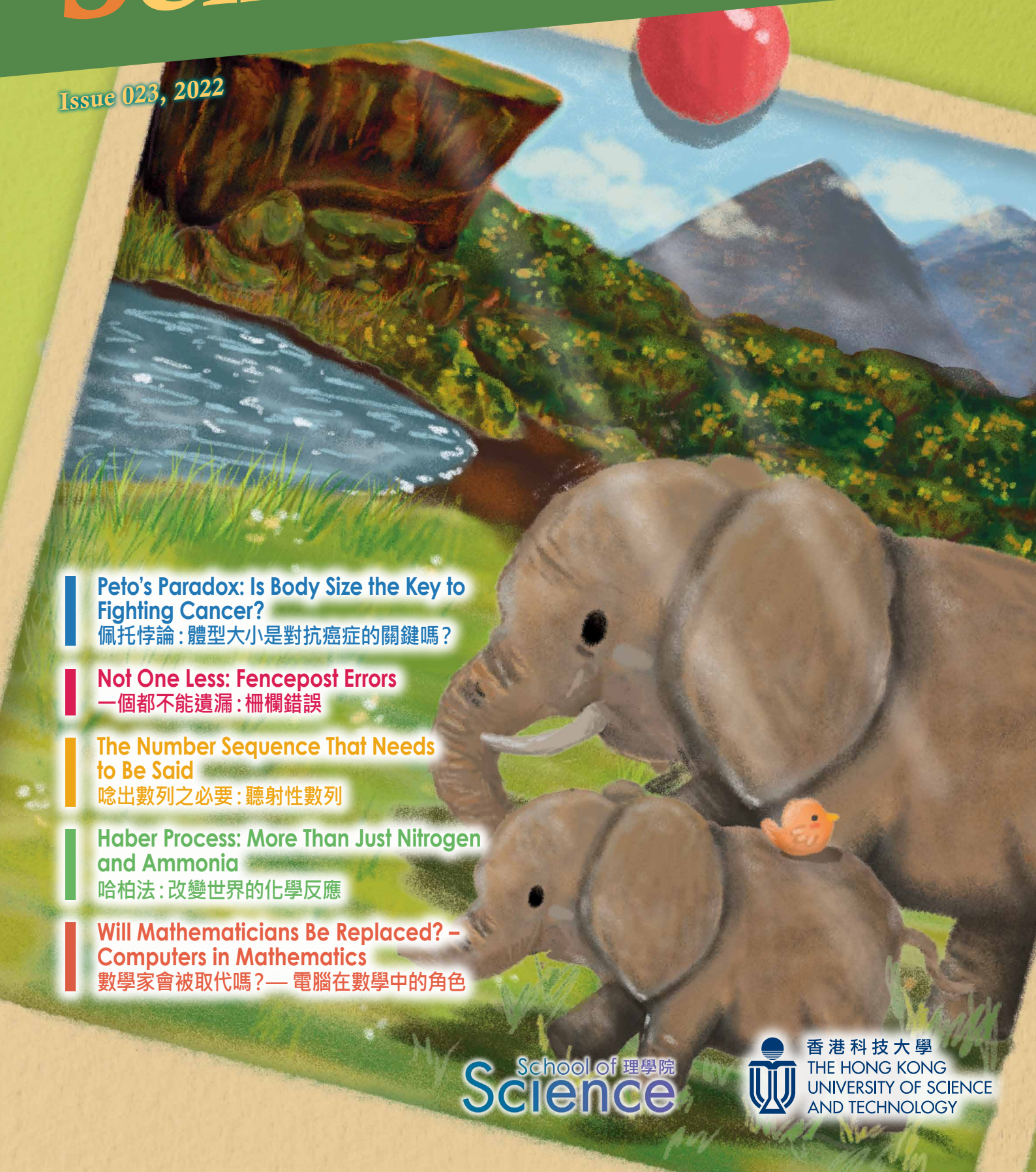
SCIENCE FOCUS 科言

Issue 023, 2022



Peto's Paradox: Is Body Size the Key to Fighting Cancer?
佩托悖論：體型大小是對抗癌症的關鍵嗎？

Not One Less: Fencepost Errors
一個都不能遺漏：柵欄錯誤

The Number Sequence That Needs to Be Said
唸出數列之必要：聽射性數列

Haber Process: More Than Just Nitrogen and Ammonia
哈柏法：改變世界的化學反應

Will Mathematicians Be Replaced? – Computers in Mathematics
數學家會被取代嗎？— 電腦在數學中的角色

School of 理學院 Science
香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Contents

Science Focus Issue 023, 2022

Message from the Editor-in-Chief 主編的話

Dear Readers,

Welcome to the new issue of *Science Focus*. In addition to being a source of your "book reports", I hope this magazine can serve as a starting point for your further exploration of specific topics. When our writers prepare each article, they normally consult multiple references to ensure the accuracy of the content. I hope you will develop the same habit during your personal journey of discovery.

Did you know that elephants are much less likely to suffer from cancer than us? Find out why from our cover story on Peto's paradox when we discussed the surprising lack of correlation between body size and cancer. One way to treat cancer is to undergo surgery. It turns out we have Joseph Lister to thank for popularizing antiseptic surgeries, which greatly increase the survival rate of patients. Turning to mathematics, we consider a seemingly simple quest of counting objects or birthdays and how it can go wrong. We will also explore how the concept of radioactive decay can be applied to a mathematical problem. For those of you who enjoy horticulture, we revisit how fertilizers are made using the Haber process. Is there more to learn beyond the chemistry textbook? Finally, don't miss our articles on the forefront of scientific breakthroughs, from using light to control brain activities to using artificial intelligence for mathematical proofs.

As always, please follow our social media channels for additional bite-size scientific stories. Drop us a line in the comments section. We would love to hear from you!

Yours faithfully,
Prof. Ho Yi Mak
Editor-in-Chief

親愛的讀者：

歡迎閱讀最新一期《科言》！除了可能作為閱讀報告的「指定書目」外，我希望這本雜誌能為你提供探索科學題材的切入點。當我們的學生編輯預備每篇文章時，他們通常都會參考多篇文獻以確保內容的準確性，希望大家踏上科學旅程時也能養成同樣習慣。

你知道大象患癌的機率遠比我們小嗎？箇中原因可以在關於佩托悖論的封面故事中找到，我們會討論體型和癌症背後出乎意料地沒有關聯的原因。進行手術是治療癌症的其中一個方法，說到手術，我們要感謝 Joseph Lister 使抗菌程序普及化，這大大增加了病人手術後的存活率。也讓我們來談談數學，看看我們如何在的數算物件或生日的過程中出亂子，亦窺探一下數學家如何把放射性衰變的概念應用在數學問題上。此外，我們亦為喜歡園藝的你重溫一下透過哈柏法製造肥料的過程，教科書以外還有沒有甚麼是值得學習的呢？最後，不要錯過介紹最新科學突破的文章，當中包括運用光來控制腦部活動和利用人工智能來寫數學證明的技術。

一如以往，我們歡迎大家追蹤我們的社交專頁，那裡會連載一些雜誌以外的科學小故事。我們亦期待在網上與大家互動，歡迎隨時留言給我們！

主編 麥晧怡教授
敬上

Copyright © 2022 HKUST E-mail: **sciencefocus@ust.hk** Homepage: **http://sciencefocus.ust.hk**

What's Happening in Hong Kong? 香港科技活動

Fun in Fall Science Activities 秋日科學好節目

Any plans for this fall? Check out these activities!

計劃好這個秋天的好去處了嗎？不妨考慮以下活動！

Dinosaurs of Antarctica 3D 極地尋龍3D

The now frozen Antarctica was once warm and forested hundreds of millions of years ago with dinosaurs roaming freely. There was fierce competition for survival on the continent. Let's follow a team of paleoecologists to examine the fossils of those prehistoric dinosaurs and witness the dramatic transformation of the southern continent through this 3D dome show which utilized next-level computer graphics to reconstruct the continent's hidden-greener-past.

Show period: July 1, 2022 – March 31, 2023

Time: 2:00 PM and 6:30 PM on Monday, Wednesday, Thursday and Friday (except public holiday)
12:30 PM and 5:00 PM on Saturday, Sunday and public holiday

Venue: Space Theatre, Hong Kong Space Museum

Admission fee: Standard admission: $32 (stalls), $24 (front stalls)
Concession admission: $16 (stalls), $12 (front stalls)

Remark: Please refer to the museum's website for more details.

現時冰封的南極洲在數億年前曾是氣候溫和的森林，有著無數恐龍在叢林裡自由遊走，然而要在猛獸橫行的南極洲生存是非常艱辛的任務。就讓我們藉著這套利用了頂尖電腦特效去重塑南極洲過往翠綠面貌的 3D 球幕電影，隨古生態學家去檢視那些史前恐龍的化石和見證南極洲的巨大變化吧！

展期：2022年7月1日至 2023 年3月31日

時間：星期一、三、四及五（公眾假期除外）下午二時正及六時三十分
星期六、日及公眾假期下午十二時三十分及五時正

地點：香港太空館天象廳

入場費：標準票：32 元（後座）；24 元（前座）
優惠票：16 元（後座）；12 元（前座）

備註：更多詳情請參閱太空館網頁。

The Shaw Prize 2022 Exhibition 2022邵逸夫獎展覽

Established in 2002, the Shaw Prize recognizes currently active scientists with recent significant breakthroughs in scientific research. Three prizes are awarded annually in Mathematical Sciences, Life Science and Medicine, and Astronomy. In the exhibition, you can learn more about the Shaw Laureates of this year and their scientific achievements.

Date: September 30, 2022 – November 30, 2022

Venue: G/F Exhibition Hall, Hong Kong Science Museum

邵逸夫獎於 2002 年設立，設有三個獎項：數學科學獎、生命科學與醫學獎和天文學獎，每年頒發予現時活躍於研究工作並在近期取得突破性成果的科學家。透過今次展覽，你可以更深入認識今年的得獎者以及他們的科研成就。

展期：2022年9月30日至 2022 年11月30日

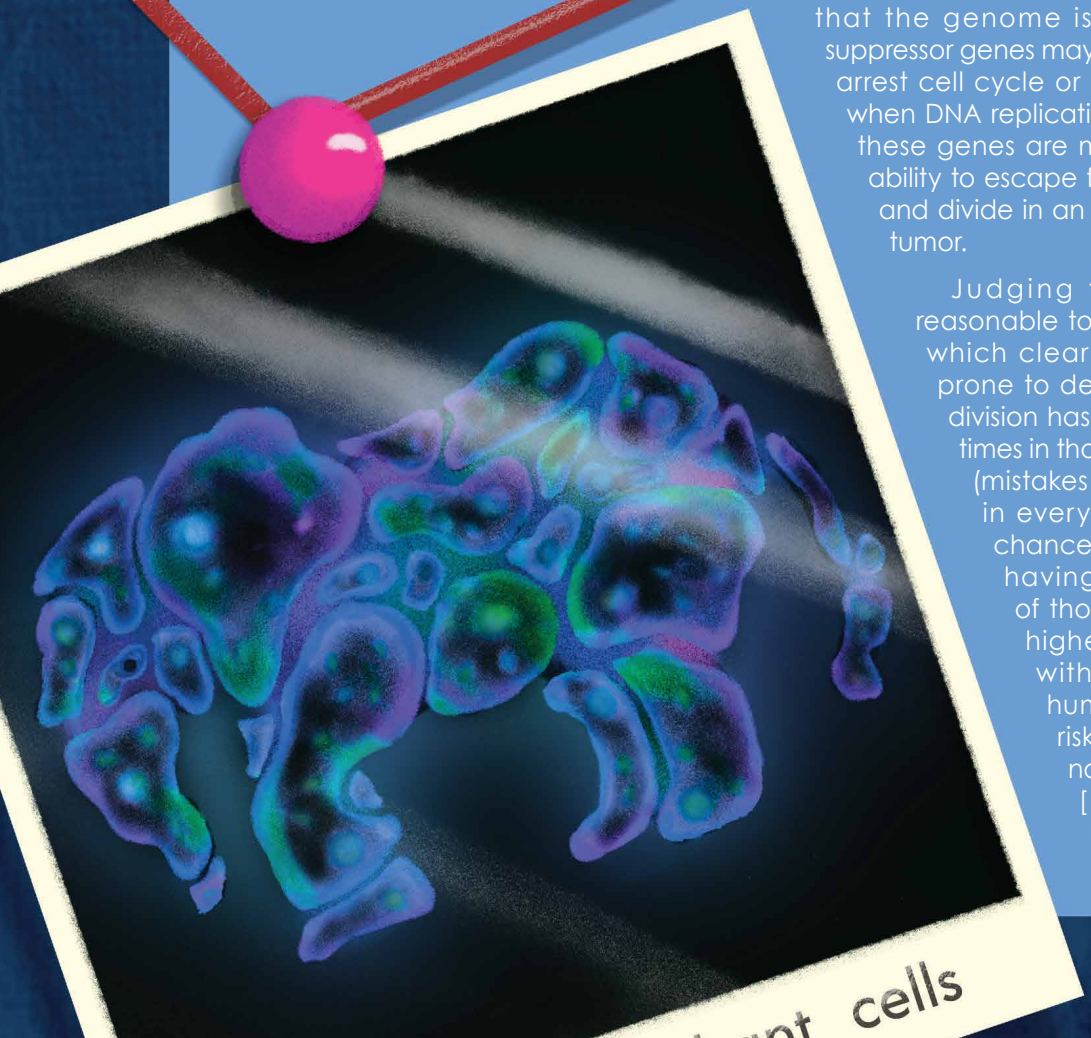
地點：香港科學館地下展覽廳

佩托悖論：體型大小是對抗癌症的關

PETO'S PARADOX : IS BODY SIZE THE KEY TO FIGHTING CANCER?

By Lambert Leung 梁卓霖

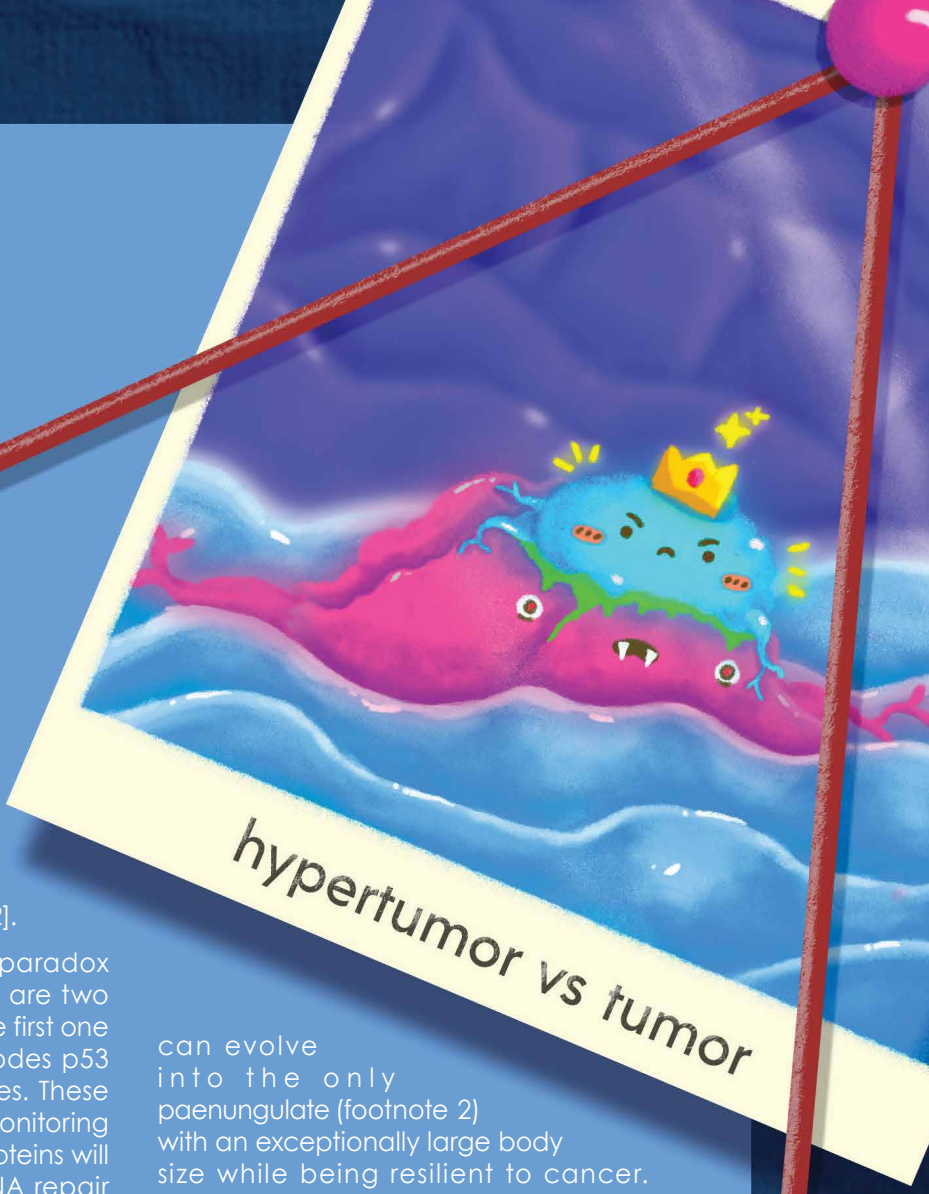
Cancer – a disease notorious for its deadliness. Till now, be it chemotherapy or immunotherapy, it still has no perfect cures. A sad truth indeed, but have you ever thought about cancers in other animals, like mice and elephants? Make a guess on the likelihood of these animals getting cancer – the answer may come as a surprise.



First, how exactly does cancer develop? In general, it can be caused by mutations in proto-oncogenes or tumor suppressor genes, both of which function to regulate normal cell growth and division. There are multiple cell cycle checkpoints to ensure that the genome is properly replicated. Tumor suppressor genes may kick in to repair damaged DNA, arrest cell cycle or induce apoptosis (cell suicide) when DNA replication goes wrong. However, when these genes are mutated, the cell may gain the ability to escape from the protective mechanisms and divide in an uncontrolled manner, forming a tumor.

Judging from this mechanism, it is reasonable to deduce that larger organisms, which clearly have more cells, are more prone to developing cancer because cell division has obviously occurred many more times in those creatures. Random mutations (mistakes in replication) may take place in every round of cell division, so the chance of larger and older individuals having cells accumulating enough of those deadly mutations should be higher. However, just as organisms with 1,000 times more cells than humans don't have an increased risk of developing cancer, we are not more cancer-prone than mice [1]. Such a lack of correlation between body size and risk of

鍵嗎？



cancer is called the Peto's paradox (footnote 1), named after the English statistician and epidemiologist Richard Peto [2].

Simply terming the phenomenon as a paradox is not enough for scientists – currently, there are two general explanations for the anomaly, with the first one being genetics. The *TP53* gene, which encodes p53 proteins, is one of the tumor suppressor genes. These proteins are located at the nuclei of cells, monitoring the DNA. When the DNA is damaged, p53 proteins will pause the cell cycle and activate other DNA repair genes if the damage is repairable, otherwise cell death will be induced to prevent further replication and perpetuation of the potentially harmful mutated DNA [3, 4]. Previous studies have demonstrated that loss-of-function mutations in *TP53* (which in turn produces p53 proteins that are not fully functional) was found in over 50% of human cancers, suggesting the importance of this mechanism in cancer suppression [5].

Subjected to higher mutation risks, elephants have evolved to contain 20 copies of the *TP53* gene, whereas humans only have one copy [6]. This means the extra copies in elephants can compensate for mutated ones, retaining the ability to kill cancerous cells in the case of mutations. In contrast, if the only *TP53* gene in humans is mutated, it can lead to an inheritable genetic condition called the Li-Fraumeni Syndrome, in which the individual is susceptible to a wide range of cancers at a young age [7]. In addition to the number of *TP53* copies, a recent study also revealed that elephants had historically restored the function of an ancient, non-functioning gene remnant called *LIF6*. In response to DNA damage, LIF6 proteins can be activated by p53 proteins to effectively induce apoptosis and kill abnormal cells before they become cancerous [8]. This could be the reason why elephant can evolve into the only paenungulate (footnote 2) with an exceptionally large body size while being resilient to cancer. Therefore, larger organisms could have unique tumor suppression mechanisms to genetically fight against cancer.

Hypertumors, the tumors of tumors, are the second explanation of Peto's paradox. Unlike normal cells, cancerous ones are competitive and not cooperative in nature. This is characterized by tumor angiogenesis, the formation of blood vessels to provide extra blood supply (with oxygen and nutrients) for proliferation [9]. Based on this fact, it is not hard to imagine that any tumor would try to capture resources by any means. Research had predicted that the competing nature of cancer cells favor the formation of parasitic hypertumors, which feed on the parent tumor's blood vessels [10]. Before a tumor can grow to a lethal size in large organisms, hypertumors would have emerged and stop the evil plan of the parent tumor by depriving their resources and keeping them at a sublethal size [10]. Nevertheless, this cannot be achieved in smaller organisms because a small tumor, which takes much less time for a single cancer cell to develop to, is already life-threatening to the host. Hypertumors simply

do not have enough time to develop. In a nutshell, there could be literally be cancers killing cancers in large organisms.

With the cruel words of "there is no treatment", the earliest description of cancer called the *Edwin Smith Papyrus* was found in Egypt in about 3000 B.C. [11]. Still being incurable at present, it is not an overstatement to say that cancer is the "king of disease". The good of studying Peto's paradox is that scientists may gain insights into how organisms cope with cancer and develop new therapeutic strategies.

1 Paradox: A contradiction that often goes against common sense.

2 Paenungulate: Members in the clade Paenungulata, which includes smaller creatures like hyrax and manatee (whose body size is still significantly smaller than that of an elephant) [8].

癌症一直以其致命性臭名昭著，即使是化療或免疫療法，今天依然沒有根治癌症的完美方法。那固然是個沉重的事實，但你又可曾想過癌症在老鼠和大象等動物身上又會是怎樣的嗎？試猜猜牠們生癌的機會，答案也許會出乎意料！

回到最根本的問題：癌症的成因是甚麼？它普遍是由原癌基因或腫瘤抑制基因突變所致，兩者都有調節細胞正常生長和分裂的功能。細胞週期裡有幾個檢查點確保整個基因組被正確地複製，當 DNA 複製過程出錯時，腫瘤抑制基因便會嘗試修復 DNA、叫停細胞週期，甚至觸發細胞凋亡（即是促使細胞自殺）。然而，如果上述兩種基因發生突變會令細胞喪失這種保護機制，使它能避過監控不受控制地分裂，形成腫瘤。

從以上機制來看，我們可以推論出體型較大的生物生癌的機會應該會較高，因為牠們體內有較多細胞，細胞分裂的次數也明顯會遠比體型較小的生物多。由於每次細胞分裂都有可能出現隨機的基因突變（即是複製上的錯誤），因此體型較大和較年老的個體擁有這些累積了足夠有害突變的細胞的機會也應該會更高。但事實上，細胞數量比人類多 1000 倍的生物並會不因此較易生癌，而我們生癌的機會也不比小鼠高 [1]。體型和生癌機會沒有關係這個現象被稱為「佩托悖論」（Peto' s paradox；註一），名稱以英國統計學和流行病學家 Richard Peto 命名 [2]。

科學家當然不會止於把現象歸類為悖論後就罷休：到目前為止，佩托悖論有兩種主流解釋，第一個與遺傳學有關。*TP53* 基因是其中一種腫瘤抑制基因，編碼著位於細胞核負責監察 DNA 的 p53 蛋白。當 DNA 受損時，如果是可以修復的損傷，p53 會煞停細胞週期並激活其他 DNA 修復基因；否則，它會觸發細胞凋亡以防細胞繼續複製已突變的 DNA，避免可能有害的 DNA 繼續存在 [3, 4]。過往研究在超過 50% 的癌症患者身上也發現 *TP53* 功能喪失型突變（loss-of-function mutations；意指會使製造出來的 p53 蛋白喪失完整功能的突變），暗示了這個保護機制對抑制癌症的重要性 [5]。

雖然體積龐大，但大象不容易有癌症的其中一個原因是因為它擁有多達 20 組 *TP53* 基因，而人類則只有一組 [6]，意味著即使大象體內的一些 *TP53* 基因發生突變，其他正常的複本也能繼續對抗癌細胞；相比之下，人類唯一的 TP53 基因發生突變會導致一種名為「李法美尼症候群」（Li-Fraumeni Syndrome）的遺傳病，令患者容易在年輕時就罹患多種癌症 [7]。除了 *TP53* 基因複本的數量外，近年研究亦發現大象曾經在演化過程中使已失去功能的遠古

基因遺骸 *LIF6* 重拾其功能：在 DNA 受損時，LIF6 蛋白能被 p53 蛋白激活而有效地觸發細胞凋亡，在異常細胞變成癌細胞前就把它殺死 [8]。這也許解釋了為甚麼大象可以進化成近蹄類（註二）裡唯一擁有龐大身軀而相對上不受高癌症風險影響的動物。由此可見，體型龐大的生物在基因上可能有著獨特的腫瘤抑制機制來抵禦癌症。

佩托悖論的第二個解釋是腫瘤上的腫瘤 — 超腫瘤（hypertumors）。癌細胞與正常細胞的不同之處在於前者喜歡爭奪一切資源，它們並不能與其他細胞和平共處。腫瘤血管新生（tumor angiogenesis）正能印證這項特徵，它是指腫瘤生長出為自己提供額外血流的新血管，藉此奪取額外的氧和營養，好讓自己快速增生 [9]。基於這個事實，我們不難想像腫瘤會用盡一切辦法霸佔周圍的資源。研究曾預測好競爭的癌細胞會有利超腫瘤的形成，超腫瘤會寄生在原本的腫瘤上，藉著吸取母腫瘤血管內的營養維生 [10]。因此腫瘤在大型生物體內生長到致命大小前，超腫瘤可能早已出現，透過耗用母腫瘤的資源使母腫瘤保持在不會致命的大小，搗破母腫瘤的邪惡計劃 [10]。然而，超腫瘤並不會出現在體型較小的生物中，因為一個細小腫瘤就已經能威脅到牠們的生命，由單一癌細胞分裂成致命腫瘤所需的時間相對較短，因此超腫瘤趕不及形成個體生命就已經危在旦夕。簡單總結一下，大型生物會出現「癌症的癌症」，從而避免了癌症發生。

最早關於癌症的描述出自約公元前 3000 年的古埃及文獻《艾德溫．史密斯紙草文稿》*(Edwin Smith Papyrus)*，當中斬釘截鐵地記述了癌症並「沒有方法醫治」這個殘酷事實 [11]，而到了現在亦依然如此，所以稱癌症為「萬病之王」一點也不為過。探討佩托悖論的好處在於科學家能從中得知大自然裡的物種是如何對抗癌症的，以此為鑑並鑽研新的治療策略。

1 悖論：看似有違常理的矛盾。

2 近蹄類：包含蹄兔和海牛等動物在內的演化支，近蹄類中很多動物的體型都比大象小（即使是海牛的體型與大象相比還是小巫見大巫）[8]。

References 參考資料：

[1] Caulin, A. F., & Maley, C. C. (2011). Peto's Paradox: Evolution's Prescription for Cancer Prevention. *Trends in Ecology & Evolution, 26*(4), 175–182. doi:10.1016/j.tree.2011.01.002

[2] Peto, R., Roe, F. J., Lee, P. N., Levy, L., & Clack, J. (1975). Cancer and ageing in mice and men. *British Journal of Cancer, 32*(4), 411–426. doi:10.1038/bjc.1975.242

[3] National Institute of Health. (2020). TP53 gene. Retrieved from *https://medlineplus.gov/genetics/gene/tp53/*

[4] Mathews, C., van Holde, K., Appling, D., & Anthony-Cahill, A. (2013). *Biochemistry* (4th ed.). Toronto: Pearson.

[5] Ozaki, T., & Nakagawara, A. (2011). Role of p53 in Cell Death and Human Cancers. *Cancers (Basel), 3*(1), 994–1013. doi:10.3390/cancers3010994

[6] Callaway, E. (2015). How elephants avoid cancer. *Nature*. Retrieved from *https://www.nature.com/articles/nature.2015.18534*

[7] American Society of Clinical Oncology. (2020). Li-Fraumeni Syndrome. Retrieved from *https://www.cancer.net/cancer-types/li-fraumeni-syndrome*

[8] Vazquez, J. M., Sulak, M., Chigurupati, S., & Lynch, V. J. (2018). A Zombie LIF Gene in Elephants Is Upregulated by TP53 to Induce Apoptosis in Response to DNA Damage. *Cell reports, 24*(7), 1765–1776. doi:10.1016/j.celrep.2018.07.042

[9] National Cancer Institute. (2018). Angiogenesis Inhibitors. Retrieved from *https://www.cancer.gov/about-cancer/treatment/types/immunotherapy/angiogenesis-inhibitors-fact-sheet*

[10] Nagy, J. D., Victor, E. M., & Cropper, J. H. (2007). Why don't all whales have cancer? A novel hypothesis resolving Peto's paradox. *Integrative and Comparative Biology, 47*(2), 317–328. doi:10.1093/icb/icm062

[11] American Cancer Society. (2018). Understanding What Cancer Is: Ancient Times to Present. Retrieved from *https://www.cancer.org/treatment/understanding-your-diagnosis/history-of-cancer/what-is-cancer.html*

Not One Less: Fencepost Errors

一個都不能遺漏：柵欄錯誤

By Peace Foo 胡適之

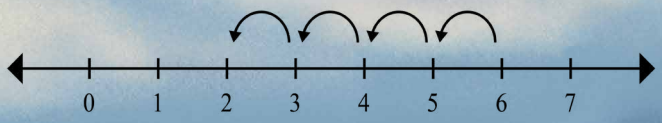
You have from 1 p.m. to 4 p.m. to work on homework assignments for subjects #1 to #4. If you complete one assignment per hour you should get them all done on time. Yes or no?

Now how about if you have from May 1 to May 4 to complete them? Can you get away with completing one assignment per day?

These kinds of problems lead to more questions [1]. Why are hours and days counted differently? Where should we start counting from anyway? Applying what we learn in school is never as simple as it seems – even with something as simple as counting.

To set things straight, let's go back to preschool. We learned to count starting from one, two, three … and we also learned that this counting process lets us know how many things there are – pencils, houses, or days. To save time we can just let the labels do the counting for us: The days of the month in June are labelled 1 to 30, so there are 30 days [1].

When we get to subtraction, the teacher holds up six pencils and takes four of them away one by one to demonstrate that 6 – 4 = 2. Subtraction is an arithmetic operation, meaning an action ("operation") is applied to change the number of objects: In this case, the act of taking a pencil away. Now we are considering a slightly different concept, the span between two numbers. This isn't the same as the numbers themselves! Once we are introduced to the number line, we get to represent the subtraction 6 – 4 = 2 as four arrows bumping down from 6 to 2:



It's clear that this refers to the four "spans" between 2 and 6. But the operation actually "touches" five numbers: 2, 3, 4, 5, 6.

The discrepancy between four and five is the fencepost error. Suppose on the way to preschool you pass by an eight-meter fence with posts every two meters [2]. How many fenceposts are there? You might do the mindless division operation and think this means there are four fenceposts in the fence. But in an ordinary fence, with fenceposts at each end, there are actually five. (Is your fence really ordinary? We'll return to this later.)

What went wrong? There are two possible things you can count in this problem: the number of segments of fence between posts, or the number of posts. When

you do the division, you start with the total length of the fence (eight meters) and divide by the length of a segment. That will give you an answer of the number of segments of fence. But the question was about the fenceposts, not the fence segments. That means that when dealing with these kinds of problems, the important question is whether you need to count the numbers or the spans [1].

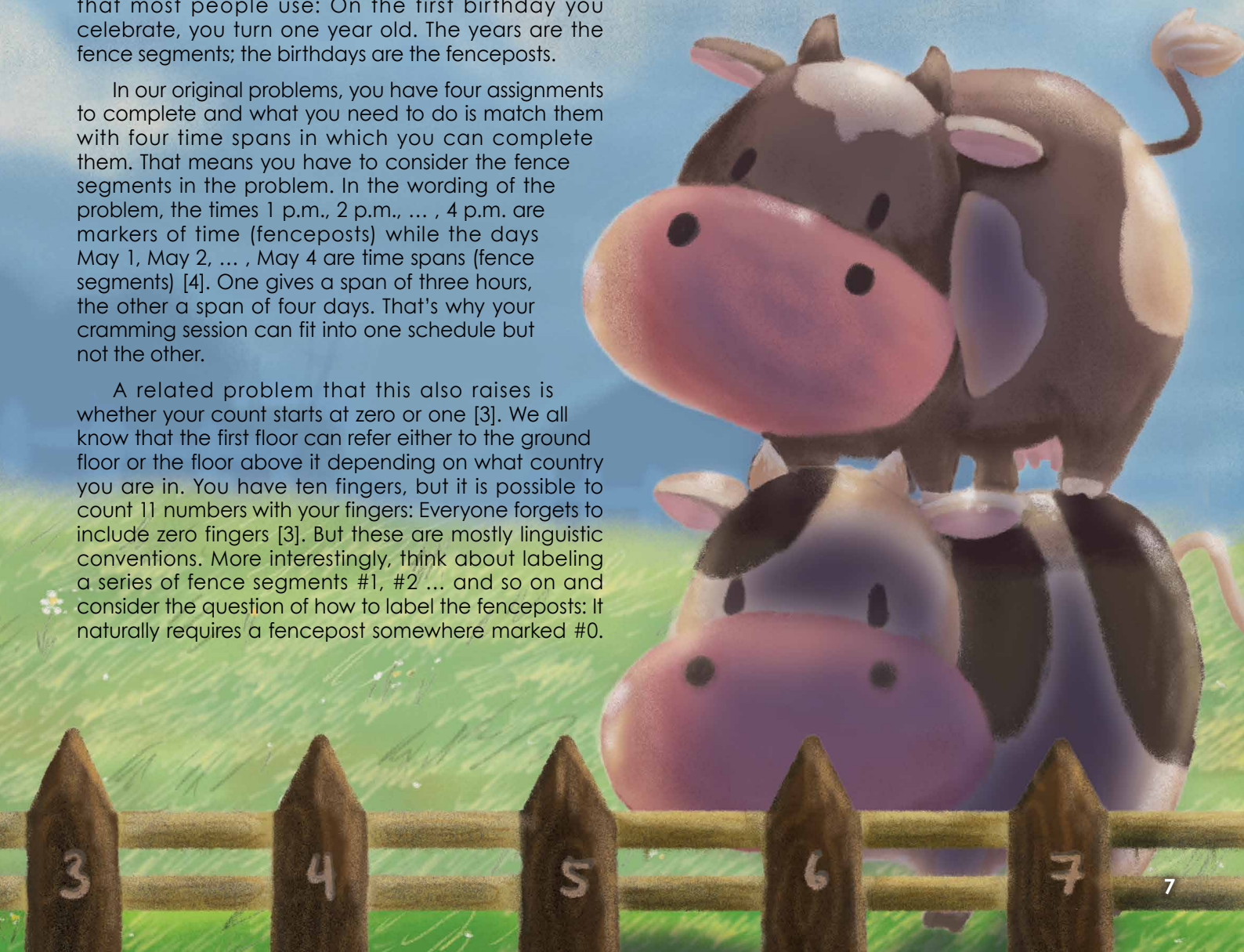
Once you can distinguish between fence segments and fenceposts, it's quite easy to see the oddities in everyday counting conventions. In music theory, a third denotes an interval of three notes: C-D-E is a third, and so is E-F-G [2]. When you put them together you get C-G, a fifth. In other words, two thirds make a fifth. This is the same as making a longer fence out of two existing fences: You'll find that there is an extra post left over since the "post" E has been counted twice. The issue is that thirds and fifths refer to the spans between notes but are named for the number of notes they contain: They count the posts instead of the segments [3].

Similarly, fence segments can be disguised as fenceposts. When you celebrate your birthday what you're actually celebrating is the number of years you were alive [2, 3]. You can even see it in the wording that most people use: On the first birthday you celebrate, you turn one year old. The years are the fence segments; the birthdays are the fenceposts.

In our original problems, you have four assignments to complete and what you need to do is match them with four time spans in which you can complete them. That means you have to consider the fence segments in the problem. In the wording of the problem, the times 1 p.m., 2 p.m., … , 4 p.m. are markers of time (fenceposts) while the days May 1, May 2, … , May 4 are time spans (fence segments) [4]. One gives a span of three hours, the other a span of four days. That's why your cramming session can fit into one schedule but not the other.

A related problem that this also raises is whether your count starts at zero or one [3]. We all know that the first floor can refer either to the ground floor or the floor above it depending on what country you are in. You have ten fingers, but it is possible to count 11 numbers with your fingers: Everyone forgets to include zero fingers [3]. But these are mostly linguistic conventions. More interestingly, think about labeling a series of fence segments #1, #2 … and so on and consider the question of how to label the fenceposts: It naturally requires a fencepost somewhere marked #0. (If you think of the fence segments as timespans spent alive, in years, and the fenceposts as birthdays then this all becomes a version of our birthday discussion above, in which the day of birth can be considered as the zeroth birthday.) This is where many of the problems of confusing fenceposts and fence segments come into play. For this reason, if you reread our original problems about the eight-meter fence, you can see that the question of whether to start counting from zero is tied to whether you want to count fence segments or fenceposts. Counting fence segments is easy. Counting fenceposts requires you to remember that extra zero – zero fingers, the zeroth floor, the zeroth birthday – and add one accordingly [1].

So once again it's very important to know the context of your question: Should you count the numbers or the spans; fenceposts or fence segments? What kind of fence are you dealing with? Just to leave you with something to think about, suppose the eight-meter preschool fence actually stretches between two walls and doesn't need posts at either end. Or maybe it closes up to form an enclosure. How many posts are needed now?



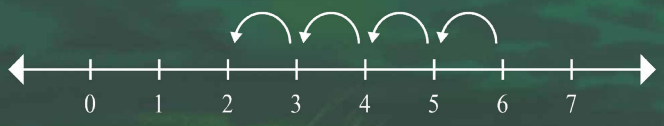
你要在下午一時至四時完成科目一至四的功課。如果你每小時完成一份就能準時完成所有功課，對或錯？

如果是五月一日至四日呢？如果每天做一份，你能及時跟功課說再見嗎？

這些問題背後引伸出更多問題 [1]：為甚麼小時跟日子的算法會有所不同？我們應怎樣數起？學校教的知識應用起來永遠不會像想像中簡單 — 顯淺如數算事物也是如此。

要搞清楚這一連串的問題，我們先要回到幼稚園。我們學會如何從一、二、三 ..... 數算物件，亦學會這個數算過程讓我們知道物件的數量 — 不論是鉛筆、房子或是日子。為了簡化事情，我們為物件標上數字並以此代替數算：六月的日子被標上數字 1 至 30，順理成章地六月就有 30 天 [1]。

然後我們學習減法。老師拿著六枝鉛筆，再把當中四枝逐枝取走，藉此示範 6 – 4 = 2。減法是算術運算的一種，指把物件數目改變的動作（「運算」），在上述例子即是取走鉛筆的動作。現在我們考慮的是稍為不同的概念：兩個數字之間的間距，這可不是數字本身！在學習數線之後，我們以由 6 到 2 之間的四個箭頭表示 6 – 4 = 2：



它們清楚表示減法考慮的是由 2 到 6 之間的四個「間距」，但這次運算其實「碰」到了五個數字：2、3、4、5、6。

四和五之間的分歧就是柵欄錯誤（fencepost error；註一）了。假設在前往幼稚園的路上你經過一道八米長的籬笆，當中每兩米豎立了一道籬杆 [2]，那麼籬杆的總數是多少？你可能不加思索便用除法計算出那裡有四道籬杆，但在首尾各豎了籬杆的正常籬笆中，籬杆的數目應該是五道。（你確定那是正常的籬笆嗎？讓我們稍後回到這個問題。）

錯誤出於哪裡？問題中有兩樣東西是你可以數算的：籬杆之間籬笆的數目，以及籬杆的數目。以除法運算時，你把籬笆的總長度（八米）除以每段籬笆的長度，那會得出籬笆有多少段這個數目，但問題問的是籬杆數目而不是籬笆，意味著解決這些問題時我們必須考慮要數算的是物件還是間距的數目 [1]。

一旦能分辨出籬笆和籬杆，就很容易發現日常生活中關於數算的習慣其實充斥著怪誕之處。在樂理中，一個三度是指相距三個音的音程：C-D-E 是一個三度，E-F-G 也是 [2]，可是把它們結合時你會得到 C-G，一個五度。換言之，兩個三度得出一個五度；這就像用兩道籬笆駁成一道更長的籬笆，數算之下會乍見多出一條「E 柱」。問題的癥結在於三度和五度理應是指兩個音之間的距離，但卻以當中有多少個音來命名，即是數算的其實是籬杆而不是籬笆 [3]。

同樣地，籬笆有時也會偽裝成籬杆。慶祝生日時你想紀念的其實是出生以後活了多少年這個事實 [2, 3]，甚至從日常用語中也能看出端倪：在你慶祝的第一個生日那時你剛好一歲。歲數是籬笆，但我們慶祝的是生日，是籬杆。

在原來的問題中，你要完成四份功課，而你要做的是把四份功課分配在四個時段內完成，所以要考慮的是問題中的籬笆。在問題的用字上，下午一時、二時.....四時等以小時計的實際上是時間的標記（籬杆），而五月一日、二日 ..... 四日的日子則是時間的間距（籬笆）[4]；前者會給出三小時的間距，後者則會給出四天的間距。因此你臨急抱佛腳的「趕功課大計」只能在後者行得通，在前者卻會碰得一臉灰。

這也帶出另一個問題，就是究竟你是由零開始數還是由一數起 [3]。我們都知道「一樓」在不同國家分別可以指地下或是地下上面的一層；你有十隻手指，但不少人會忘記如果比手勢的話你也可以由零的手勢比起，這樣便能數算 11 樣物件 [3]，但這些大多都只是語言上的習慣。如果你想想，把一列籬笆逐段標記成 #1、#2.....後到底應該如何標記當中的籬杆，那就更有趣了，因為其中一道籬杆少不免要被標記為 #0。（如果你把籬笆想成歲數，籬杆想成生日，那其實與上面生日的例子同出一轍，當中你出生那天正是你的「零歲生日」。）這就是為甚麼很多關於籬杆和籬笆的問題使人困惑的根源了。因此，如果你再看一次上面關於八米籬笆的問

題，你會發現是否應該由零數起取決於你想數的是籬笆還是籬杆。數算籬笆很容易，數算籬杆就需要你在前面加上那個額外的零 — 零的手勢、零樓（地下或「G樓」）、零歲生日 — 並在總數加上一 [1]。

因此理解問題的背景非常重要：你要數算的是物件數目還是間距，籬杆還是籬笆？此外你要處理的籬笆是哪一種？最後留下一個問題給你想想：假設幼稚園的八米籬笆兩端原來各自連著牆壁而不需要籬杆，又或是籬笆圍成一個圓圈而當中沒有缺口，那麼，現在又需要多少道籬杆呢？

1 柵欄錯誤：英文「fencepost error」中「fencepost」指的是籬桿，而中文較常見的翻譯「柵欄錯誤」卻把重點放於籬笆。

References 參考資料：

[1] Azad, K. (2009, April 28). Better Explained: Learning How to Count (Avoiding the Fencepost Problem). Retrieved from *https://betterexplained.com/articles/learning-how-to-count-avoiding-the-fencepost-problem/*

[2] Propp, J. (2017, November 16). Mathematical Enchantments: Impaled on a fencepost. Retrieved from *https://mathenchant.wordpress.com/2017/11/16/impaled-on-a-fencepost/*

[3] Parker, M. (2019). *Humble Pi: A Comedy of Maths Errors*. London, UK: Allen Lane.

[4] Lamb, E. (2016, May 10). Roots of Unity: How to Confuse a Traveling Mathematician. Retrieved from *https://blogs.scientificamerican.com/roots-of-unity/how-to-confuse-a-traveling-mathematician/*

The Number Sequence That

By Peace Foo 胡適之

Here's a list of numbers: **1, 11, 21, 1211, 111221**. Read it to yourself. Now try and guess the next number.

Here's the next number: **312211**.

And the next one: **13112221**.

Can you guess the number after that?

This puzzle was given to the very famous mathematician John Conway by one of his students. He couldn't guess it [1]. But the answer is really quite simple: The first number is **1**. When you read it to yourself, that's "one 1", or **11**. Read 11 to yourself: "two 1's", or **21**. Read 21 to yourself: "one 2, one 1", or **1211**. Read 1211 to yourself: "one 1, one 2, two 1's", or **111221**. Because it's generated by reading aloud, Conway called this an audioactive sequence [2, 3], which is also known as a look-and-say sequence.

This puzzle apparently started at the 1977 International Mathematical Olympiad [4]. When Conway heard it from a Cambridge math student who had a friend attending the competition, and after he failed to solve it, he decided to make the problem even harder. Why? This is quite standard for mathematicians – when you make a problem harder and more general, it helps you think about how to solve all possible versions of the problem.

The obvious way to make the problem harder is to start with any number you like. But one of the first things Conway noticed when starting to work on this problem was that only the digits 1, 2, and 3 "occur naturally" [1]. If you want other digits you'll need to include them in the first number. So the "interesting" digits are 1 to 3 and we should focus on them if we want to find out more about the problem.

More subtly, if you extend your example out far enough you may find something about your problem worth investigating. The next number after **13112221** is **11132 | 13211**, then **311312 | 11131221**, then **1321131112 | 3113112211**. Look at the bars we added that split each number into two parts. If we take just the first part, **11132**, and treat it as a single number, we get **311312**, then **1321131112** ... which are the first parts of the next few numbers. The same thing is true for the second part, **13211**. From this point onward the two parts in fact never interact with each other again [1], so Conway called this a "split" and the two parts its "descendants". He then started looking for numbers that can't be split in this way. Although there are infinitely many of these numbers, there are exactly 92 of them which must ultimately all appear as the descendants of every possible sequence, except **22**, which repeats itself [3]. Since Conway obviously missed studying chemistry in school, he called them "atoms" or "elements". More complex numbers like **1113213211** that can be split in this way are called "compounds". So the process of splitting compounds into elements is called "audioactive decay" [5]. Surprise!

When do these splits happen? For the string **11132 | 13211**, you can see that the first part **11132** ends with 2, so every step from then on ends with "some number of 2's" and keep ending with 2 no matter what the second part does. On the other side, **13211** begins with a 1 and will continue to begin with either 1 or 3, but not 2, so it will never mess with the first part [1].

Each audioactive element is assigned to one of the first 92 elements of the periodic table, as shown in Figure 1, from hydrogen to uranium. **11132** is hafnium, and **13211** is tin.

Element	Length	String
92 Uranium	1	**3**
91 Protactinium	2	**13**
90 Thorium	4	**1113**
...		
1 Hydrogen	2	**22**

The full table can be found on:

Figure 1 Lengths and strings of some Conway's elements [5].

The names are assigned to resemble the real physical process of radioactive decay into lighter

Needs to Be Said

唸出數列之必要：聽射性數列

elements. For example, uranium (**3**) decays into protactinium (**13**), then into thorium (**1113**), and so on. One might expect a lighter element to have a longer string under audioactive decay. However, sometimes an element also decays into a combination of "shorter" elements, which is why some elements are shorter than their predecessors [**3**] – for instance, promethium (**132**) follows samarium (**311332**). The next number after **311332** should be **13212312**, but it can split into three lighter elements, promethium (**132**), calcium (**12**) and zinc (**312**). In addition, the original starting point of **1**, **11**, **21**, **1211** until **13112221** are referred to as "primordial elements" because they can't be split either but are not involved in every possible sequence [5]. Conway focused on these 92 numbers, again, because they are "sufficiently general" mathematically; this means they can tell us something interesting about the sequence no matter which starting point we have.

It should be possible for you to guess that since all elements occur in a decay process, any possible decay process will eventually result in only these 92 elements. But as all mathematicians know, a guess is not enough. Conway proved this result over about a month of work with assistance from a fellow mathematician and called it the "Cosmological Theorem" [3]. Soon later another simpler proof was announced, but unfortunately both proofs were not published. Eventually the result was proven again by several others [6].

A consequence of this theorem is that the number of digits of successive numbers increases at a constant rate for all sequences

[3]. In our original sequence starting from **1**, these lengths are 1, 2, 2, 4, 6, 6, 8, 10, 14, 20, … which you can calculate for yourself gradually approximates toward a ratio of 1.303577… times the number of digits in the previous term [1]. Based on the proof of the Cosmological Theorem, some application of linear algebra can tell you that this ratio is a solution to some equation of degree at most 92 (i.e. the highest power of the unknown is x^{92}) [3]. Conway and his colleagues then deduced that 1.303577… is actually the largest real root of a specific 71-degree polynomial so unnecessarily complex that we're afraid to print it here [2].

Why does the answer to such a simple sequence involve monstrous decay chains, matrices and 71-degree equations? We don't know, but it shows how even the simplest questions can produce genuinely engaging mathematics if you know how to look for it. Mathematics is a sport; mathematicians love to challenge themselves. Challenge yourself enough, ask your questions in the right way, and like Aladdin's cave the door to a whole world of interesting new insights will open.

以下是一個數列：**1**, **11**, **21**, **1211**, **111221**。用中文（或英文）唸一次，然後猜猜下一個項。

下一個項是：**312211**。

再下一個是：**13112221**。

你知道再之後一個項嗎？

知名數學家 John Conway 的學生曾經叫他猜這道謎題，但 Conway 沒有成功 [1]。然而答案並不難：第一個項是 **1**，對自己唸一次：「一個一」（one 1），那就是 **11**。唸一次 11：「兩個一」（two 1's），所以是 **21**。唸一次 21：「一個二，一個一」（one 2, one 1），**1211**。唸一次 1211：「一個一，一個二，兩個一」（one 1, one 2, two 1's），**111221**。由於這個數列是透過朗讀而產生，Conway 稱之為「audioactive sequence」（像「radioactive」般具「放射性」，而「audio-」表示與聽力有關）[2, 3]，中文則對應另一個名字「look-and-say sequence」而常被譯作「外觀數列」。

這道題目據說源於 1977 年國際數學奧林匹克 [4]。當 Conway 從劍橋數學學生口中得知而一時未能解答後，他決定把問題改得更難，為甚麼？這其實是數學家解答問題的標準程序：把問題弄得更困難而更具概括性能有助思考如何一舉解決問題的所有版本。

使這道問題複雜化的方法明顯就是允許數列從任何數字開始。Conway 著手解決後其中一樣注意到的是只有數字 1、2 和 3 是「自然存在」的 [1]；如果想其他數字出現在數列中，就必須在數列的第一個項加入該數字。因此數字 1 至 3 應該內藏著某些「玄機」，如果我們要探究這個問題，就得專注於這三個數字。

事實上，事情比我們想像的還要巧妙，如果我們把數列的項都一一寫出，寫到某個位置以後你會發現一些值得研究的事情。**13112221** 的下一個項是 **11132 | 13211**，然後是 **311312 | 11131221**，再之後是 **1321131112 | 3113112211**。留意中間把字串分開的直線，如果我們僅把前面部分的 **11132** 取出並當成一個獨立的項的話，之後我們會得到 **311312**、**1321131112**……它們正是原來數列之後兩個項的前面部分。後面部分的 **13211** 也是一樣。從這個項起，字串中的前後部分再不會互相干涉 [1]，Conway 稱這個現象為「分裂」（split）。及後他著手尋找不能分裂的字串，儘管這樣的字串是無限多的，然而他發現有 92 組字串既不能分裂，但最終會全數出現在所有可能數列所產生的項當中（除了 **22**，因為之後所有項都只會是 **22**）[3]。由於 Conway 實在太掛念在學時讀的化學科，他叫這些字串「原子」或「元素」，而例如 **1113213211** 這些稍為複雜、可以分裂的字串則叫作「化合物」，化合物分裂成元素的過程就被稱為「聽射性衰變」（audioactive decay）[5]。想像力夠豐富了吧？

那甚麼時候會分裂呢？以字串 **11132 | 13211** 為例，你可以看見第一部分 **11132** 的結尾數字為 2，不管第二部分的字串是甚麼，第一部分之後的項都會以 2 結尾。另一方面，**13211** 以 1 起首，而之後的項也只能以 1 和 3 起首，絕不可能是 2，因此它永遠不會與第一部分的字串混在一起 [1]。

每種「聽射性元素」都依照真實元素表上的首 92 種元素被冠以一個名字，從氫到鈾（見表一）；**11132** 是鈴，而 **13211** 是錫。

元素	長度	字串	全表可見於：
92 鈾	1	**3**	
91 鏷	2	**13**	
90 釷	4	**1113**	
…			
1 氫	2	**22**	

表一　一些 Conway 元素的長度和字串 [5]

名字分配的方式希望反映真實放射性衰變中元素會衰變成質量較輕元素的特點，例如鈾（**3**）衰變成鏷（**13**），然後是釷（**1113**），如此類推。由此我們可能會推論較輕元素應該在聽射性衰變下擁有較長的字串，但元素有時也會衰變成由數個較短元素組成的字串組合，令較輕的元素反而比上一個元素短 [3]，例如接著釙（**311332**）的是鉅（**132**）。**311332** 的下一個項應該是 **13212312**，但它可以分裂成三種較輕的元素：鉅（**132**）、鈣（**12**）和鋅（**312**）。此外，最初 **1**、**11**、**21**、**1211** 到 **13112221** 這數個項被特別命名為「太初元素」（primordial elements），因為它們不會衰變，但也不會出現在所有可能的數列中 [5]。重申一點，Conway 只專注於這 92 組字串是因為它們在數學上「足以概括」（sufficiently general）整條問題；簡而言之，亦即是無論數列以甚麼字串開首，這 92 組字串都足以告訴我們一些關於這個數列的有趣事情。

從上文已能猜測出既然全部元素都會出現在一個衰變過程中，那任何可能的衰變過程最終都只會得出這 92 種元素。可是，所有數學家都深知「猜測」是不足夠的。Conway 在另一個數學家的幫助下用了一個月時間證明這個猜想，並稱之為「宇宙論定理」（Cosmological Theorem）[3]。不久之後出現了更簡單的證明，但不幸地兩份證明都沒有被公開發表。結果後來被其他數學家重新證明 [6]。

這個定理亦意味著所有數列裡前後兩個項的長度會以一個固定的率增加 [3]。在我們原來以 **1** 開始的數列裡，每個項字串的長度分別為：1、2、2、4、6、6、8、10、14、20……稍作計算就能發現前後項長度之比會趨近 1.303577…… [1]。參考宇宙論定理的證明，運用一些線性代數後便會得知這個比例的值是某高次方程（high-ordered equation）的解，該方程最高次項的次數可達 92 次（即最高次項可為 x^{92}）[3]。Conway 和同事之後推論出 1.303577……其實是某一條 71 次方程的最大實根，但我們深怕把這條長得可怕的方程列出會嚇怕讀者，故不在此展示了 [2]。

為何一個如此簡單的數列最終會涉及駭人的衰變鏈、矩陣和 71 次方程呢？的確很難想像事情為何會向這方向發展，但它告訴我們只要你懂得怎去尋找，即使是最簡單的問題也可以衍生出最有趣的數學。數學是一項運動，數學家喜歡挑戰自己。不時為自己尋求新挑戰，運用智慧問正確的問題，然後像阿拉丁的奇幻洞窟一樣，一扇通往奇妙新世界的知識大門就會為你而打開。

References 參考資料：

[1] Numberphile. (2014, August 8). *Look-and-Say Numbers (feat John Conway) – Numberphile* [Video file]. Retrieved from *https://www.youtube.com/watch?v=ea7lJkEhytA*

[2] Bonato, A. (2018, May 2). Audioactive Sequences. Retrieved from *https://anthonybonato.com/2018/05/02/audioactive-sequences/*

[3] Conway, J. H. (1987). The Weird and Wonderful Chemistry of Audioactive Decay. In T. M. Cover, & B. Gopinath (Eds.), *Open Problems in Communication and Computation* (pp. 173–188). New York, NY: Springer. doi:10.1007/978-1-4612-4808-8_53

[4] Mowbray, M., Pennington, R., & Welbourne, E. (1986). Prelude. *Eureka*, (46), 4. Retrieved from *https://www.archim.org.uk/eureka/archive/Eureka-46.pdf*

[5] Hilgemeier, M. (1993). Audioactive Decay. Retrieved from *http://www.se16.info/mhi/Part1.htm*

[6] Ekhad, S. B., & Zeilberger, D. (1997). Proof of Conway's Lost Cosmological Theorem. *Electronic Research Announcements of the American Mathematical Society*, *3*, 78–82.

The Father of Antiseptic Surgery: Joseph Lister

抗菌手術之父：Joseph Lister

By Dana Kim 金娥凜

Let's start with a riddle. What is the translucent, blue liquid that sits around the bathroom shelves in typical households? You can swish this around your mouth, but you cannot ingest it.

Perhaps, you know the answer by now. That's right – we are referring to a mouthwash!

To understand the inspiration behind the invention of modern mouthwash, we need to go back to history, specifically to the 19th century, to catch a glimpse of the development of antisepsis, which refers to the prevention of infections on living tissues by eliminating harmful microorganisms. The British medical scientist, Joseph Lister, was an important figure who paved the way for the development of antisepsis.

Unexplained High Surgery Mortality Rates in the 19th Century

Back in the 1860s, people were not aware that the culprit of infections was actually microorganisms. Although the idea that microorganisms can cause disease was established and promoted by Louis Pasteur's famous experiments around that time, medical professionals failed to associate the germ theory of disease with wound infections. Therefore, medical procedures remained surprisingly primitive by today's standard. For example, surgical tools were solely cleansed before being placed away for storage. Bed linens were not washed [1] and the same probe was used on the wounds of multiple patients to search for undrained pus [2]. Furthermore, pus was even thought to be a natural component of the healing process – known as "laudable pus" – but they are, in fact, a clear indication of inflammation [3].

As a result, cross infection was very common, especially in surgery, which could lead to gangrene (footnote 1) and deaths [4]. Without knowing the root cause, many surgeons have accepted wound infection to be an inevitable complication of surgery [4]. Some doctors were therefore calling for the abolishment of all surgery, advocating that "the abdomen, chest and brain will forever be closed to operations by a wise and humane surgeon [2]."

Introduction of Antiseptic Procedure by Lister

This was when Joseph Lister stepped in – inspired by a paper by Pasteur on the use of creosote to disinfect sewage, Lister selected to use carbolic acid (phenol), a newly isolated active ingredient of creosote, to kill germs on the wound and the air above, preventing them from invading the surgical wounds [2, 5].

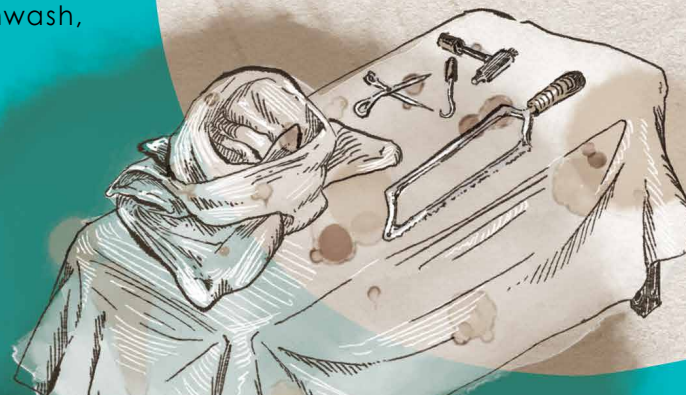
At first, many scientists were against his idea, as some were nonbelievers of the germ theory [6]. The mainstream belief was that it is the miasma released by the wound itself that causes infections, instead of pathogenic microorganisms, so doctors would wrap the wound with impermeable dressings to insulate the wound from miasma [2, 5]. In addition, some surgeons argued that the antiseptic process would hinder the progress of an operation, in which every second could be crucial to the patient's survival [4]. However, it was evident that the incorporation of antisepsis prevented major infections from operations, reducing the death rate of Lister's patients from 46% to 15% [4].

Consequently, the concept of antisepsis became accepted in other countries. Initially, doctors in Germany adopted Lister's antiseptic procedure, followed by those in the United States, France, and the United Kingdom [2, 7]. It can also be concluded that Lister allowed a broader range of operations to be performed, for instance, abdominal and other intracavity surgery [2].

Inspiration for the Invention of Modern Mouthwash

Inspired by the use of carbolic acid for antiseptic surgery, Dr. Joseph Lawrence and pharmacist Jordan Wheat Lambert set out to invent an alcohol-based surgical antiseptic. They eventually formulated an antiseptic in 1879 with no specified application [8]. The product failed to take off as a floor cleaner or a cure to dandruff, but became popular after being marketed as an oral care product to prevent plaque and gingivitis [8]. Dr. Lawrence named the modern mouthwash after Joseph Lister – Listeri...a household name you must know!

Mouthwashes are now a daily necessity sitting around our bathroom shelves, but its development involved at least three prominent scientists in history – Louis Pasteur, Joseph Lister and Joseph Lawrence. So, next time you rinse your mouth with mouthwash,



why not take a moment to thank these scientific superheroes?

1 Gangrene: The death of body parts caused by a lack of blood supply. Bacterial infection can cause tissue swelling that reduces or stops the blood supply, and eventually lead to gangrene.

先猜一道迷語：一般家庭浴室的櫃子裡都會放著透明的藍色液體，那是甚麼？你可以使它在你的嘴裡盤旋，但不可吞下。

也許你已經知道答案。對！那是漱口水。

要理解是甚麼促使漱口水的誕生，我們需要回到 19 世紀，看看當時抗菌術（antisepsis）的發展，那是指透過消除有害微生物來預防活組織受感染的方法。英國醫療科學家 Joseph Lister 是令抗菌術得以發展的重要人物。

在 19 世紀難以解釋的手術後死亡率

回到 1860 年代，人們還未意識到感染背後的元兇是微生物。儘管「微生物可以致病」的觀念已被確立，並因 Louis Pasteur 的著名實驗被廣傳，但醫護人員還是未能以細菌致病學說解釋傷口感染。因此，當時的醫療程序以今天的標準來說可謂相當「原始」，例如手術工具僅會在用後放回原來位置前才被清潔，床單正常不會被清洗 [1]，同一枝探針會被用於多個病人的傷口上找出未被清理的膿 [2]。更甚的是，膿被認為是痊癒過程中的自然產物，是樣好東西，故英文上稱之為「laudable pus」（意指「值得讚揚的膿」）；但那明顯就是發炎的跡象 [3]。

結果，尤其在手術裡，交叉感染非常普遍，導致壞疽（註一）和大量死亡個案 [4]。在對成因毫無頭緒的情況下，外科醫生已經接受了傷口感染是手術後無可避免的併發症之一 [4]，部分醫生提出應該停止一切手術，提倡「有良知的醫生應該把腹、胸、腦永遠關起而不進行手術」[2]。

Lister 引入抗菌程序

這時是 Joseph Lister 步入故事的時候：受到 Pasteur 一篇關於使用木餾油消毒污水的論文啟發，Lister 選擇採用一種當時從木餾油提取到的新活性成分石碳酸（又名苯酚；即 carbolic acid 或 phenol）來殺滅傷口及附近空氣中的病原體，避免它們入侵手術後的傷口 [2, 5]。

最初不少科學家反對他的看法，當中有不少人並不相信細菌致病學說 [6]。那時的主流觀念認為導致感染的是由傷口本身釋放的瘴氣（miasma）而不是致病的微生物，所以醫生會用不透氣的物料包紮傷口，隔開傷口和瘴氣 [2, 5]。此外，也有醫生擔心抗菌程序會阻慢手術進度，因為手術中的每分每秒都性命攸關 [4]。然而，數據表明採用抗菌程序能預防手術後的嚴重感染，把 Lister 負責病人的死亡率由 46% 減至 15% [4]。

結果抗菌的概念亦慢慢在其他國家被接納，最初德國的醫生採用 Lister 的抗菌術，然後是美國、法國和英國 [2, 7]。Lister 的發明亦令我們能在身體更多部位進行手術，譬如腹腔和其他體腔 [2]。

對發明現代漱口水的啟發

受到在抗菌手術中使用石碳酸啟發，Joseph Lawrence 博士和藥劑師 Jordan Wheat Lambert 著手研發以酒精為主要材料的手術抗菌劑，他們最終於 1879 年調製了一款沒有指定用途的抗菌劑 [8]。產品曾以地板清潔劑及去頭屑洗髮液銷售，但反應都不太理想，直至以預防牙菌膜和牙周炎的口腔護理產品作招徠後才大賣 [8]。Lawrence 博士以 Joseph Lister 的名字為產品命名，那就是家傳戶曉以「李」字開首的漱口水了！

漱口水現在是我們浴室必備的日用品之一，它的出現涉及最少三個在歷史上鼎鼎有名的科學家：Louis Pasteur、Joseph Lister 和 Joseph Lawrence。下次使用漱口水時，不妨默默感謝一下這些科學上的超級英雄。

1 壞疽（gangrene）：身體部分因缺乏血液供應而死亡。細菌感染能使組織腫脹，令血液供應減少或停止，最終導致壞疽。

References 參考資料：

[1] Ackerknecht, Erwin H., and Charles Rosenberg E. *A Short History of Medicine*. Revised and Expanded edition, Johns Hopkins University Press, 2016.

[2] Pitt, Dennis, and Aubin, Jean-Michel. "Joseph Lister: Father of Modern Surgery." *Canadian Journal of Surgery*, vol. 55, no. 5, 2012, pp. E8-E9, doi:10.1503/cjs.007112

[3] Freiberg, Jeffrey A. "The mythos of laudable pus along with an explanation for its origin." *Journal of Community Hospital Internal Medicine Perspectives*, vol. 7, no. 3, 2017, pp. 196-198, doi:10.1080/20009666.2017.1343077

[4] "Joseph Lister and the Use of Antiseptics - Attempts to Treat and Cure Illness and Disease – WJEC - GCSE History Revision - WJEC." *BBC*, *www.bbc.co.uk/bitesize/guides/zwkm97h/revision/5*. Accessed 21 Jan. 2022.

[5] Ambrose, Charles T. "The Conquest of Pus -- a History of Bitumen, Creosote and Carbolic Acid." *Journal of Infectious Diseases & Preventive Medicine*, vol. 6, no. 2, 2018, pp. 1-8, doi:10.4172/2329-8731.1000179

[6] Walker, Kenneth M. *Joseph Lister*. Hutchinson & Company, 1956.

[7] Metchnikoff, Elie. *The founders of modern medicine: Pasteur, Koch, Lister*. Books for Libraries Press, 1971.

[8] Fine, Daniel H. "Listerine: past, present and future - a test of thyme." *Journal of Dentistry*, vol. 38 Suppl 1, 2010, pp. S2-S5, doi:10.1016/S0300-5712(10)70003-8

HABER PROCESS: MORE THAN JUST NITROGEN AND AMMONIA

哈柏法：改變世界的化學反應

By Lambert Leung 梁卓霖

Have you ever doubted how the agricultural supply manages to feed the exponentially growing world population? One of the keys is to provide nitrogen, an essential nutrient for good yields, through the use of nitrogenous fertilizers. The question then becomes: Where does all that nitrogen come from?

Haber Process

The answer is the Haber process – it should ring a bell to some of you, especially those taking chemistry classes. To recap briefly, it was discovered by the German chemist Fritz Haber, who was later awarded the Nobel Prize in Chemistry in 1918. Unlike its suboptimal predecessors, the Haber process was more energy-efficient and scalable in converting nitrogen into ammonia, which could be further processed into nitrogenous fertilizers like ammonium nitrate (NH_4NO_3) and urea ($(NH_2)_2CO$) [1].

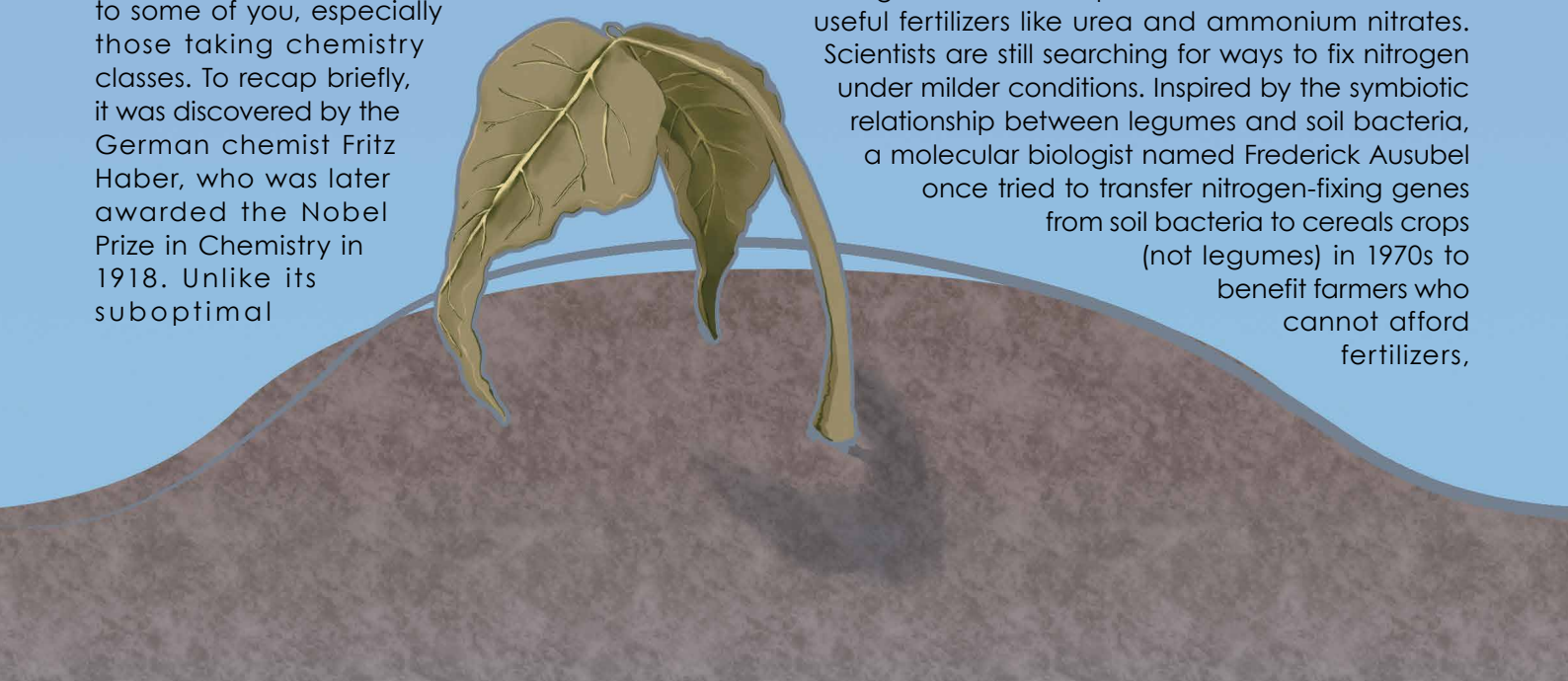
$$N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(g)$$

Such an improvement can be attributed to the Le Chatelier's principle (footnote 1), where chemists maximize yields by casting some magic on chemical equilibrium and kinetics – in industrial practice, the Haber process operates at both high temperature (around 450 °C) and high pressure (around 200 atm) [2]. The above is only a general picture of the process, if not a tip of the iceberg (most probably you have seen it in textbooks!). Let's dig deeper into the steps behind the short equation.

Nitrogen Fixation

It may be tempting to think that since about 78% of atmospheric air is made up of nitrogen, the source for fertilizers should be more than abundant. Yet, atmospheric nitrogen exists as inert diatomic molecules (N_2) held together by extremely strong N–N triple covalent bonds. This is why nitrogen won't react with hydrogen to form N–H bond under normal conditions; moreover, it also prevents plants from converting N_2 molecules to other useful forms by themselves, at least not without the help of nitrogen-fixing soil bacteria. It simply had posed a challenge for scientists. For this reason, early fertilizers were highly limited to natural sources such as manures and niter mines (for KNO_3).

To address the issue, scientists have made attempts, both chemically and biologically, in "cracking" nitrogen. In Haber process, the role of nitrogen-fixing bacteria is chemically substituted by high temperature, pressure, and an iron catalyst. This breaks the nitrogen molecules into atoms for forming ammonia, which will be converted into nitric acid through the Ostwald process as the feedstock of useful fertilizers like urea and ammonium nitrates. Scientists are still searching for ways to fix nitrogen under milder conditions. Inspired by the symbiotic relationship between legumes and soil bacteria, a molecular biologist named Frederick Ausubel once tried to transfer nitrogen-fixing genes from soil bacteria to cereals crops (not legumes) in 1970s to benefit farmers who cannot afford fertilizers,



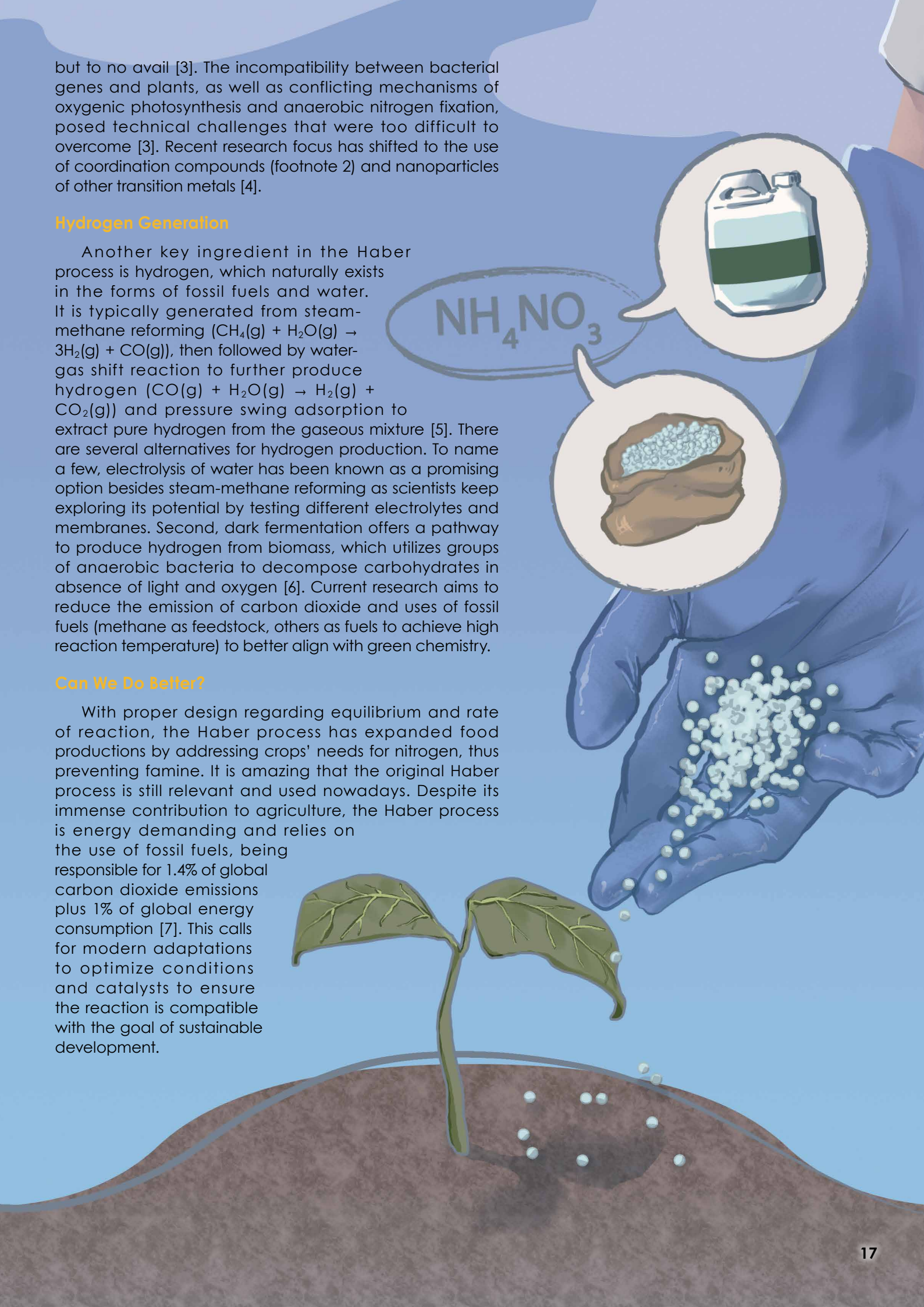
but to no avail [3]. The incompatibility between bacterial genes and plants, as well as conflicting mechanisms of oxygenic photosynthesis and anaerobic nitrogen fixation, posed technical challenges that were too difficult to overcome [3]. Recent research focus has shifted to the use of coordination compounds (footnote 2) and nanoparticles of other transition metals [4].

Hydrogen Generation

Another key ingredient in the Haber process is hydrogen, which naturally exists in the forms of fossil fuels and water. It is typically generated from steam-methane reforming ($CH_4(g) + H_2O(g) \rightarrow 3H_2(g) + CO(g)$), then followed by water-gas shift reaction to further produce hydrogen ($CO(g) + H_2O(g) \rightarrow H_2(g) + CO_2(g)$) and pressure swing adsorption to extract pure hydrogen from the gaseous mixture [5]. There are several alternatives for hydrogen production. To name a few, electrolysis of water has been known as a promising option besides steam-methane reforming as scientists keep exploring its potential by testing different electrolytes and membranes. Second, dark fermentation offers a pathway to produce hydrogen from biomass, which utilizes groups of anaerobic bacteria to decompose carbohydrates in absence of light and oxygen [6]. Current research aims to reduce the emission of carbon dioxide and uses of fossil fuels (methane as feedstock, others as fuels to achieve high reaction temperature) to better align with green chemistry.

Can We Do Better?

With proper design regarding equilibrium and rate of reaction, the Haber process has expanded food productions by addressing crops' needs for nitrogen, thus preventing famine. It is amazing that the original Haber process is still relevant and used nowadays. Despite its immense contribution to agriculture, the Haber process is energy demanding and relies on the use of fossil fuels, being responsible for 1.4% of global carbon dioxide emissions plus 1% of global energy consumption [7]. This calls for modern adaptations to optimize conditions and catalysts to ensure the reaction is compatible with the goal of sustainable development.



The Dark Side of the Haber Process

In many textbooks, the Haber process is commonly introduced as a life-saving contribution to mankind. It really would have been great if the process is used in producing fertilizers only.

Aside from being a nutrient for plants, nitrogen constitutes explosives. As aforementioned, nitrogen tends to exist as a highly stable, triple-bonded diatomic form (N_2). However, nitrogen atoms are bound by much weaker N–O or N–H bonds in nitrogen compounds compared to the strong N–N bond in nitrogen gas. When those weak bonds in the nitrogen compounds are broken by ignition, large amounts of energy are released when the triple bonds are reformed to form the more stable nitrogen gas (footnote 3), which can expand rapidly as a gas and create a shock wave. Hence, fertilizers themselves are explosives, leading to the Beirut explosion in 2020 [8].

During the World War I (1914-1918), not only was the process widely used in manufacturing explosives for the German army, but Haber himself also advocated chemical warfare: He proposed and supervised the deployment of extremely toxic chlorine gases in the Second Battle of Ypres, the first use of chemical weapons in war [9]. Shocked and unprepared, the Allies suffered severe casualties and had to retreat.

1 Le Chatelier's principle: When changes in conditions of temperature, concentrations, pressure, and volume are applied to a system, the equilibrium position would respond by counteracting those changes.

2 Coordination compound: A central metal atom that is chemically bound to, or surrounded by, other groups of non-metal atoms.

3 Editor's remark: Recall that the breaking of bonds requires energy (endothermic), while the forming of bonds releases energy (exothermic).

你可曾想過農業供應是怎樣餵飽正以幾何級數上升的世界人口嗎？氮是提高農作物產量至關重要的植物營養，因此其中一個關鍵是藉著施以氮肥為農作物提供氮。可是，這些氮又從何而來呢？

哈柏法

答案就是哈柏法，相信有修讀化學的同學對這個單詞並不陌生。簡單重溫一下，它由德國化學家 Fritz Haber 發明，亦使其在 1918 年被授予諾貝爾化學獎。與之前的同類方法相比，哈柏法能較節省能源地將氮氣轉化成氨，亦兼容大規模生產，製造出來的氨可被加工成硝酸銨（NH_4NO_3）或尿素（$(NH_2)_2CO$）等氮肥 [1]。

$$N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(g)$$

哈柏法有如此進步是因為它善用了勒沙得利爾原理（Le Chatelier' s principle）（註一）。化學家衡量過反應的化學平衡和動力學後，決定在工業上採取高溫（大約 450 ℃）和高壓（大約 200 atm）以將氨的產量最大化 [2]。不過以上只是哈柏法冰山一角的簡介（亦相信你已經在課本看過更詳盡的介紹），讓我們深入探討這條簡短化學式背後的步驟吧！

固氮作用

你可能在想反正大氣裡有 78% 都是氮氣，生產氮肥的原料應該俯拾皆是吧？然而，大氣中的氮是以惰性雙原子分子（N_2）的方式存在，原子間由強大的 N–N 三共價鍵連接，使其不會在正常情況下與氫產生化學反應並形成 N–H 鍵，這也是植物不能自行將氮轉換成其他有用形態的原因，即使要轉換亦要靠固氮細菌的幫助。這為科學家帶來一道難題，亦出於這個原因，早期的肥料大多來自糞肥和硝石礦（KNO_3）等天然來源。

有見及此，科學家嘗試在化學和生物層面把氮分子「拆開」。在哈柏法中，固氮細菌的角色由高溫、高壓和鐵催化劑等化學條件取代，當中氮分子被拆成原子用於製造氨，氨稍後會經一組名為奧斯托惠爾特法（Ostwald process）的程序被轉化成硝酸，進而用作原料製造尿素和硝酸銨等有用的肥料。直到現在，科學家仍在探索如何在較溫和的條件下固氮。分子生物學家 Frederick M. Ausubel 曾受豆科植物和土壤細菌之間的共生關係啟發，在 1970 年代嘗試把參與固氮的基因從土壤細菌轉移到穀類農作物（而不是豆科植物）來幫助買不起肥料的農夫，但研究以失敗告終 [3]。技術上的難處在於細菌和植物的相關基因並不兼容，而光合作用和固氮作用兩者亦是互相矛盾的過程，光合作用會產生氧氣但固氮作用卻是厭氧的 [3]。近年研究焦點被轉移到配位化合物（註二）和過渡金屬納米粒子的應用上 [4]。

氫氣的來源

哈柏法的另一樣原料是氫，它在大自然中以化石燃料和水的形式存在。氫氣通常由蒸汽甲烷重整（$CH_4(g) + H_2O(g) \rightarrow 3H_2(g) + CO(g)$）產生，再經水煤氣轉化反應（$CO(g) + H_2O(g) \rightarrow H_2(g) + CO_2(g)$）產生更多氫氣，最後經變壓吸附由氣體混合物中抽取純氫氣 [5]。除此之外，水電解也是一種能寄予厚望的方法，科學家正探索使用不同電解質和隔膜的可能，在未來有望充分發揮這個方法的潛能。另外，一種叫暗發酵的方法能利用厭氧細菌在無光和缺氧的環境下把碳水化合物分解，為我們帶來將生物質（biomass）轉化成氫氣的方法 [6]。為了更加符合綠色化學的原則，現時的研究致力於減少生產氫氣時所排放的二氧化碳以及所使用的化石燃料（包括作為原料的甲烷和用於產生高溫的燃料）。

有改進空間嗎？

哈柏法的設計巧妙地顧及到化學平衡和反應速率，透過滿足農作物對氮的需求來大大擴充農作物的生產規模，成功避免了糧食短缺。儘管哈柏法已有上百年的歷史，我們今時今日卻仍然依靠著它。但就如上文所述，哈柏法雖然對農業有莫大貢獻，卻有著消耗大量能源和依賴化石燃料的缺點。數據指出哈柏法排放的二氧化碳和耗用的能源分別佔了全球二氧化碳排放總量的 1.4% 和全球能量消耗的 1% [7]。

因此，科學家有需要從哈柏法的反應條件和催化劑著手去改良這個歷史悠久的化學反應，以確保哈柏法能符合現代可持續發展的概念。

1 勒沙得利爾原理：當改變系統的溫度、濃度、壓力或體積時，化學反應的平衡位置會移向相反方向以抵消改變。

2 配位化合物：由一顆中心金屬原子連接著非金屬原子，或被非金屬原子包圍而成的化合物。

3 編按：破壞化學鍵需要能量（吸熱反應），形成化學鍵則釋出能量（放熱反應）。

哈柏法的黑暗面

在很多課本以「拯救生命的發明」來介紹哈柏法，如果它僅被用於生產肥料，那的確再理想不過，可是……

氮除了是植物的營養，還是炸藥的重要成份。上文提到氮傾向以雙原子分子（N_2）的形式存在，原子被三鍵連接故非常穩定，但氮原子在氮化合物中卻是被N—O或N—H鍵等比氮氣中N—N鍵弱很多的鍵連接。當這些較弱的鍵被點燃並破壞時，氮原子便得以釋放，它們之間的三鍵就能得以重組，再次產生較穩定的氮氣，過程會釋出大量能量（註三），使作為氣體的氮氣迅速膨脹並產生衝擊波。因此，肥料本身就是炸藥，過往這個危險的特質曾經釀成2020年的貝魯特（Beirut）爆炸事故 [8]。

哈柏法不但在第一次世界大戰（1914-1918）中被德軍廣泛用於生產炸藥，發明家Haber本人也是化學武器的提倡者：他提出並監督了在第二次伊珀爾戰役（Second Battle of Ypres）中使用氯氣這種劇毒的策略，亦是第一次在戰爭中使用化學武器的實例 [9]。這殺盟軍一個措手不及，使其嚴重受挫，被逼撤退。

References 參考資料：

[1] Barona, A., Etxebarria, B., Aleksanyan, A., Gallastegui, G., Rojo, N., & Diaz-Tena, E. (2018). A Unique Historical Case to Understand the Present Sustainable Development. *Science and Engineering Ethics, 24*(1), 261-274.

[2] BBC. (n.d.). Ammonia and sulfuric acid. Retrieved from *https://www.bbc.co.uk/bitesize/guides/z234tyc/revision/2*

[3] Ausubel, F. M. (2018). Tracing My Roots: How I Became a Plant Biologist. *Annual Review of Genetics, 52*, 1-20. doi:10.1146/annurev-genet-120417-031722

[4] Morlanés, N., Almaksoud, W., Rai, R. K., Ould-Chikh, S., Ali, M. M., Vidjayacoumar, B., . . . Basset, J. (2020). Development of catalysts for ammonia synthesis based on metal phthalocyanine materials. *Catalysis Science & Technology, 10*(3). doi:10.1039/C9CY02326G

[5] U.S. Department of Energy. (n.d.). Hydrogen Production: Natural Gas Reforming. Retrieved from *https://www.energy.gov/eere/fuelcells/hydrogen-production-natural-gas-reforming*

[6] Ghavam, S., Vahdati, M., Grant Wilson, I. A., & Styring, P. (2021). Sustainable Ammonia Production Processes. *Frontier in Energy Research, 9*. doi:10.3389/fenrg.2021.580808

[7] Capdevila-Cortada, M. (2019). Electrifying the Haber–Bosch. *Nature Catalysis, 2*(1055). doi:10.1038/s41929-019-0414-4

[8] Guglielmi, G. (2020, August 10). Why Beirut's ammonium nitrate blast was so devastating. *Nature*. doi:10.1038/d41586-020-02361-x

[9] King, G. (2012, June 6). Fritz Haber's Experiments in Life and Death. *Smithsonian Magazine*. Retrieved from *https://www.smithsonianmag.com/history/fritz-habers-experiments-in-life-and-death-114161301/*

OPTOGENETICS: 光遺傳學：運用光控制腦部活動 CONTROLLING BRAIN ACTIVITIES WITH LIGHT

By Sirius Lee 李揚

Have you read the article *The Amazing Cephalopods* from our previous issue? We introduced to you how large and thick axons from marine organisms like squids have contributed to our understanding of nerve impulse. This time, we are excited to present to you another useful technique in the study of neuroscience. In earlier days of research, scientists struggled to control a single type of cell in the brain without altering other variables in the vicinity [1]. Previous techniques which involve electrodes to stimulate neurons and record signals were not ideal because they failed to target specific cell types and electrical recording could be affected by simultaneous stimulation at the same site [1]. Administration of drugs to control neurons is possible but drugs act slowly [1] and may lead to unwanted side effects [2]. In recent years, scientists have invented a novel technology which can overcome the problems above. They were inspired by the discovery of bacteriorhodopsin found in archaea and its sensitivity towards light [3]. Voila! Welcome to the era of optogenetics.

Breaking down the word, "opto" indicates "optical", while "genetics" implies the modification of the genetic material. With a specific set of neurons in mind, their genome could be rewritten by editing techniques such that the cells would express the newly inserted gene encoding the light-sensitive bacteriorhodopsin. By exposing the genetically modified organisms to the light of specific wavelengths, scientists could therefore switch on or off neurons, similar to a video gamer when they use their remote control to command the avatar.

The haloarchaeal bacteriorhodopsin was the first to be discovered. This membrane protein pumps protons (H^+) from the cytoplasm to the extracellular fluid when activated by photons. Since its initial discovery in 1971, researchers also discovered two other classes of light-responsive proteins, halorhodopsin in 1977 and channelrhodopsin in 2002 [1]. Yellow light-activated halorhodopsin is a chloride pump which actively pumps negative chloride ions (Cl^-) into the cell upon excitation by yellow light; and blue light-activated channelrhodopsin is a cation channel which allows positive ions to flow into the cell under concentration gradient upon excitation by blue light [1, 4]. It is known that a neuron fires when it is depolarized, such as when positive ions rush into the cell. Therefore, by expressing suitable light-sensitive proteins and exciting them with suitable colors of light, one could excite or inhibit a neuronal cell by manipulating the ion flow [4]. With the key to instructing the networks of genetically modified neurons, causal links between neuronal activities and molecular or behavioral outcomes can be established.

Optogenetics has quickly become a gold standard for neuroscience research. For instance, neuroscientists could first use a pharmacological agent (drug), tetrodotoxin, to inhibit neuronal activity in hippocampal slices *in vitro* (footnote 1), and then repeat the experiment through light-directed inhibition on transgenic mice expressing halorhodopsin *in vivo* to confirm the neuronal activity-dependent expression of a key protein [5]. In terms of clinical use, researchers have explored the therapeutic use of optogenetics, expanding its functionality. Promising results recently published in *Nature Medicine* support such an approach in restoring partial vision to a patient suffering from a neurodegenerative eye disease, retinitis pigmentosa [6]. In the study, viral vector containing channelrhodopsin gene was injected into a patient's eye to genetically engineer the retinal cells in fovea. Meanwhile, researchers designed a pair of goggles which could detect the light intensity of the surroundings, and convert the information to light signal for stimulating the channelrhodopsin expressed in the retinal cells. With the aid of the goggles, retinal cells were appropriately activated and partial vision of the patient could be restored.

It could not be more amazing when nature's terrific designs can be harnessed and transformed into powerful tools in our pursuit of science. This wisdom of using optogenetics has brought us one step closer to illuminating the neural circuitry of our brains, or perhaps, given us the key to untangling all mysteries in neuroscience.

1 *In vivo* & *in vitro*: *In vivo* means "within the living" literally in Latin. It often refers to experiments conducted in or on a living organism, as opposed to *in vitro*, meaning "in glass (labware)".

你讀過我們上一期的文章《地球上的「外星智慧生物」：頭足類動物》了嗎？我們向你介紹過魷魚那些海洋生物大而厚的軸突如何幫助我們了解神經脈衝。今次，我們會向大家介紹一種對神經科學研究極為有用的技術。在早期研究中，科學家面對的困難是無法在不影響附近環境的其他變數下控制單一類型的腦細胞 [1]。以前所用的技術涉及以電極刺激神經元及記錄信號，因為無法針對特定的細胞類型，而且記錄的電脈衝可能會受到在同一部位同時施以的刺激影響，因此效果並不理想 [1]。利用藥物來操控神經元雖然可行，但藥物發揮作用需時 [1]，更可能會導致不必要的副作用 [2]。科學家近年發明了一項可以克服上述問題的新技術，他們的靈感來自發現於古細菌的細菌視紫紅質（bacteriorhodopsin）及其感光性 [3]。太好了！— 歡迎來到光遺傳學的年代。

光遺傳學，顧名思義，是一門涉及使用光和改寫基因編碼的技術。選定了一種特定的神經元作研究對象後，我們可以透過基因編輯技術改寫它們的基因組，使細胞表達新插入的感光蛋白細菌視紫紅質的基因。之後透過把生物暴露在相應波長的光下，科學家就可以控制神經元開關，情況類似於遊戲玩家使用遙控器控制遊戲角色。

最早發現的是來自古細菌的細菌視紫紅質，這種細胞膜蛋白被光子激活時會將質子（H^+）從細胞質泵出細胞外液。自 1971 年首次發現細菌視紫紅質以來，研究人員還發現了另外兩類感光蛋白，分別是在 1977 年發現的嗜鹽視紫紅質（halorhodopsin）和 2002 年的光敏感通道蛋白（channelrhodopsin）[1]。對黃光敏感的嗜鹽視紫紅質是一種氯離子泵，在黃光激發下能主動轉運負極性的氯離子（Cl^-）進入細胞；對藍光敏感的光敏感通道蛋白是一種正離子通道，在藍光激發下允許正離子順濃度梯度流入細胞 [1, 4]。神經元會在去極化時被激發而發放神經脈衝，例如在正離子湧入細胞的時候，因此透過表達合適的感光蛋白並用適當顏色的光把蛋白激活，就可以操縱離子進出細胞以刺激或抑制神經元 [4]。掌握著這把操控基因改造神經元網絡的鑰匙，我們就可以設計實驗驗證神經元活動與某分子活動或個體行為間的因果關係。

光遺傳學很快就成為了神經科學研究上的黃金標準，例如神經科學家可以一方面用河豚毒素這種藥物在活體外（註一）抑制海馬體切片中的神經元活動，另一方面在表達嗜鹽視紫紅質的轉基因小鼠上用光在活體內抑制相關神經元活動來重覆實驗，以驗證某重要蛋白是否會因神經元活動而被表達 [5]。研究人員亦開始嘗試尋找光遺傳學在臨床治療上的用途，擴展這項技術的用處。最近發表在《自然醫學》（*Nature Medicine*）的有力證據支持光遺傳學在治療視網膜色素病變（retinitis pigmentosa）上的應用，採用光遺傳學的方法能使罹患這個神經退化疾病的患者恢復部分視力 [6]。在這項研究中，含有光敏感通道蛋白基因的病毒載體被注射到患者其中一隻眼睛以基因改造黃點上的視網膜細胞。與此同時，研究人員設計了一副眼鏡，它可以探測周圍環境的光強度，並將資料轉化成光訊號以刺激視網膜細胞中表達的光敏感通道蛋白。在這副眼鏡的幫助下，視網膜細胞就能被適當地激活，令患者得以恢復部分視力。

在探究科學的路上，沒有事情比能夠把大自然的美妙設計轉化成我們的強大工具更為奇妙。這種運用光遺傳學的智慧點亮了神經科學研究的道路，使我們對完全理解大腦神經迴路邁進了一步；或者，這給了我們一把釐清神經科學上許多謎團的鑰匙。

1 在活體內（外）：在中文上也許從字面就能看出含義；而在英文上「*in vivo*」一詞在拉丁文中解作「在活的生物內」，通常指在活生生的生物上進行的實驗，其對應的相反詞為「*in vitro*」，解作「在玻璃（實驗器具）內」。

References 參考資料：

[1] Deisseroth K. Optogenetics. *Nat Methods*. 2011;8(1):26-29. doi:10.1038/nmeth.f.324

[2] Hartsough LA, Park M, Kotlajich MV, et al. Optogenetic control of gut bacterial metabolism to promote longevity. *Elife*. 2020;9:e56849. doi:10.7554/eLife.56849

[3] Oesterhelt D, Stoeckenius W. Rhodopsin-like protein from the purple membrane of Halobacterium halobium. *Nat New Biol*. 1971;233(39):149-152. doi:10.1038/newbio233149a0

[4] Yizhar O, Fenno LE, Davidson TJ, Mogri M, Deisseroth K. Optogenetics in neural systems. *Neuron*. 2011;71(1):9-34. doi:10.1016/j.neuron.2011.06.004

[5] Wang Y, Fu WY, Cheung K, et al. Astrocyte-secreted IL-33 mediates homeostatic synaptic plasticity in the adult hippocampus. *Proc Natl Acad Sci U S A*. 2021;118(1):e2020810118. doi:10.1073/pnas.2020810118

[6] Sahel JA, Boulanger-Scemama E, Pagot C, et al. Partial recovery of visual function in a blind patient after optogenetic therapy. *Nat Med*. 2021;27(7):1223-1229. doi:10.1038/s41591-021-01351-4

WILL MATHEMATICIANS BE REPLACED? – Computers in Mathematics

數學家會被取代嗎？— 電腦在數學中的角色

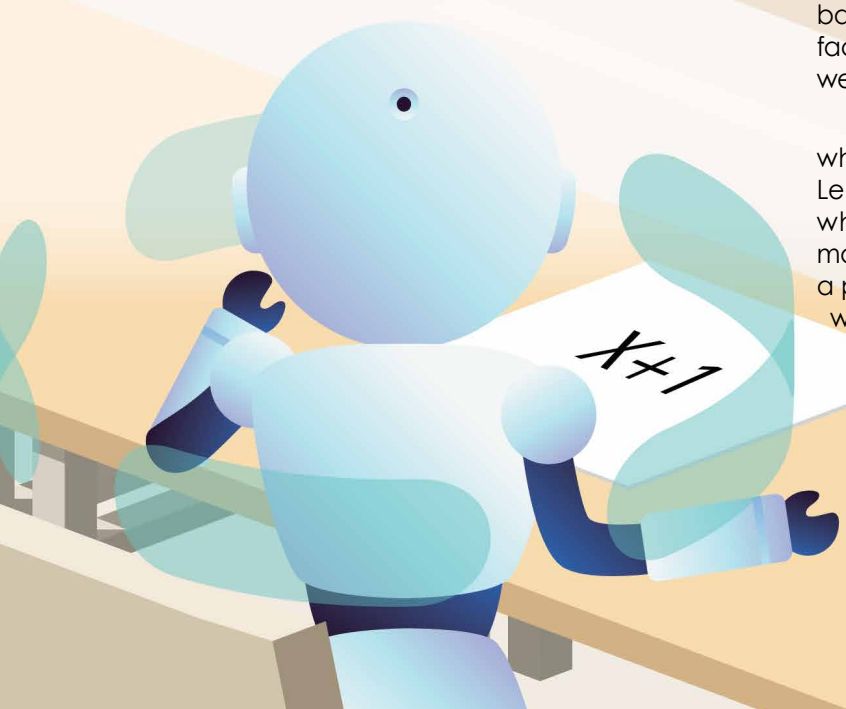
By Sonia Choy 蔡蒨珩

Could a computer do math?

Well, you may think, of course it can – you only need to type "1 + 1" into the Google search bar for it to show the number two. Computers, in fact, were built to do math; the name comes from the root "compute", and was formerly used to refer to actual people who do calculations by hand before we had computers. Indeed, by using Microsoft Excel or programming languages such as Python or C++, a computer can do plenty of difficult calculations far faster than a human being could. But mathematicians are far more interested in another type of math – proof writing.

Consider the following problem: Start with any number. If it is even, divide the number by two. If it is odd, multiply it by three, then add one. Does this sequence always get to 1 eventually? This innocent-looking problem is known as the Collatz conjecture, and remains unsolved to this day. A computer could check whether this is true for an arbitrarily large number (given enough time and computing power), but to prove that it goes to 1 for any number takes far more effort and new techniques; we do not know how to do it yet.

Consider another problem that we can solve. Given the quadrilateral in Figure 1, how do you find the highlighted angle? Well, we apply the theorem "exterior angle of cyclic quadrilateral" to prove that the two are the same. Could a computer do the same thing? If we teach it the theorem, then yes. Can computers, then, prove things that we do not yet know?



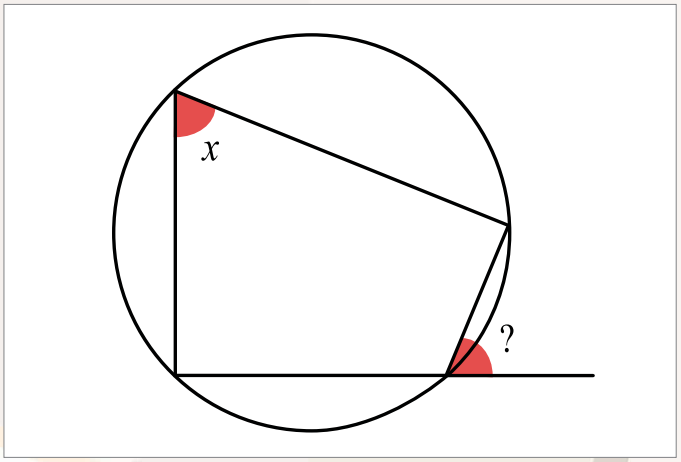


Figure 1 Exterior angle of a cyclic quadrilateral.

Well, maybe. We need to teach a computer how to do mathematics.

The first task is to formalize the mathematics we know now. One recent contender for this is Lean, a software developed by Microsoft engineer Leonardo de Moura in 2013 [1]. Big swathes of basic university-level mathematics have been formalized (rewritten in a way a computer understands) since its foundation, and there are multiple ongoing efforts to build up a bigger mathematical library for Lean that it can draw on.

To get a taste of how this is done, there is a game available online called *The Natural Number Game* [2]. There you can try teaching a computer to count and do basic addition and multiplication – even the most basic facts, like "the natural number after x is $x + 1$," are things we need to tell the computer.

Being able to count is, of course, a long way from what current research in mathematics looks like. But Lean can be fed more and more complex theorems which allow it to do far more: At the end of 2020, mathematician Peter Scholze had doubts about parts of a proof he wrote. Scholze, a Fields Medalist (footnote 1), works in arithmetic geometry, a field known for being extremely technical and abstract; he and Dustin Clausen had proved a fundamental statement to the field a few years ago, but there were parts of it he was unsure about – and who better to check it than a machine? Hence began the Liquid Tensor Experiment, an attempt to verify their proof using Lean [3].

Six months later, the results were positive – their proof was correct, albeit with a few slight imprecisions. A group of mathematicians, with the vast majority of work done by Johan Commelin, were able to follow their written out proof and "teach" it to a computer, hence verifying their proof [4].

So it is actually possible that a computer could actively participate in research-level mathematics – with a slight catch. Here Lean is a proof assistant, not a proof generator; the computer's checking of the proof relies on the human effort of formalizing the proof, since it cannot read and digest human prose by itself. It also certainly cannot come up with a proof of this very technical statement all on its own; while Lean can make simplifications that might not come as completely obvious to humans, it still follows the skeleton of the argument given by Scholze and Clausen and merely checks whether it is logically correct.

There is also another dimension to this problem; research in mathematics is largely about proving statements, but part of it is also about choosing and proving the "right" statements. Often when mathematicians cannot prove a certain statement, they go about proving a simpler version that is still "interesting". These words, however, are much more difficult to define than absolute concepts like true and false; for example, while we might not yet have enough machinery to prove the Twin Prime Conjecture (i.e. "there are infinitely many pairs of primes that have a difference of two"), we are able to prove a weaker version of this statement, that there are infinitely many pairs of primes with a difference less than or equal to 246 [5], and notably this weaker version is still of interest to us.

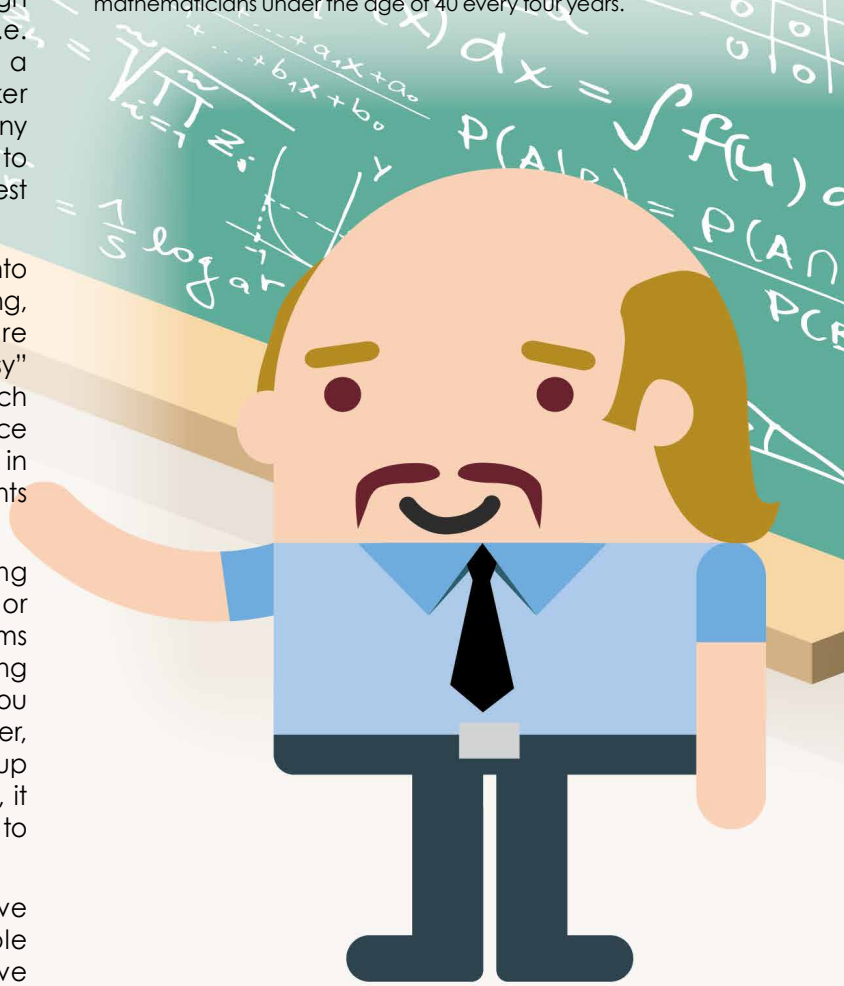
To look for ways to simplify a given conjecture into the right statement is straightforward to a human being, but not necessarily easy to a computer. There are ongoing efforts to study the difference between "easy" and "hard" problems by using machine learning, such as a recent effort began by Timothy Gowers [6]; once this barrier is cleared, we might have more progress in getting computers to prove "not-too-hard" statements that are "useful" to mathematicians.

We are, also, still a long way away from solving problems that need some sort of "cheat" or "hint", or involve constructing certain examples; these problems are often extremely difficult by brute force searching but are almost trivial once you know what method you need. Without a human feeding the hint to the prover, at present it is impossible for a computer to come up with the "hint" itself, and even for a mathematician, it takes years of training to be able to spot the best way to attack a problem.

But recently, AIs have also been able to solve harder problems independently – it is now possible for certain engines (ROBOT [6], Lean [7]) to solve "routine" undergraduate-level math problems, as well as problems from the International Mathematical Olympiad, a worldwide competition for high schoolers often called the Olympics of Mathematics. So it is only natural to expect that computers will be able to solve harder and harder problems in the future – perhaps, at some point, even surpassing human capabilities.

So what role will computers play in research-level mathematics in the future? The honest answer is that nobody knows; AI-optimists will claim that computers will eventually replace mathematicians, in the time frame of ten years to a century; those less optimistic will claim that mathematicians are irreplaceable. As an aspiring research mathematician, I surely hope that we will not be replaced in the near future (very unlikely). But it is also exciting to see progress both in improving the knowledge base which software can draw from, and in computers' ability to solve problems hands-on. Perhaps one day they will replace mathematicians – who knows? But first there will surely be a time when computers will aid mathematicians greatly in their research, when they gain the ability to prove tedious statements that nobody wanted to do.

1 Fields Medal: Regarded as the "Nobel Prize in mathematics" (which does not exist), it is one of the most prestigious prizes in mathematics, which is awarded to two to four outstanding mathematicians under the age of 40 every four years.

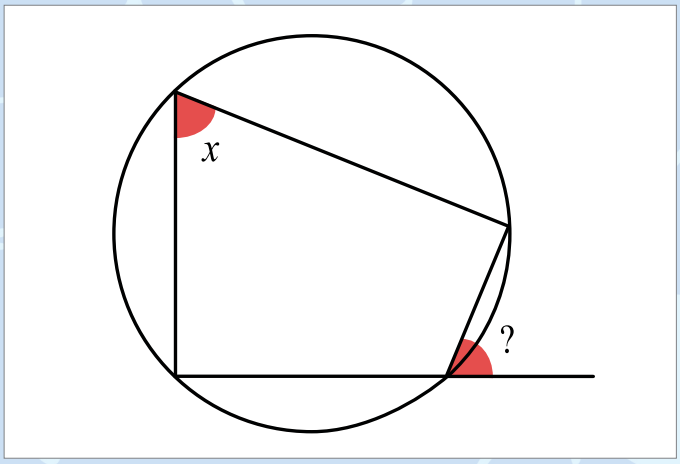


電腦能從事數學工作嗎？

聽到這個問題的時候，你的第一個反應必定是「可以」吧，畢竟我們只需在 Google 搜尋「1 + 1」，它就會自動顯示答案「2」。事實上最初發明電腦的目的真的是為了計算；英文名稱「computer」（計算機）來自字根「compute」（計算），在沒有電腦的年代指從事計算工作的計算員。透過微軟 Excel 或 Python 及 C++ 等程式語言，電腦的確可以迅速地進行複雜的計算工作，所需時間遠比人類短，可是數學家感興趣的卻是另一種「數學」：寫數學證明。

考慮以下的問題：任意選一個數字作起點，如果這個數字是雙數，把它除以二；如果是單數，先將數字乘以三，再加上一。重覆以上步驟，到底數列最終是不是永遠都會得出 1？這個看似簡單的問題就是考拉茲猜想（Collatz Conjecture），但現在還沒有人能夠解答。電腦可以檢查一個任意大的數字最後會得出甚麼（假設有足夠的時間和計算能力給電腦去計算），但這與要證明任何一個數字最後都會得出 1 還有一段距離，要解決這個問題肯定需要一些新方法，但擁有電腦的我們卻一籌莫展。

那有甚麼問題是我們可以用電腦解決的呢？例如這道：試找出圖一四邊形中所示的角度。我們都懂得用「圓內接四邊形外角」這條定理去證明兩者相等。那麼，電腦又懂得這樣做嗎？如果我們教曉電腦用這條定理的話，那沒問題；然而，電腦又能不能解開人類還沒解開的數學問題呢？



圖一　圓內接四邊形外角

嗯，也許可以吧。但我們要先給電腦上一課數學課。

第一個任務是把我們現在知道的數學知識形式化（formalize）。近年這方面有一個在 2013 年由微軟工程師 Leonardo de Moura 開發，叫 Lean 的軟件 [1]。自 Lean 問世而來很多大學程度的基本數學知識都已經被形式化（以電腦能理解的形式改寫），現時多方仍在努力豐富 Lean 的數學知識庫，令其可以引用更多數學知識和定理。

網上有一個叫《*The Natural Number*》的遊戲讓玩家可以體驗把數學知識形式化的過程 [2]。在遊戲中你可以嘗試教電腦數算、基本加法和乘法，但在這之前你需要告訴電腦例如「x 的下一個自然數是 $x + 1$」等最基本的知識。

當然，現時數學家研究的課題遠比數算難，但我們可以對 Lean 輸入更多更複雜的定理讓其發揮更多。2020 年末，數學家 Peter Scholze 對自己寫的一部分證明有所懷疑；他是菲爾茲獎得主（註一），主要研究以極其專門和抽象見稱的算術幾何。事源他和 Dustin Clausen 幾年前證明了一條對該範疇非常重要的定理，但他對當中一些細節並不確定。如果要檢查這份證明，還有比電腦更適合的「人選」嗎？於是他們展開了液體張量實驗（Liquid Tensor Experiment），嘗試用 Lean 來驗證證明 [3]。

六個月後，Scholze 和 Clausen 得到了正面的結果：儘管他們的證明有幾個小漏洞，但大致正確。實驗中，以 Johan Commelin 貢獻最多的一組數學家在理解 Scholze 和 Clausen 寫的證明後，以電腦明白的語言「教」電腦讀懂證明，使其能驗證證明 [4]。

因此電腦的確有能力參與數學研究並作出貢獻，但有一個細節必須注意：上文提到的 Lean 只能輔助我們檢查證明，但它並不能憑空寫出證明。由於電腦不能讀懂用我們文字寫的證明，因此即使要驗證也要靠我們把證明形式化。電腦亦肯定不能自行創造出這類極為複雜的證明，儘管 Lean 能洞察到證明中可以作出簡化的地方，並在這些對我們來說未必太明顯的位置精簡證明，但總的來說它還是僅僅順著 Scholze 和 Clausen 所寫的論證來檢查證明是否邏輯上正確。

這延伸出另一方面的問題：雖然數學研究大多都是關於證明敘述，但當中亦涉及如何找「正確」的敘述來證明。當數學家不能證明某個敘述時，他們往往會將敘述簡化，先證明比較簡單但仍然「有用」的版本。然而判別何謂「正確」和「有用」並不像分辨對錯那樣非黑即白，這需要用到更高階的思維，譬如我們現在並不能證明孿生質數猜想（其敘述為「前後兩者相差為 2 的質數對有無限多對」），但是我們卻能退而求其次，證明前後兩者相差少於或等於 246 的質數對有無限多對 [5]，這個安慰獎對我們來說還是有用的。

把猜想簡化成仍然有用的敘述對人類來說並不難，但對電腦來說可能一點也不容易。數學家正嘗試用機器學習（machine learning）研究一個問題到底何謂「簡單」和「困難」，例如 Timothy Gowers 最近就正著手處理這個問題 [6]。一旦有所突破，也許我們就能讓電腦自行找出一些「不太困難」但仍對數學家有用的敘述來證明。

對於要用上一些「捷徑」或「提示」，或是要建構特別例子才能解決的數學問題，電腦還是鞭長莫及。這些問題通常很難透過蠻力搜尋（brute force searching）以窮舉的方式解決，但當你知道所需的方法後一切卻會迎刃而解。如果沒有人對電腦輸入這個「提示」，現今的電腦是不能自己洞察到的；儘管是數學家，這種直覺也是讀書和實戰多年後的成果。

可是近年人工智能也開始能獨自解決一些較難的問題：一些程式（ROBOT [6] 和 Lean [7] 等）都能解開大學本科課程中比較簡單的數學問題，以及高中數學比賽國際數學奧林匹克（International Mathematical Olympiad）的試題，因此不難想像電腦在未來將可以解決越來越難的問題 — 也許有一天，它的解題能力甚至會超越人類。

那麼電腦未來在數學研究上會擔任什麼角色呢？說真的，沒有人知道。對人工智能前景樂觀的人會認為電腦數十年或一個世紀內將超越並取代數學家；沒那麼樂觀的人會認為數學家是不可取替的。筆者作為一個希望成為數學家的學生，當然希望我們在短期內不會被電腦取代（其實也不太可能會被取代），但也樂見這些輔助軟件的知識庫變得越來越豐富，電腦的解題能力又不斷在進步。也許電腦真的有一天會取代人類，誰知道？但在這天來臨之前，電腦或許會先掌握書寫數學證明的能力，屆時沒有人想花時間解決的繁複證明工作將得以解決，數學界也必定會經歷一段黃金時期，那時電腦將會成為數學家的得力助手。

1 菲爾茲獎：被喻為「諾貝爾數學獎」（這個獎項並不存在），它是數學界其中一個最崇高的獎項，每四年頒發一次予二至四位年齡為 40 歲以下的傑出數學家。

References 參考資料：

[1] Hartnett, K. (2020, September 21). At the Math Olympiad, Computers Prepare to Go for the Gold. *Quanta Magazine*. Retrieved from *https://www.quantamagazine.org/at-the-international-mathematical-olympiad-artificial-intelligence-prepares-to-go-for-the-gold-20200921/*

[2] Buzzard, K., & Pedramfar, M. (n.d.). Natural Number Game. Retrieved from *https://www.ma.imperial.ac.uk/~buzzard/xena/natural_number_game/*

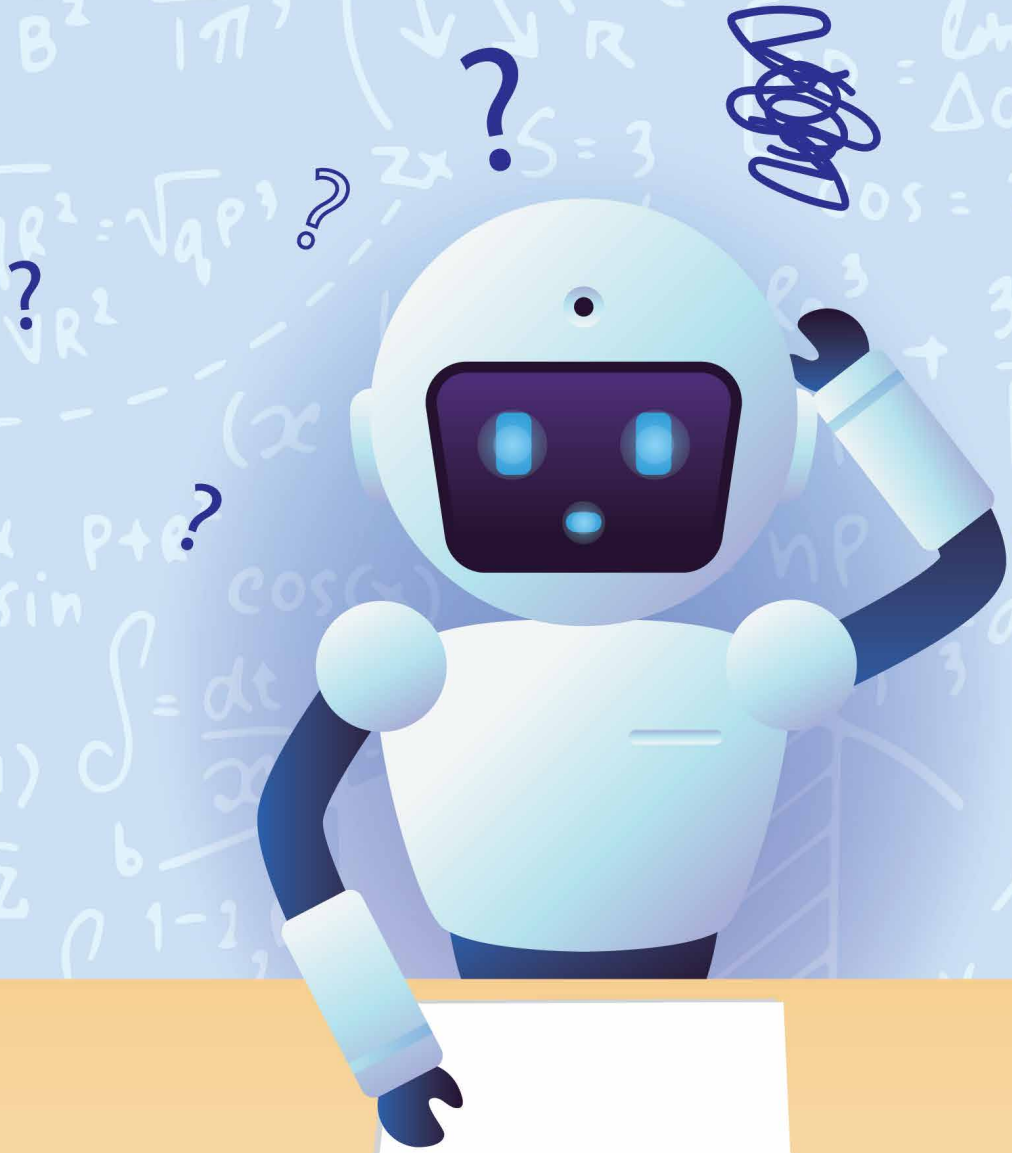
[3] Scholze, P. (2020, December 5). Liquid tensor experiment [Web log post]. Retrieved from *https://xenaproject.wordpress.com/2020/12/05/liquid-tensor-experiment/*

[4] Scholze, P. (2021, June 5). Half a year of the liquid tensor experiment: Amazing Developments [Web log post]. Retrieved from *https://xenaproject.wordpress.com/2021/06/05/half-a-year-of-the-liquid-tensor-experiment-amazing-developments/*

[5] Polymath, D. H. J. (2014). Variants of the Selberg sieve, and bounded intervals containing many primes. *Research in the Mathematical Sciences, 1*. doi:10.1186/s40687-014-0012-7

[6] Gowers, W. T. (2022, April 28). Announcing an automatic theorem proving project [Web log post]. Retrieved from *https://gowers.wordpress.com/2022/04/28/announcing-an-automatic-theorem-proving-project/*

[7] Polu, S., Han, M. J., & Sutskever, I. (2022, February 2). Solving (Some) Formal Math Olympiad Problems. Retrieved from *https://openai.com/blog/formal-math/*



Let's Be Social!

Follow us on Instagram (@sciencefocus.hkust) and get the latest updates of Science Focus:

SCIENCEFOCUS.HKUST

Fun Facts

Memes

Bite-Size articles

Instagram Photo Competition

Acknowledgements 特別致謝

Print Advising 印刷諮詢
HKUST Media Technology and Publishing Center
香港科技大學媒體科技及出版中心

Like us on Facebook

FOLLOW US ON Instagram
@sciencefocus.hkust