學跟快樂多一點關係，跟痛苦少一點關係，是挺重要的。

長遠來說，還得學生在生活中想起數學，才會讀多一點。在他們眼中，讀數學是快樂的，就會主動去想多一點。

數學看起來只是紙上的知識，只是在教與學之中，令學生有興趣學習，才是關鍵所在。

● 張志基



● 「綠得開心最傑出大使」詹曉雯同學親身體驗水耕的種植方法。

作者供圖

天台體驗水耕 學習環保意義

綠得開心@校園

環境保護近年逐漸成為全球各地關注的重要議題，全球暖化的問題迫在眉睫，為人類生活及地球生態帶來源源不絕的影響。保良局羅傑承 (一九八三) 中學致力推行環保教育，鼓勵學生從小建立綠色生活習慣，減少碳排放，為地球未來、可持續發展出一分力，履行公民責任，從根源上減少人類活動對地球的影響。

自 2019/2020 學年起，我校積極改善校園天台環境，並建設有機水耕場地，給予學生親身體驗水耕的種植方法。水耕場地種有羅馬生菜、菜芯、羽衣甘藍、小白菜等品種，並採用全有機種植，如運用苦楝油清除植物上的害蟲，而非採用殺蟲劑。在活動中，學生學習到水耕方法的原理、水循環、收割菜類的方法及其環保意義。此外，同學



● 師生齊參與校內環保發電單車吸氣候變化教育活動。

作者供圖

為學校資訊日拍攝了一段關於校園水耕的介紹短片，向公眾宣傳水耕及其好處，很有教育意義。

我於 2019 年獲環保組織贈予一架環保發電單車，並於 2021 年 4 月舉行一連三日的發電單車工作坊，同學在活動中能夠親身嘗試以腳踏單車的方式發電，為電腦、風扇等電器供電運作，同場亦擺放有關全球暖化的展板。是次活動讓同學學習到課外的環保知識，亦親身體驗了發電單車的運作，明白到電力得來不易，要好好珍惜地球資源。

未來保良局羅傑承 (一九八三) 中學將繼續秉持環保理念，以多元化的學習活動推動環保教育，從根源著手，師生共同宣揚綠色生活，為下一代打造更美好的地球環境。

● 保良局羅傑承 (一九八三) 中學

(港燈「綠得開心計劃」「綠得開心學校」之一，兩位同學於 2020 年參加「綠得開心推廣大使計劃」。當中詹曉雯同學更獲選為「最傑出大使」。) 港燈「綠得開心計劃」，致力透過多元化活動，協助年輕一代及公眾人士培養良好的用電習慣、多認識可再生能源和實踐低碳生活，目前已有五百多間全港中小學校加入「綠得開心」學校網絡。如欲加入一同學習和推動環保，歡迎致電 3143 3727 或登入 www.hkelectric.com/happygreencampaign。

簡介：奧校於 1995 年成立，為香港首間提供奧數培訓之註冊慈善機構(編號：91/4924)，每年均舉辦「香港小學數學奧林匹克比賽」，旨在發掘在數學方面有潛質的學生。學員有機會選拔成為香港代表隊，獲免費培訓並參加海內外重要大賽。詳情可瀏覽：www.hkmos.org。

