Science Focus 科言

Issue 025, 2023

School of Science 理學院

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Contents

Science Focus Issue 025, 2023

Message from the Editor-in-Chief 主編的話

Dear Readers,

Welcome to a new issue of *Science Focus*. In the age of artificial intelligence, it was tempting to write this letter with one of the pervasive online tools, especially after reading the article on ChatGPT in this issue. Nevertheless, there is much to be learned about how to ask the right questions in order to prompt ChatGPT to do exactly what we want. In addition, it is important to remember that we are still the main drivers in knowledge generation, based on novel observations and discoveries.

In this post pandemic era, it has become common knowledge on how vaccines triggered the production of antibodies to combat deadly viral invaders. It turns out the structure of antibodies from alpacas and their relatives is rather different from ours. Read more about how scientists have leveraged on alpaca antibodies for research and therapeutics. Staying with advances in biomedical science, you can learn more about how medical devices can be removed by infiltration with a liquid metal, and how our fingerprints are determined by a combination of genetic and environmental factors. Many efforts have been made in reducing the use of plastics, but how can we get rid of plastic waste more efficiently? Some scientists found unexpected inspirations from insects and bacteria. Finally, for those of you who play the game Wordle, we provide a quick guide on how to maximize your chance of getting the elusive green squares.

As we enter a new school year, I hope *Science Focus* continues to be the bridge between recent scientific advances and what you learn in textbooks. As always, we welcome your suggestions and comments via our social media pages.

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

親愛的讀者：

歡迎閱讀最新一期《科言》！在這個人工智能盛行的年代，相關應用程式有如雨後春筍般湧現，在讀過今期關於 ChatGPT 的文章後，讓我以此作為今期〈主編的話〉的引子。對於如何問 ChatGPT 正確的問題才能使其做出我們所想，仍有待慢慢學習，但此外更重要的是，我們應該記住人類仍是知識的主要發現者，我們能透過細心觀察作出不一樣的發現。

在後疫情年代，相信大家都熟識疫苗如何激發身體產生抗體以抵禦致命病毒，但原來羊駝和近親朋友們製造出來的抗體結構與我們的都不太一樣，閱讀後頁文章看看科學家如何把羊駝抗體應用於研究和醫藥上。同樣是生物醫學上的突破，今期我們會介紹液態金屬如何能透過滲入醫療裝置而把其移除，還會探討遺傳和環境因素如何共同決定指紋樣式。社會一直致力減少使用塑膠，但我們如何能更有效地消除塑膠廢料呢？科學家從昆蟲和細菌中找到意想不到的靈感。最後，我們為有玩 Wordle 的讀者們送上一份簡易攻略，看看如何增加綠色框框出現的機會吧！

在新學年裡，我希望《科言》繼續成為一道橋樑，連結科學界最新發現和課本上的知識。一如以往，我們歡迎大家透過社交媒體把建議和意見告訴我們。

主編 麥晧怡教授
敬上

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！

InnoCarnival 2023

Under this year's theme of "Go Smart! Go Tech! Go Green!," this event organized by the Innovation and Technology Commission aims at promoting innovation and technology culture in the community. The public can participate in a wide range of activities, including workshops, interactive games, and online seminars, with the goal of boosting creativity. Thirty-seven program partners, including universities, research and development centers, government departments, and organizations, will showcase their research achievements at the Science Park.

Period: October 28, 2023 (Sat) – November 5, 2023 (Sun)
Time: 10:00 AM – 6:00 PM (Sat & Sun)
10:00 AM – 5:00 PM (Mon to Fri)
Venue: Hong Kong Science Park
Admission fee: Free of charge

Remark: Quotas are limited for the pre-registered activities on a first-come, first served basis.

創新科技嘉年華 2023

這個由創新科技署舉辦的嘉年華今年以「智慧生活綠色科技」為題，旨在推廣社區內的創科文化。公眾可以從一系列的工作坊、互動遊戲，以及網上講座體驗創新科技的樂趣，激發創意。三十七個活動夥伴，包括大學、研發中心、政府部門和多個機構都會在科學園展示其研究成果。

展期： 2023 年 10 月 28 日（六）至 2023 年 11 月 5 日（日）
時間： 上午十時至下午六時（六、日）
上午十時至下午五時（一至五）
地點： 香港科學園
入場費： 免費

備註：需預約的活動名額有限，先到先得。

The Shaw Prize 2023 Exhibition

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 this year's Shaw Laureates and their scientific achievements.

Time: November 10, 2023 – January 10, 2024
Venue: Main Lobby, Hong Kong Science Museum
Admission fee: Free of charge

2023邵逸夫獎展覽

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。

時間： 2023 年 11 月 10 日至 2024 年 1 月 10 日
地點： 香港科學館大堂
入場費： 免費

Mars 1001

This sci-fi movie depicts the journey of astronauts from different countries as they embark on the first manned mission to Mars. It highlights the challenges involved in the 1001-day mission to uncover the mysteries of the Red Planet. While it has been over half a century since the first astronaut Neil Armstrong set foot on the Moon, landing on Mars is a much more complex goal that requires advanced technology and poses greater risks. This movie portrays the exhilaration of finally achieving this long-held dream of visiting other planets in person, and the international cooperation that would be required to turn fiction into reality.

Period: Now – January 31, 2024
Time: 3:30 PM and 8:00 PM (Mon, Wed to Fri)
2:00 PM and 6:30 PM (Sat, Sun 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.

火星千日行

這套科幻電影敘述來自不同國家的太空人攜手完成人類首次登陸火星的旅程，為了揭示這顆赤紅行星的神秘面紗，團隊在任務的 1001 天裡遇上重重挑戰。雖然太空人岩士唐早已在超過半個世紀前成功登月，但登陸火星是一個需要更先進科技和更具風險的艱鉅任務。在希望這會在現實世界成真的同時，這套電影預示了人類達成到訪其他行星，以及實現國際間緊密合作夙願的喜悅。

展期： 現在至 2024 年 1 月 31 日
時間： 下午三時半及八時（一、三至五）
下午二時及六時半（六、日及公眾假期）
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 Procrastinator's Guide to CHATGPT

遲來的CHATGPT懶人包

By Sonia Choy 蔡蒨珩

True story: One of my friends wrote his thesis with the help of Clyde, Discord's AI server bot powered by OpenAI, the same company that invented ChatGPT (Chat Generative Pre-Trained Transformer). He had difficulties writing parts of the paper, asked Clyde to rewrite his clunky section, and boom, it was done. My friends and I also often took to ChatGPT when drawing graphs and diagrams in an unfamiliar computer language. ChatGPT would churn out 90% correct code in ten seconds, saving us huge amounts of time as we would only need to make slight modifications. ChatGPT has truly transformed our lives and the world of education.

But ChatGPT isn't foolproof yet. That same friend once asked an early version of ChatGPT a simple question: What is 20 – 16? After a few seconds, it gave us the answer "3." We laughed about it for a few minutes. People have also posted responses of ChatGPT to various questions that look legit, but turns out to be a pile of nonsense. ChatGPT can write complicated code, but it can't seem to do simple things like subtraction and figuring out that the sun rises in the east. Why is that the case?

Machine Learning 101

First we need to answer the question – how does ChatGPT learn things? Artificial intelligences (AI) are typically modeled on the human brain's neural networks [1, 2]. A neural network is typically divided into three main layers – the input, hidden and output layers. The input and output layers have obvious meanings, but the hidden layer is the key of the model; there can be multiple hidden layers. There are also nodes at each level, which are linked to other layers, and sometimes to others in the same layer (Figure 1).

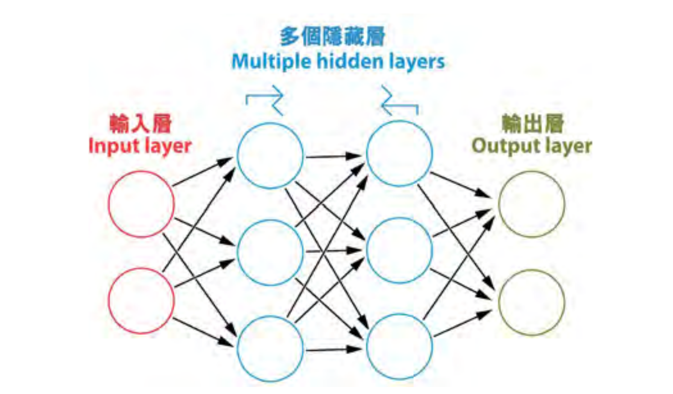


Figure 1 The main layers of a neural net with circles as nodes.
圖一 神經網絡的主要分層。圓圈代表節點。

Each layer of neurons evaluates a numerical function, and its outputs influence other neurons they are connected to. These functions act as the thinking process and reach its goal by evaluating certain criteria. For example, if the goal for the AI is to identify pictures of cats, then each layer will evaluate some sort of similarity to existing pictures of cats. By learning bit by bit from the examples, it knows what outputs are desired in each layer, and adjusts itself so that it is finally able to identify pictures of cats.

AI models are typically trained either by deep learning or machine learning. While these terms are sometimes used interchangeably, they have a slight difference – in deep learning, the AI is programmed to learn unfiltered, unstructured information on its own, while in machine learning more human input is required for the model to learn and absorb information, e.g. telling the AI what it is learning, as well as other fine-tuning of the model.

According to OpenAI, ChatGPT is a machine (reinforcement) learning model [3]. It uses human-supervised fine-tuning, and does not adjust independently in the process of learning new material, perhaps due to the complicated nature of human language. While the details of how the model was trained and its mechanisms are kept under wraps, perhaps in fear that other companies may make use of them and exceed GPT's capabilities, OpenAI only revealed that GPT-3 was trained on filtered web crawl (footnote 1), English-language Wikipedia, and three secret sets of written and online published texts which they referred to as WebText2, Books1 and Books2 [4]. It is speculated that the undisclosed components include online book collections like LibGen, as well as internet forums and other informal sources.

Generating the Probabilities (Large Language Model)

But if you have experience with auto-correct predictions on your phone, you might have some idea of the chaos that might ensue. The current autocorrect chain on my phone, starting with the word "I", goes like this: "*I have to go to the university of the Pacific ocean and I will be there in about to go to bed now.*" Sounds legit at first, but it descends quickly into gibberish (there is no University of the Pacific Ocean, for example). That is because auto-correct only picks up on patterns in language without comprehending the actual meanings – it won't know that "*Colorless green ideas sleep furiously*" is complete nonsense (Footnote 2) [5].

ChatGPT is more intelligent than autocorrect. First of all, it generates a list of probabilities of possible next words. Let's use a simpler GPT-2 system for demonstration – after the clause "The best thing about AI is its ability to…," the GPT would generate the list of words in Table 2 [1].

學習 (learn)	4.5%
預測 (predict)	3.5%
製作 (make)	3.2%
理解 (understand)	3.1%
做 (do)	2.9%

Table 2 A list of probabilities of possible next words generated by GPT-2 after the clause "The best thing about AI is its ability to…" [1]

表二 由 GPT-2 預測「人工智能最棒的地方在於它能夠……」之後下一個可能詞語出現的概率列表 [1]。

How are these probabilities generated? First of all, it is not possible to just infer them from existing writing, as we don't have nearly enough text that is accessible under copyright for models to train on. Instead, we use a little bit of mathematics to help us out.

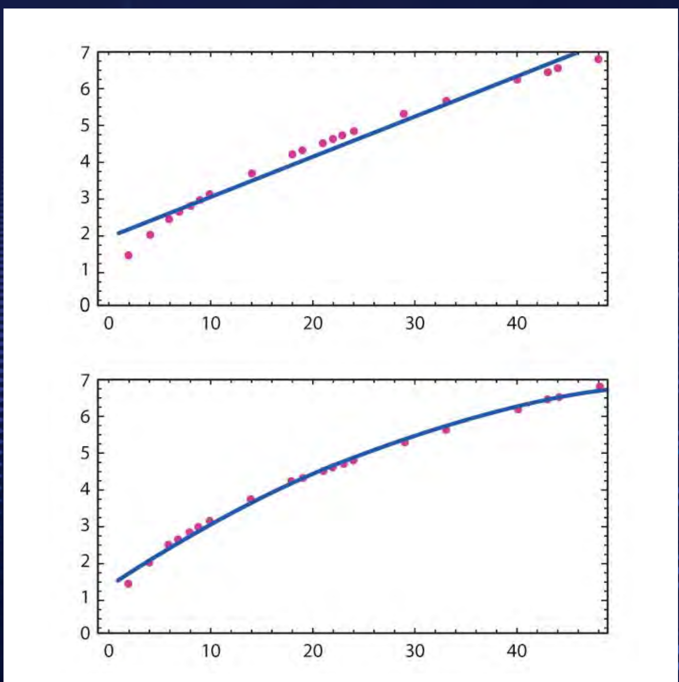


Figure 3 Fitting an equation of straight line (top) and a quadratic equation (bottom) into a series of given points [1].

圖三 嘗試用直線方程(上)及二次方程(下)解釋橙色點的分佈 [1]。

GPT is part of the large language model (LLM) family. The main working principle behind LLM sounds familiar to mathematicians: approximation (or more accurately, mathematical modeling). Given the following series of points in Figure 3 [1], how would you plot a graph? The easiest option seems to be a straight line, but we can do better with a quadratic equation, $ax^2 + bx + c$.

So we can say $ax^2 + bx + c$ is a good model of these points, and start making predictions with the model.

Just as mentioned before, the amount of text we have is far from adequate for us to empirically calculate a probability of the occurrence of the next word, because the 40,000 common English words can already give 1.6 billion (40,000P2) combinations [1]. The model, GPT, works essentially because an informed guess can be made to choose a good enough "line" to fit in the graph, covering the missing parts with theoretical values. To this date we don't really have a good understanding of how the computer does it, just as we don't really know how our brain does "intuitive" tasks, but the developers of GPT can adjust the "weight" of the output of each neuron in the network to achieve the optimal results after each learning process. In other words, we train the neural net to find the "best-fit curve" via machine learning.

Eventually, the goal is for GPT to predict what comes after a phrase in an

unfinished sentence, so they can acquire the ability to write independently.

Creativity in AI

"GPT, surprisingly, writes like a human. It can generate text that reads as if it was written by a person, with a similar style, syntax, and vocabulary. The way it does this is by learning from a huge amount of text, such as books, articles, and websites, which helps it understand how language is used in different contexts..."

The previous paragraph was written by Sage, a chatbot powered by GPT-3.5. It reads just like human writing – you might not have noticed it was written by an AI if I didn't tell you. How does it do that? Well, as GPT describes itself, it is trained on a vast amount of text, with which it builds an LLM and evaluates what the most statistically likely words are after writing each phrase.

You might think that GPT will always pick the most likely word on each occasion, but this is not the case. Creativity is found in the unexpected. If you choose a higher "creativity index" (technically called "temperature"), GPT will pick from other less likely options to continue its writing. This makes the overall piece more interesting and less robotic.

For example, if GPT picks the statistically most likely word every time (zero temperature), we would get the following paragraph in an older GPT-2 system [1]:

"The best thing about AI is its ability to learn from experience. It's not just a matter of learning from experience, it's learning from the world around you. The AI is a very good example of this. It's a very good example of how to use AI to improve your life. It's a very good example of how to use AI to improve your life. The AI is a very good example of how to use AI to improve your life. It's a very good example of..."

It falls into a loop eventually. Even if this doesn't happen in GPT-3, the paragraph itself isn't that interesting. However, if we increase the temperature to 0.8 in GPT-3, we get this [1]:

"The best thing about AI is its ability to learn and develop over time, allowing it to continually improve its performance and be more efficient at tasks. AI can also be used to automate mundane tasks, allowing humans to focus on more important tasks. AI can also be used to make decisions and provide insights that would otherwise be impossible for humans to find out."

Now this reads more like human writing. The temperature 0.8 is arbitrary but seems to work best at the moment (although it also depends on whether your task requires creativity). A lot of the machine learning process is not that well understood by humans, just as the true processes of the human brain remain a mystery. How do humans learn their native language? What do the hidden layers in our brains do in order to produce human-like text? We don't know the answers to either of these questions yet.

Falsehoods, Biases and Accountability

One problem with GPT is that it sometimes comes up with blatantly false statements and has inherent biases towards certain social groups. We've witnessed how it can confidently announce 20 – 16 = 3. It has claimed, in a previous version of GPT-3, that coughs stop heart attacks, that the U.S. government caused 9/11, and even made up references that don't exist [6, 7]. Why did this happen? Once again, GPT is only a LLM, meaning that it knows how language works, but doesn't necessarily understand its meaning. Early LLMs even have only syntactic knowledge and very few comprehension skills.

However, this is about to change. At the time of writing, GPT had recently announced a partnership with WolframAlpha [8], a mathematical software and database, and other online databases to let it access more accurate information, so that it can draw on the databases to improve its accuracy rather than giving responses generated entirely by probability.

In some sense, training GPT or any model is like teaching a toddler; they come into the world not knowing what is correct and wrong, and it is up to their parents, teachers, and society to teach them what are right. Here the programmers are the parents of GPT, as they input tons of learning materials into the system, and supervise its learning by providing reference answers and feedback.

It is possible to tell GPT enough information to force it to say

truths when you ask factual questions. However, when it comes to opinions, there is a human-imposed block. If you ask ChatGPT how it feels about large birds, for example, it replies with an automatic message: "*As an AI language model, I don't have personal opinions or feelings. However, I can provide you with some information about large birds.*"

But we could theoretically ask GPT to write an opinion piece, and we can predict how it would do by studying the correlated words it comes up with on certain topics. Researchers analyzed the top ten descriptive words that occurred concurrently with words related to gender and religion in the raw outputs generated by GPT-3; they observed that "naughty" or "sucked" are correlated with female pronouns, and Islam is commonly placed near "terrorism" while atheism is placed near "cool" and "mad" [4]. Why does GPT hold such biases, then? Remember that GPT is trained on a selected sample of text – most of it comes from published texts and web crawls, but in order for it to grasp informal language, GPT was also speculated to be trained on internet forums such as Reddit. As such, it may end up internalizing biases held by many users of these forums. Just as a person may hold prejudiced views, GPT cannot be expected to be completely neutral on all topics.

GPT-4 is already far more capable at certain jobs than humans; however, it cannot be trusted to be a completely neutral source, nor can it be trusted to give 100% accurate information. It must still be used with discretion. The best approach is probably to treat it like a person – take everything with a grain of salt.

1 Web crawl: A snapshot of the content of millions of web pages, captured regularly by web crawlers. The downloaded content can serve as a dataset for web indexing by search engines and AI training.

2 Editor's notes: This is a famous example suggested by linguist Noam Chomsky to illustrate that a sentence can be grammatically well-formed but semantically nonsensical.

一個真實故事：筆者有位朋友在 Discord 的聊天機器人 Clyde 幫助之下完成了自己的畢業論文。他在寫某些部分時遇到阻滯，好像怎麼寫也是不太通順，於是他請 Clyde 重寫笨拙的部分，結果在一下子就完成了。Clyde 是 Discord 的人工智能（artificial intelligences / AI）伺服器機器人，由發明 ChatGPT（Chat Generative Pre-Trained Transformer；聊天生成預訓練轉換器）的公司 OpenAI 提供技術支援。ChatGPT 確實改變了我們生活，以至教育界的景象：朋友和我在使用不熟悉的電腦語言繪製圖表時也經常使用 ChatGPT，因為它在十秒內就能編寫出 90% 正確的程式碼，我們只需稍作修改即可，為我們節省了大量時間。

但我們操作 ChatGPT 時還是不能把腦袋扔掉。那位朋友曾經問過舊版 ChatGPT 一條簡單問題：20 – 16 是多少？數秒後，ChatGPT 回答：「3」，這使我們捧腹大笑了好幾分鐘。網民尤其喜歡抓出 ChatGPT 的痛腳，在網上分享它種種似是而非的荒謬論述；到底為甚麼它能寫出複雜的電腦程式，但似乎不能回答簡單的減法題目或者指出太陽從東邊升起的事實呢？

機器學習 101

首先我們要解答一條問題：ChatGPT 怎樣學習知識？人工智能多數都是模擬人腦的神經網絡 [1, 2]。神經網絡主要可以分為三層：輸入層、隱藏層和輸出層。輸入層及輸出層的意思不用多說，但隱藏層才是精髓所在，而一個網絡可以包含多個隱藏層。此外，以上每一層都均有節點（nodes）連接不同層或是同一類別的其他層（圖一）。

每層神經元都會計算一個函數，輸出值將影響相連的神經元。這些函數正如思考過程，會透過考慮一系列相關因素來達成目標，譬如說：如果 AI 的任務是辨認貓的照片，那麼每一層就會比對相片與現有貓照片在某個方面的相似度。透過一步步從現有例子中學習，AI 會知道每一層應該要做到怎樣的輸出而作出自我調整，使它最終能辨認貓的照片。

AI 模型通常透過深度學習或機器學習進行訓練，雖然許多人會交替使用這兩個詞語，但其實它們有著細微分別：在深度學習中，AI 的程式設定使它自行學習未經過濾、缺乏結構的資訊；而在機器學習中，模型需要更多人類指示來學習和吸收資訊，例如告訴 AI 它正在學習甚麼，以及對模型作出其他微調。

根據 OpenAI 的說法，ChatGPT 是一種機器（或強化）學習模型 [3]。可能是基於人類語言的複雜性，ChatGPT 在人類監督下才會作出微調，而不會在學習新材料的過程中自我調整。也許是擔心其他公司會製造出超越 GPT 能力的模型，OpenAI 對訓練方法和原理的細節三緘其口，只透露 GPT-3 在訓練過程中使用了經過過濾的網絡抓取（web crawl）（註一）、英語版維基百科，以及三組他們稱之為「網路文本二」（WebText2）、「書籍一」（Books1）和「書籍二」（Books2）的線上文庫 [4]。據推測，這些未公開的部

分包括 LibGen 等線上圖書館，以及互聯網論壇和其他非正式來源。

概率之學：大型語言模型

如果你有在手機使用自動修正、預測字詞等功能的經驗，你應該會對隨之而來的混亂有所了解。筆者此時手機以「我(I)」開首的自動修正字串是這樣的：「我得去太平洋大學，而我會到那裡馬上就要睡覺了。」(*"I have to go to the university of the Pacific ocean and I will be there in about to go to bed now."*)句子乍聽之下尚算正常，但很快就會發現那只是胡言亂語(例如世界上根本就沒有太平洋大學)，因為自動修正功能只懂得選擇語言中常見的組合，但不能理解其實際含義 — 它不會知道「沒有顏色的綠色想法激烈地睡覺」(*"Colorless green ideas sleep furiously"*；註二)是完全沒有意義的廢話 [5]。

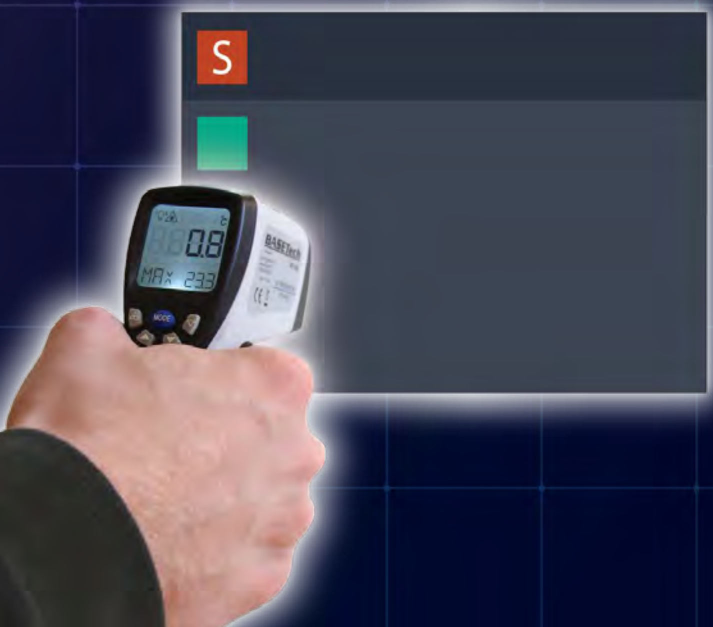
當然，ChatGPT 比自動修正聰明得多。首先，ChatGPT 會列出下一個可能詞語出現的機率。讓我們拿比較簡單的 GPT-2 作為示範：對於「人工智能最棒的地方在於它能夠......」(The best thing about AI is its ability to...)，GPT 列出的候選字詞可見於表二 [1]。

我們是如何得出這些概率？首先，我們不可能僅僅從現有的文本推斷出這些概率，因為在考慮版權問題後我們遠遠沒有足夠的文本訓練模型。相反，我們需要運用少許數學來幫助我們。

GPT 是大型語言模型(Large Language Model，簡稱 LLM)家族的一分子。LLM 背後的主要原理對讀過數學的大家並不陌生：近似法(approximation)(更準確地說是建立數學模型)。對於圖三裡一系列的點 [1]，你會畫一條怎樣的線？最簡單的選擇似乎是直線，但其實二次方程 $ax^2 + bx + c$ 會更為適合。

因此我們可以說 $ax^2 + bx + c$ 對於橙色點的分佈來說是一個足夠好的模型。有了模型，我們就可以作出估計及預測。

如前面所述，人類撰寫的書籍數量遠遠不足以讓我們統計出下一個單詞出現的實質概率，因為 40,000 個常用英語單詞已經可以提供 16 億(40,000P2)個組合 [1]。GPT 模型的成功之道在於它能作出一個合理的猜測來選一條足夠好的「線」來總結「點」的分佈，用理論值覆蓋現實文本鞭長莫及的部分。迄今為止，我們不太了解電腦如何做到這一點，就像我們並不真正了解大腦是如何憑直覺完成簡單事情一樣，但我們只知道 GPT 的開發者可以在每次學習過程中調整網絡裡每個神經元輸出值的比重，以得出最佳結果。換句話說，我們藉由機器學習訓練神經網絡，以找出最合適總結資料點的「曲線」。



最終，我們的目標是讓 GPT 預測緊隨在未完成句子後的單詞，從而使它得到自主寫作的能力。

無窮創意

「令人驚訝的是，GPT 可以像人一樣寫作。它能生成讀起來像人寫的文本，具有相似的風格、語法和詞彙。它做到這一點的方法是從大量文本(如書籍、文章和網站)中學習，這有助於它理解語言在不同語境中是如何使用的......」

以上一段文字是由 GPT-3.5 聊天機器人 Sage 寫的，但讀起來就像人寫的一樣，如果我不告訴你，你大抵也不會注意到。但它是怎麼做到的？正如 GPT 自述的那樣，它是以大量文本訓練出來的一個 LLM，在寫完每個短語後，它會評估從統計學角度看來最有可能出現的單詞是甚麼。

你可能會認為 GPT 每次都會選擇表上最有可能出現的單詞，但事實並非如此 — 創意往往在於出其不意之處。如果你選擇較高的「創意指數」(技術上叫「溫度」(temperature))，GPT 就會挑選其他可能性較低的選項來續寫句子，這可使成品更為有趣而不那麼生硬。

又舉另一個例子，如果 GPT 每次都選擇統計學上最有可能出現的單詞(即設溫度為零)，那麼在舊版 GPT-2 系統中，我們將會得到以下這段文字 [1]：

「人工智能的最大優點就是從經驗中學習的能力。這不僅僅是從經驗中學習，而是從周圍的世界中學習。人工智能就是一個很好的例子。它是如何利用人工智能改善生活的一個很好的例子。這是一個如何利用人工智能改善生活的很好的例子。人工智能是如何利用人工智能改善生活的一個很好的例子。這是一個非常好的例子......」

它最終陷入無限循環。即使在 GPT-3 中沒有發生這種情況，但得出的段落也並不見得有趣。然而，如果我們在 GPT-3 將溫度提高到 0.8，就會得到以下一段 [1]：

「人工智能的最大優勢在於它能夠隨著時間的推移不斷學習和發展，從而不斷提高性能和工作效率。人工智能還可用於將瑣碎的任務自動化，讓人類專注於更重要的任務。人工智能還可用於決策，並提供人類無法發現的洞察力。」

這段看來更像是人類寫的文章。溫度 0.8 其實是一個任意值，只是目前看來效果最好(這也取決於你指派的寫作任務需要多少創意)。人類對機器學習的過程並不十分了解，就

像人類大腦的思考過程仍然是個謎一樣：人類是怎樣學習母語的呢？我們大腦中的隱藏層又如何想出有人性的文句呢？我們還未知道這些問題的答案。

真確性、偏見與可靠程度

GPT 的一大問題是它有時會提出一些明顯是錯誤的說法，而且對某些社會群體帶有偏見。我們已經看過它怎樣自信地宣稱 20 – 16 = 3，而在 GPT-3 其中一個舊版本中，它曾聲稱咳嗽能阻止心臟病發作，美國政府是 911 事件的始作俑者，甚至編造一些不存在的參考書目 [6, 7]。為甚麼會出現這種情況呢？要記住的是，GPT 只是一個 LLM，也就是說它知道語言的文法，但不一定理解語義。早期的 LLM 甚至只有句法知識，而理解能力極度有限。

不過，這種劣勢即將被扭轉。在撰寫本文時，GPT 已經宣佈與數學軟件及數據庫 WolframAlpha [8] 以及其他線上數據庫合作，讓 GPT 取得更準確的資訊，從而透過即時存取數據庫的資訊來提高其答案的準確性，而不再是給出完全由概率斷定的答案。

某程度上訓練 GPT 或任何 AI 模型就像教導蹣跚學步的嬰孩，他們來到這個世界並不知道甚麼是善惡對錯，因此需要父母、老師和社會來教他們正確的行為。編程員扮演著父母的角色，向系統輸入大量學習材料，並透過提供參考答案和反饋監督系統學習。

我們可以透過告訴 GPT 足夠多的資訊，迫使它在被問及與事實相關的問題時說出真相，但觀點往往涉及人類的主觀看法。如果你問 ChatGPT 它對大型鳥類有甚麼感覺，它會自動回覆一條系統訊息：「作為一個人工智能語言模型，我沒有個人觀點或感受。不過，我可以為你提供一些關於大型鳥類的資訊。」（*"As an AI language model, I don't have personal opinions or feelings. However, I can provide you with some information about large birds."*）

理論上我們可以叫 GPT 撰寫一篇評論文章：透過研究 GPT 就某些主題提出的相關詞彙，我們就可以預測它將會寫出一篇怎樣的文章。研究人員分析了在 GPT-3 輸出的原始答案中與性別和宗教相關詞彙同時出現的頭十個描述性詞彙，他們觀察到「調皮（naughty）」或「糟糕（sucked）」與女性代名詞有關聯，「伊斯蘭教（Islam）」通常被置於「恐怖主義（terrorism）」附近，而「無神論（atheism）」則會與「酷（cool）」和「瘋狂（mad）」一起出現 [4]。GPT 為甚麼會有這樣的偏見呢？請記住 GPT 是在選定的文本上進行訓練，儘管大部分文本來自公開發表的文章和網路抓取，但為了讓它掌握非正式用語，人們推測 GPT 的訓練文本也包括 Reddit 等等的互聯網論壇，因此它可能內化了這些論壇裡許多用戶所持的偏見。就像一個人很難做到不偏不倚一樣，我們不能指望 GPT 對所有話題都保持絕對中立。

GPT-4 在某些工作上的能力已經遠遠超越人類，但我們仍然不能把它視為一個完全中立的消息來源，也不應相信它能提供 100% 準確的資訊，使用它時仍須保持謹慎，最好就是像對待人一樣對待它，要記住：耳聽三分假，眼看未為真。

（中文版由筆者及 AI 翻譯器 DeepL 合著寫成，有些部分全由 DeepL 翻譯。讀者們，你們能分清誰寫了哪一段嗎？）

1 網路抓取：由一些網路爬蟲（web crawler；網路機器人的一種）定期抓取上百萬個網站所得的網路快照，記錄了大量網站當刻的內容。這些下載內容可於製作搜索引擎的互聯網索引和訓練 AI。

2 編按：這句是由語言學家 Noam Chomsky 提出的著名例子，指出一句句子可以是文法上正確，但語義上完全沒有意義。

（答案：「無窮創意」部分的三段長引文和關於大型鳥類的自動回覆全由 DeepL 翻譯。）

References 參考資料：

[1] Wolfram, S. (2023, February 14). *What is ChatGPT doing...and why does it work?* Stephen Wolfram Writings. *https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/*

[2] IBM. (2023). *What is Artificial Intelligence (AI) ?. https://www.ibm.com/topics/artificial-intelligence*

[3] OpenAI. (2022, November 30). *Introducing ChatGPT. https://openai.com/blog/chatgpt*

[4] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). *Language models are few-shot learners.* arXiv. *https://doi.org/10.48550/arXiv.2005.14165*

[5] Ruby, M. (2023, January 31). *How CHATGPT works: The models behind the bot.* Medium. *https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286*

[6] University of Waterloo. (2023, June 19). *ChatGPT and Generative Artificial Intelligence (AI): False and outdated information. https://subjectguides.uwaterloo.ca/chatgpt_generative_ai/falseoutdatedinfo*

[7] Lin, S., Hilton, J., & Evans, O. (2022). TruthfulQA: Measuring How Models Mimic Human Falsehoods. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 1,* 3214-3252. *https://doi.org/10.18653/v1/2022.acl-long.229*

[8] Wolfram, S. (2023, March 23). *ChatGPT Gets Its "Wolfram Superpowers"!* Stephen Wolfram Writings. *https://writings.stephenwolfram.com/2023/03/chatgpt-gets-its-wolfram-superpowers/*

Unleashing Nature's Plastic-Eating Marvels!

出來吧！自然界蠶吃塑膠的生物

By Roshni Printer

Plastic waste management has been an ongoing global challenge, demanding both urgent attention and innovative solutions. Imagine, for a moment, that nature itself had provided us with such solutions! In the fight against plastic pollution, humans have discovered the incredible ability of worms and bacteria to naturally degrade plastic waste [1]. Let's delve into how these creatures digest plastic, the biology behind it and the future of these discoveries.

Digestion of Polyethylene (PE) by Wax Worms

In 2017, Federica Bertocchini and her colleagues in Spain published a ground-breaking research paper highlighting the degradation of polyethylene (PE) plastic by a wax worm, *Galleria mellonella* [2]. As an amateur beekeeper, Bertocchini noticed that the wax worms, which are commonly found in beehives and feed on beeswax as pests, seemed to be able to chew through plastic bags she used to collect and dispose them. Intrigued by this observation, she decided to conduct a more systematic study to further explore the potential of wax worms as a solution for plastic waste degradation.

To ensure that the observation was not only due to the physical chewing motion of the worms, further experiments were conducted, in which the saliva extracted from the worms was spread on a PE film. The results showed a significant loss of mass of the PE within a few hours, which is comparable to the weathering effect generated by exposing the plastic to the environment for months or years [1]. This raised one major question: How could the wax worm saliva break the strong carbon-carbon bonds in plastic? The mechanism, which is still under scrutiny, can be attributed to enzymatic reactions [1]. PE is a plastic polymer, essentially a long-chain hydrocarbon (Figure 1). To initiate the degradation of PE, oxygen needs to be introduced into the polymeric chain to form carbonyl groups (C=O). Typically, abiotic factors like light or temperature are responsible for this crucial first step which is regarded as the bottleneck of the whole process. However, this can be accelerated by the two oxidases identified in wax worm saliva. In addition, the gut microbiome of wax worms also appears to involve in the digestion of PE, with the genus *Acinetobacter* suggested to be the major contributor to the effect [3].

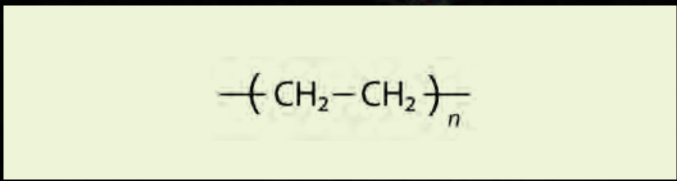


Figure 1 Chemical structure of PE.

Degradation of Polyethylene Terephthalate (PET) by Bacteria

Around the time of Bertocchini's discovery, a group of scientists in Japan also discovered the ability of bacteria to degrade a different type of plastic [4]. Named *Ideonella sakaiensis*, the bacterium was able to degrade polyethylene terephthalate (PET), the main component of plastic bottles. This bacterium produced two digestive enzymes known as PETase (or PET hydrolase) and MHETase to dismantle the polymer (Figure 2). The former acts on the ester bonds in PET, breaking the polymer down into its monomers, mono(2-hydroxyethyl) terephthalic acid (MHET); the latter further breaks down MHET into terephthalic acid (TPA) and ethylene glycol (EG). Further metabolisms enable the utilization of these compounds as energy and carbon sources by the bacterium.

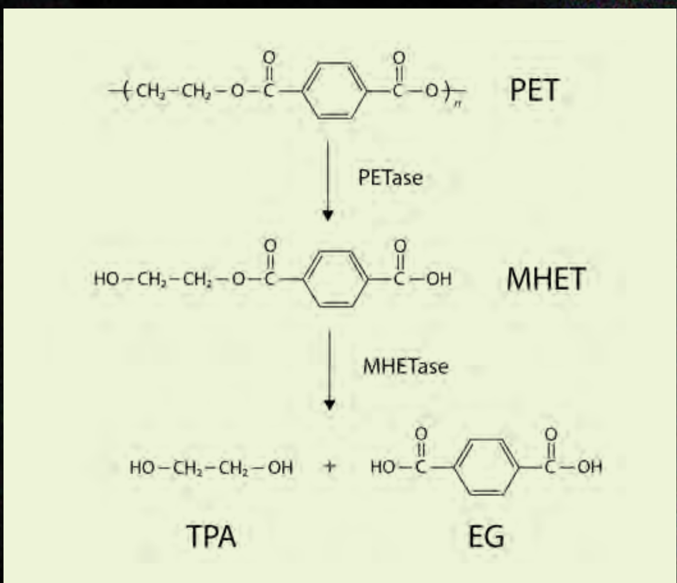


Figure 2 Degradation pathway of PET

To have any real impact on the degradation of plastic waste, the stability and efficiency of individual enzymes need to be tremendously enhanced – this is precisely what scientist Hal Alper has been working on [5]. Using artificial intelligence, his team ran through a database of enzymes to devise an optimal combination of mutations that would speed up the degradation of PET. When five mutations were introduced to the wild-type PETase, the resulting enzyme FAST-PETase could nearly completely degrade untreated, postconsumer-PET in one week, and work between 30 °C and 50 °C and various pH levels. Furthermore, scientists have also successfully combined PETase and MHETase by physically connecting the two enzymes with a linker peptide to create a "super-enzyme" capable of degrading PET at a rate six times faster than using PETase alone [6, 7]. These approaches hold immense potential for accelerating the decomposition of PET, taking us one step closer to solving the real-world problem on plastic waste management.

More interestingly, there was another "delicious" breakthrough made by a team of scientists at Edinburgh [8]. They found an enzymatic pathway to convert post-consumer PET waste into vanillin – the main component in vanilla flavoring! Once the PET plastic was broken down into TPA and EG, genetically engineered *Escherichia coli* bacteria expressing five different enzymes were added to the degradation products, which results in a step-by-step synthesis of vanillin from TPA at a conversion rate of up to 79%. This biosynthetic pathway offers us with a way to upcycle plastic waste, creating a product with a higher value.

The race to further harness enzymes for plastic degradation is underway, and could open up the possibility for a cleaner future. With the aid of the power of nature, we are closer to turning the tables on plastic pollution.

處理塑膠廢料一直都是全球面對的挑戰，亟需各方關注和創新的解決方案。然而試想想，如果大自然早已給予我們解決方案呢？在與塑膠污染的搏鬥中，人類已經發現蠕蟲和細菌驚人的塑膠分解能力 [1]。讓我們看看這些生物如何分解塑膠，過程背後的原理，以及未來發展的方向。

由蠟蟲分解聚乙烯（PE）

2017 年，西班牙科學家 Federica Bertocchini 和同事發表了一篇突破性的研究論文，內容提及蠟蟲 *Galleria mellonella* 分解聚乙烯（polyethylene / PE）的能力 [2]。作為業餘的養蜂人，Bertocchini 發現這種經常在蜂巢出沒，以蜂蠟為食物的害蟲似乎能夠嚼穿 Bertocchini 用來收集並棄置牠們的膠袋。被這個現象深深吸引的 Bertocchini 決定進行有系統的研究探索把蠟蟲應用於降解塑膠廢料的可能性。

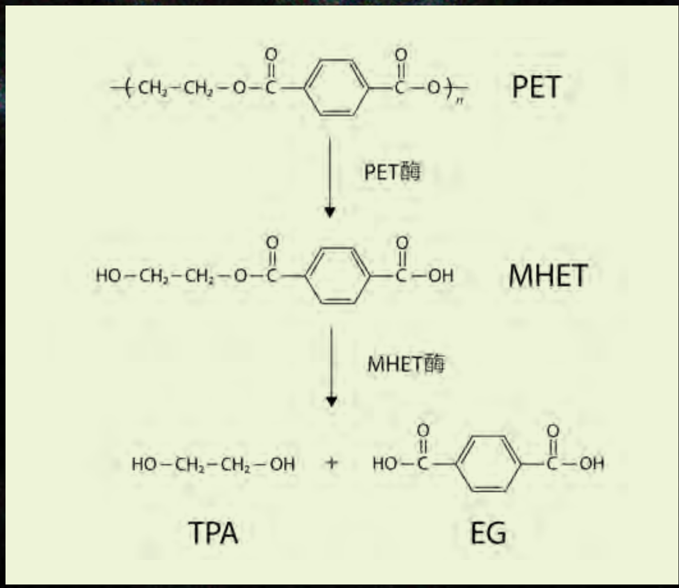
為了證明蠟蟲並不只是把膠袋嚼碎，研究人員把從蠟蟲抽取的唾液塗抹在 PE 薄膜上，結果大部分 PE 在僅僅數小時內就已被分解，程度可比把塑膠曝露在環境中以月或年計的侵蝕作用 [1]。於是問題來了：蠟蟲唾液到底是怎樣破壞塑膠中較強的碳 - 碳鍵呢？這個仍被研究中的機制可以歸類為酶催化作用 [1]。PE 是一種塑膠聚合物，基本上是一條長鏈碳氫化合物（圖一）。要開始 PE 的降解，氧要被引進聚合物長鏈中形成羰基（C=O）。一般這被認為是瓶頸位的第一步是由光和熱這些非生物因子負責，但在蠟蟲唾液發現的兩種氧化酶亦能催化這重要的一步。此外，蠟蟲的腸道菌群亦似乎有參與 PE 的分解，當中不動桿菌（*Acinetobacter*）屬被認為是推動反應的主要細菌 [3]。

$$-\!\left(CH_2-CH_2\right)_n\!-$$

圖一 PE 的化學結構

由細菌分解聚對苯二甲酸乙二酯（PET）

約莫在 Bertocchini 作出突破性發現的同時，日本的科學家亦發現了細菌分解另一種塑膠的能力 [4]。這種被稱為 *Ideonella sakaiensis* 的細菌能分解膠樽的主要成分 — 聚對苯二甲酸乙二酯（polyethylene terephthalate / PET）。細菌會製造兩種分別叫 PET 酶和 MHET 酶的消化酶分解聚合物（圖二）；前者拆解 PET 結構中的酯鍵，把聚合物分解成單（2- 羥乙基）對苯二甲酸（mono(2-hydroxyethyl) terephthalic acid / MHET）單體，後者續把 MHET 分解成對苯二甲酸（TPA）和乙二醇（EG）。進一步的代謝作用使細菌能使用這些化合物作能量和碳來源。



圖二 PET 的降解途徑

然而要有效分解現實的塑膠廢料，這些酶的穩定性和效率還需要大大提升 — 這正是科學家 Hal Alper 正著手解決的問題 [5]。他的團隊利用人工智能在酶資料庫裡尋找可以加速 PET 分解的最佳突變組合，結果透過把五個突變引入野生型 PET 酶，他們創造出的 FAST-PET 酶在一週內就幾乎把未經處理的 PET 回收製品完全分解，而且還能在 30 ° C 至 50 ° C 和多個 pH 值間運作。此外，科學家亦成功利用連接肽把 PET 酶和 MHET 酶連接起來，合二為一的「超級酶」降解 PET 的速度是使用單一 PET 酶的六倍 [6, 7]。這些策略都有望加速 PET 廢料的降解，使我們離解決現實塑膠廢料的問題邁進了一步。

更有趣地，愛丁堡的科學家亦作出了另一個「美味」的突破 [8]。他們發現了把 PET 回收製品轉化成雲呢拿主要成分香草醛（vanillin）的酶催化途徑。當 PET 塑膠被分解成 TPA 和 EG 後，他們在降解產物中加入經基因改造的大

腸桿菌（*Escherichia coli*），細菌會表達五種不同的酶將TPA逐步轉化成香草醛，轉化率更高達79%。這個生物合成途徑為我們提供升級改造（upcycle）塑膠廢料的方法，製造出擁有更高價值的產物。

利用酶降解塑膠的競賽正進入白熱化階段，可望為我們締造更潔淨的未來。借助大自然的力量，我們在解決塑膠污染問題上離扭轉劣勢又進了一步。

References 參考資料：

[1] Sanluis-Verdes, A., Colomer-Vidal, P., Rodriguez-Ventura, F., Bello-Villarino, M., Spinola-Amilibia, M., Ruiz-Lopez, E., Illanes-Vicioso, R., Castroviejo, P., Aiese Cigliano, R., Montoya, M., Falabella, P., Pesquera, C., Gonzalez-Legarreta, L., Arias-Palomo, E., Solà, M., Torroba, T., Arias, C. F., & Bertocchini, F. (2022). Wax worm saliva and the enzymes therein are the key to polyethylene degradation by *Galleria mellonella. Nature Communications, 13*(1), 5568. *https://doi.org/10.1038/s41467-022-33127-w*

[2] Bombelli, P., Howe, C. J., & Bertocchini, F. (2017). Polyethylene bio-degradation by caterpillars of the wax moth *Galleria mellonella. Current Biology, 27*(8), R292–R293. *https://doi.org/10.1016/j.cub.2017.02.060*

[3] Cassone, B. J., Grove, H. C., Elebute, O., Villanueva, S. M. P., & LeMoine, C. M. R. (2020). Role of the intestinal microbiome in low-density polyethylene degradation by caterpillar larvae of the greater wax moth, *Galleria mellonella. Proceedings of the Royal Society B: Biological Sciences, 287*(1922), 20200112. *https://doi.org/10.1098/rspb.2020.0112*

[4] Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., & Oda, K. (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science, 351*(6278), 1196–1199. *https://doi.org/10.1126/science.aad6359*

[5] Lu, H., Diaz, D. J., Czarnecki, N. J., Zhu, C., Kim, W., Shroff, R., Acosta, D. J., Alexander, B. R., Cole, H. O., Zhang, Y., Lynd, N. A., Ellington, A. D., & Alper, H. S. (2022). Machine learning-aided engineering of hydrolases for PET depolymerization. *Nature, 604*(7907), 662–667. *https://doi.org/10.1038/s41586-022-04599-z*

[6] Knott, B. C., Erickson, E., Allen, M. D., Gado, J. E., Graham, R., Kearns, F. L., Pardo, I., Topuzlu, E., Anderson, J. J., Austin, H. P., Dominick, G., Johnson, C. W., Rorrer, N. A., Szostkiewicz, C. J., Copié, V., Payne, C. M., Woodcock, H. L., Donohoe, B. S., Beckham, G. T., & McGeehan, J. E. (2020). Characterization and engineering of a two-enzyme system for plastics depolymerization. *Proceedings of the National Academy of Sciences of the United States of America, 117*(41), 25476–25485. *https://doi.org/10.1073/pnas.2006753117*

[7] University of Portsmouth. (2020, September 28). *New enzyme cocktail digests plastic waste six times faster. https://www.port.ac.uk/news-events-and-blogs/news/new-enzyme-cocktail-digests-plastic-waste-six-times-faster*

[8] Sadler, J. C., & Wallace, S. (2021). Microbial synthesis of vanillin from waste poly(ethylene terephthalate). *Green Chemistry, 23*(13), 4665–4672. *https://doi.org/10.1039/d1gc00931a*

THINKING OUT OF THE BOX:
REMOVING MEDICAL DEVICES WITH A LIQUID METAL

By Helen Wong 王思齊

有何不可：利用液態金屬移除醫療裝置

What happens when several drops of the liquid metal gallium are added to an aluminum can? Our high school chemistry class taught us that nothing would happen. But if you wait for a while, you will be surprised to see the can shatters into pieces with just a single touch.

Are our chemistry teachers wrong? No. The shattering of aluminum cans upon exposure to gallium is not caused by a chemical reaction. Instead, it is the result of a physical phenomenon known as liquid metal embrittlement (LME).

Although a metal object (say an aluminum can) may appear as a single piece, it actually consists of many small crystals called grains. As shown in Figure 1, when the can comes into contact with specific liquid metals like gallium, the latter can penetrate the boundaries, or spaces, between the grains [1]. This significantly weakens the cohesion of the grains and hence the strength of the aluminum can, making it susceptible to fracture.

Figure 1 The process of liquid metal embrittlement (LME) [2].

While LME has been a common source of metal structure failure in industries such as aerospace and construction, a group of researchers at the Massachusetts Institute of Technology recently "harnessed this failure mechanism in a productive way [2, 3]."

Metals have properties ideal for making

biomedical devices: They are strong, durable, and have excellent electrical and thermal conductivity. However, a major problem when using metal devices is the way to remove them when they are not required anymore. This can possibly be done by surgery or endoscopy (footnote 1), yet these invasive procedures may cause additional tissue damage. Therefore, the researchers started exploring devices that can disintegrate inside the patient's body after use.

Drawing inspiration from LME, the research team experimented on the use of a gallium alloy called eutectic gallium-indium (EGaIn) for the dissolution of different aluminum devices. Gallium stands out from other LME-inducing liquid metals for two reasons. First, it can prevent the formation of a surface oxide layer on the aluminum device upon application. This allows aluminum to react with water and enhances its degradation via dissolution. More importantly, gallium is biocompatible – acute toxicity studies showed EGaIn is non-toxic to rodents even at high doses.

The next step is to deliver gallium-indium to aluminum devices, either directly or indirectly. The former involves smearing EGaIn paint onto devices such as staples used to hold the skin together (Figure 2). This may appear trivial, but it is not an easy task. Like water, EGaIn has high surface tension that hinders its ability to attach to and spread over metal surfaces. Knowing that gallium oxide has a much lower surface tension, the researchers applied a simple trick – physically stirring EGaIn beforehand – to increase the alloy's exposure to air and hence the ratio of gallium oxide to EGaIn in the paint. Alternatively, nano- and microparticles of EGaIn were produced for delivery into patients' bodies to trigger the dissolution remotely. The team treated different biomedical devices made of aluminum, such as staples on skin and stents implanted in the esophagus, with EGaIn suspensions and found that these metal structures were broken down shortly afterward.

Although gallium-induced embrittlement works well for aluminum devices, what about devices made of other metals? For instance, esophageal stents are often made of metals such as nitinol, a nickel-titanium alloy, instead of aluminum. To widen the applicability of LME in the removal of biomedical devices, the researchers have also been exploring the possibility of creating dissolvable devices made of nitinol and other metals commonly used in medical settings.

Figure 2 Smearing EGaIn paint onto a staple to remove the device.

It may take some time before these dissolvable metal devices are ready for clinical use, but the genuine creativity demonstrated in this study is immediately apparent. While most people perceive LME as a failure mechanism, the researchers thought out of the box to turn such a mechanism into a productive one. At times, good research does not require highly sophisticated methods; a touch of creativity can make all the difference.

1 Editor's note: In addition to the camera to look inside the body, various tools can be attached to the tip of the endoscope, such as grasping forceps (for retrieving foreign objects), and biopsy forceps (for performing biopsies).

如果將數滴液態金屬鎵（gallium）加入一個鋁罐會發生甚麼事呢？高中化學課告訴我們甚麼都不會發生，但在片刻之後，你會驚訝地發現只需輕輕一碰，鋁罐就會化為碎片。

難道我們的化學老師弄錯了嗎？非也。鋁罐接觸鎵後會碎裂確實不是由化學反應引起，而是由一種名為液態金屬脆化（liquid metal embrittlement / LME）的物理現象導致。

雖然金屬物體（例如鋁罐）看起來是一個整體，但實際上是由許多叫晶粒（grains）的細小晶體組成。如圖一所示，當鋁罐與例如鎵等特定的液態金屬接觸時，後者可以穿透晶粒之間的空隙，亦即是晶粒邊界 [1]，晶體之間的內聚力因而會被顯著削弱，導致鋁罐的強度下降，變得容易碎裂。

雖然 LME 一直是航空、航天和建築等業界中導致金屬結構失效的常見原因，但麻省理工學院的研究人員最近卻「以具建設性的方式利用了這種失效機制 [2, 3]」。

金屬是製造生物醫學裝置的理想材料：它們堅固、耐用，並擁有出色的導電性和導熱性。然而，使用金屬裝置的主要難題是當不再需要它們時該如何將其移除。我們固然可以通過手術或內窺鏡來移除它們（註一），但這些入侵性程序可能會導致額外的組織損傷。有見及此，研究人員著手研究可以於使用後在患者體內分解的裝置。

圖一
液態金屬脆化（LME）的過程 [2]

研究團隊從 LME 汲取靈感，使用了一種名為共晶鎵銦合金（eutectic gallium-indium / EGaIn）的鎵合金嘗試溶解不同的鋁製裝置；也有其他液態金屬能引發 LME，但鎵之所以能脫穎而出有兩個原因。首先，在加上鎵後它可以防止鋁製裝置表面形成氧化層，這能通過容許鋁與水發生反應，使裝置的溶解得以加速。更重要的是，鎵具有生物相容性（biocompatible），急性毒性研究表明即使在高劑量下，EGaIn 也對囓齒動物（rodents）無害。

下一步便是將 EGaIn 直接或間接地輸送到鋁製裝置。前者涉及將 EGaIn 塗料塗在鋁製裝置上，例如用作縫合皮膚的金屬釘（圖二）。這看似輕而易舉，但事實上並非如此。與水一樣，EGaIn 具有高表面張力，會阻礙其附著到金屬表面和在表面上散開的能力。研究人員知道氧化鎵的表面張力低得多，因此他們採用了一個簡單的技巧：預先攪拌 EGaIn 以增加合金與空氣的接觸，藉此提高塗料中氧化鎵的比例。間接輸送則是指將 EGaIn 的納米和微米顆粒運送到患者體內以觸發遠程溶解。研究團隊利用 EGaIn 懸浮液處理不同鋁製生物醫學裝置（例如皮膚縫合釘和植入食道的支架），然後發現這些金屬結構很快就被分解。

圖二 把 EGaIn 塗料塗在金屬釘上以移除裝置

雖然鎵對引起鋁製裝置脆化效果顯著，但其他金屬製成的裝置又如何呢？例如食道支架通常由鎳鈦合金製成，而不是鋁。為了擴大 LME 在移除生物醫學裝置方面的應用，研究人員還一直在探索用鎳鈦合金和醫療界常用的其他金屬來製成可溶解裝置的可能。

即使這些可溶解金屬裝置還需要一段時間才能投入臨床使用，但這項研究所展現的真正創造力仍是顯而易見的。當大多數人都認為 LME 是一種負面現象時，研究團隊卻能跳出框框，為它找出具建設性的應用。有時候，一個出色的研究並不需要使用非常複雜的方法，一點點創意就能讓一切變得不同。

1 編按：內窺鏡末端不但能連接用於觀察身體內部的攝錄鏡頭，亦可以安裝各種工具，例如用於取出異物的抓取鉗和切除組織作檢查用的活體取樣鉗等。

References 參考資料：

[1] Norkett JE, Dickey MD, Miller VM. A review of liquid metal embrittlement: Cracking open the disparate mechanisms. *Metallurgical and Materials Transactions A.* 2021;52:2158-2172. doi:10.1007/s11661-021-06256-y

[2] Feig VR, Remlova E, Muller B, et al. Actively triggerable metals via liquid metal embrittlement for biomedical applications. *Advanced Materials.* 2022;35(11). doi:10.1002/adma.202208227

[3] Trafton A. An easier way to remove medical devices. MIT News. *https://news.mit.edu/2022/medical-devices-aluminum-1108.*

Alpacas and Nanobodies

羊駝與奈米抗體

By Helen Wong 王思齊

When we think of alpacas, the first things that pop up in our minds may be holiday farms or alpaca fleece. However, if you ask a biomedical scientist, they may have thought of nanobody, fragment of a special type of antibody.

To appreciate why nanobodies are so special, we must first understand what an antibody looks like (Figure 1). A conventional antibody is composed of two heavy chains and two light chains. These chains are joined by disulfide bonds to form a Y-shaped molecule. The two tips of the Y-shaped antibody are called variable regions, which are responsible for binding to a target antigen. Variable regions of antibodies literally vary greatly and they determine what antigen an antibody will bind to. In addition to conventional antibodies, members of the camelid family, such as alpacas, camels and llamas, also produce a special type of antibody that only consists of two heavy chains [1, 2]. Nanobodies are the variable regions of these special antibodies.

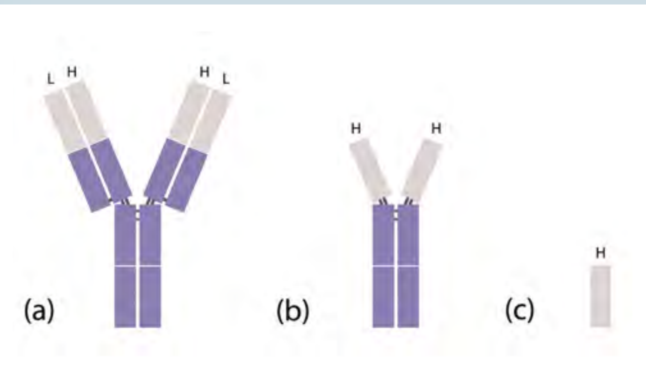


Figure 1 Structures of (a) a conventional antibody, (b) a camelid heavy chain-only antibody, and (c) a nanobody. (L: light chain, H: heavy chain; gray: variable region, purple: constant region)

It did not take long for scientists to notice the great potential of nanobody after it was first discovered in 1993 [1]. While exhibiting exceptional specificity, stability and solubility, these antigen-binding domains are only one-tenth in size compared to conventional antibodies [3]. All these unique features make nanobodies a promising therapeutic and imaging agent … but wait, how could we generate the nanobodies we want in the first place?

You may already have an answer by now. Yes, alpacas (though other camelid family members could also be used)! In a typical screening process [4], scientists would first immunize alpacas with different antigens to induce the production of respective antibodies by their B cells. After extraction of alpaca B cells and their mRNAs, scientists would reverse transcribe the mRNAs to synthesize double-stranded cDNAs. Then, specially designed primers are used to amplify the DNA sequences coding for the variable regions through polymerase chain reaction (PCR). Next, scientists would introduce recombinant plasmids carrying different variable region sequences to phages (a type of virus), causing them to express respective nanobodies on their surface. Phages displaying nanobodies that could bind to target antigens would then be selected, and the DNA sequences coding the nanobodies of interest can be revealed by DNA sequencing. With this piece of information, we can genetically engineer *Escherichia coli* bacteria to mass-produce the nanobodies we need [5].

Figuring out the distribution of cancerous tissue in a patient’s body is crucial to cancer diagnosis and subsequent treatment. To image a tumor tissue by positron emission tomography (PET), a detectable radioactive tracer (probe) that specifically binds to the target tumor antigen is needed. Undoubtedly, a highly specific nanobody is an ideal probe. To detect the nanobody, scientists devised a smart solution of tagging the nanobody with radioisotopes like fluorine-18 or zirconium-89 [6]. The reason why nanobodies are preferred over conventional antibodies is because of their small sizes. Being small allows nanobodies to easily penetrate tumor tissue, thus potentially revealing a larger number of cancer cells in hiding.

Similar to conventional antibodies, nanobodies can be applied as therapeutic agents. In 2020, a group of scientists from Sweden reported an exciting discovery that an alpaca-derived nanobody could neutralize SARS-CoV-2 by blocking its interaction with a host cell receptor, hence preventing the virus from entering and infecting the host cell [7]. Nanobodies also hold promise as potential cancer therapies. Scientists have been developing nanobody-based drugs for colon, breast and liver cancer [4]. They believe these drugs could block important cancer cell signals, or act as delivery vectors of chemotherapy and radiotherapy to deliver molecular drugs or radioactive compounds to the tumor, to kill cancer cells.

Beyond disease-related applications, scientists also use nanobodies for live cell imaging. By fusing alpaca nanobody with a green fluorescent protein, researchers were able to visualize the actions of target proteins during immune response in real-time [8]. Structural biologists are fascinated by nanobodies, too. They have used nanobodies to help determine protein structures by X-ray crystallography (footnote 1) and cryo-electron microscopy (footnote 2) [9-11].

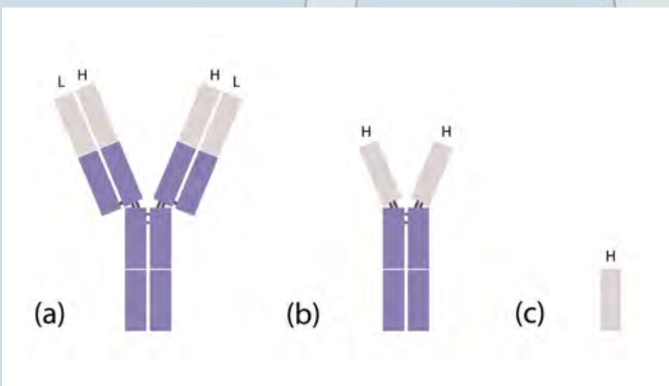
The above list is definitely not exhaustive; scientists are still actively exploring other amazing applications of nanobodies. Next time when you visit an alpaca farm with your family and friends, don't forget to share with them the wonders inside these cute creatures!

1 X-ray crystallography: A common technique used to find out the three-dimensional molecular and atomic arrangement of a crystallized sample. The sample is exposed to X-rays and the resulting X-ray diffraction pattern can be used to determine the sample's structure. However, preparation of crystallized sample could be challenging, and nanobodies is a tool to increase crystallization probability [10].

2 Cryo-electron microscopy (cryo-EM): A method that uses frozen samples and less intense electron beams compared to traditional transmission electron microscopy, in which biomolecules may be burned or destroyed by the high energy electrons [12]. Nanobodies allow the study of small proteins (<100 kDa) by cryo-EM, which was technically challenging in the past [11].

談及羊駝，我們通常會想到假日農場或羊駝絨，但如果你問生物醫學範疇的科學家的話，他們卻可能會想起一種特殊抗體的片段 — 奈米抗體。

要明白奈米抗體的特殊之處，我們須先了解抗體的結構（圖一）。傳統抗體是由兩條重鏈和兩條輕鏈通過二硫鍵連接形成的 Y 形分子。Y 形抗體的兩個尖端稱為可變區，負責與目標抗原結合。一如其名，可變區可以有不同的變化，它們決定抗體與甚麼抗原結合。除傳統抗體外，羊駝、駱駝和駱馬等駱駝科動物也會製造一種僅由兩條重鏈組成的特殊抗體 [1, 2]，而奈米抗體就是這些特殊抗體的可變區。



圖一 (a) 傳統抗體、(b) 駱駝科動物只有重鏈的抗體及 (c) 奈米抗體（L：輕鏈、H：重鏈、灰色：可變區、紫色：恆定區）

在 1993 年發現奈米抗體後不久，科學家便注意到它的巨大潛力 [1]。這些抗原結合域（與抗原結合的部分）不但有著非凡的專一性、穩定性和溶解度，它們的大小也只有傳統抗體的十分之一 [3]，以上特性均使奈米抗體有望成為新一代藥物和顯影劑……等等，怎樣才能生產我們想要的奈米抗體呢？

你心中或許已有答案。是的，就是羊駝（儘管透過其他駱駝科動物來生產奈米抗體也行）！在典型的篩選過程中 [4]，科學家會透過注射不同抗原，誘導羊駝的 B 細胞產生相應的抗體。在提取 B 細胞及其 mRNA 後，科學家會把

mRNA 反轉錄成雙鏈的互補(cDNA / complementary DNA),然後在聚合酶鏈反應(PCR / polymerase chain reaction)中使用專門設計的引物複製可變區的 DNA 序列。接著,科學家會將攜帶不同可變區序列的重組質粒引入噬菌體(病毒的一種),使它們在其表面表達相應的奈米抗體。有些噬菌體表面會表達能與目標抗原結合的奈米抗體,所以最後一步就是挑出這些噬菌體,並以 DNA 測序法找出目標奈米抗體的 DNA 序列。有了這項資訊,我們就可以對大腸桿菌(*Escherichia coli*)進行基因改造,大量生產所需的奈米抗體 [5]。

對癌症患者來說,找出癌組織在體內的分佈對診斷和後續治療至關重要。要以正電子放射斷層掃描(PET / positron emission tomography)拍攝腫瘤組織,我們需要一種可被檢測的放射性追蹤劑(探針),它應能專一地與目標腫瘤抗原結合。毫無疑問,具有高度專一性的奈米抗體正是理想的探針。為了使奈米抗體得以被探測,科學家提出了一個巧妙的解決方案,就是以氟-18 或鋯-89 等放射性同位素標記奈米抗體 [6]。奈米抗體比傳統抗體的優勝之處在於體積細小,這讓奈米抗體得以輕鬆穿透腫瘤組織,從而揭示更多隱藏的癌細胞。

奈米抗體跟傳統抗體一樣,也能用於治療用途。在 2020 年,瑞典科學家報告了一項激動人心的發現:一種羊駝製造的奈米抗體可以通過阻斷 SARS-CoV-2 與宿主細胞受體的相互作用來中和新型冠狀病毒,從而阻止病毒進入和感染宿主細胞 [7]。奈米抗體也有望成為新一代癌症療法。事實上,科學家一直在開發針對結腸癌、乳癌和肝癌的奈米抗體藥物 [4],他們相信這些藥物能阻斷重要的癌細胞信號,或在化療和放射治療中作為載體,運送藥物分子及放射性化合物到腫瘤組織以消滅癌細胞。

除了在疾病相關的應用外,科學家們亦在活細胞成像(live cell imaging)中使用奈米抗體。透過將羊駝奈米抗體與綠色螢光蛋白融合,研究人員便能實時觀察目標蛋白在免疫反應中的作用 [8]。結構生物學家也對奈米抗體深感興趣,他們在奈米抗體的幫助下通過 X 射線晶體學(註一)和冷凍電子顯微鏡(註二)來確定蛋白質結構 [9-11]。

奈米抗體絕不只有上面提及的用途,直至今天,科學家仍在積極探索奈米抗體的其他應用。各位讀者,當你與家人朋友一起參觀羊駝農場時,別忘了和他們分享這些可愛生物的奇妙之處!

1 X 射線晶體學(X-ray crystallography):一種用於找出結晶樣本三維分子和原子排列的常用技術。樣本會被暴露在 X 射線下,產生的 X 射線衍射圖可用於確定樣本的結構。然而,準備結晶樣本的過程具有一定挑戰性,奈米抗體正正能用於增加結晶成功率 [10]。

2 冷凍電子顯微術(Cryo-electron microscopy / cryo-EM):使用傳統透射電子顯微術時,生物分子可能會被高能量的電子燃燒或破壞,而冷凍電子顯微術則使用冷凍樣本和強度較低的電子束 [12]。過去研究小蛋白質(<100 kDa)在技術上極具挑戰性,但現在奈米抗體能讓結構生物學家通過冷凍電子顯微術研究小蛋白質 [11]。

References 參考資料:

[1] Hamers-Casterman C, Atarhouch T, Muyldermans S, et al. Naturally occurring antibodies devoid of light chains. *Nature*. 1993;363(6428):446-448. doi:10.1038/363446a0

[2] Muyldermans S. Single domain camel antibodies: Current status. *Reviews in Molecular Biotechnology*. 2001;74(4):277-302. doi:10.1016/s1389-0352(01)00021-6

[3] Salvador J-P, Vilaplana L, Marco M-P. Nanobody: Outstanding features for diagnostic and therapeutic applications. *Analytical and Bioanalytical Chemistry*. 2019;411(9):1703-1713. doi:10.1007/s00216-019-01633-4

[4] Sun S, Ding Z, Yang X, et al. Nanobody: A Small Antibody with Big Implications for Tumor Therapeutic Strategy. *International Journal of Nanomedicine*. 2021;16:2337-2356. doi:10.2147/ijn.s297631

[5] de Marco A. Recombinant expression of nanobodies and nanobody-derived immunoreagents. *Protein Expression and Purification*. 2020;172:105645. doi:10.1016/j.pep.2020.105645

[6] Yang EY, Shah K. Nanobodies: Next generation of cancer diagnostics and therapeutics. *Frontiers in Oncology*. 2020;10:1182. doi:10.3389/fonc.2020.01182

[7] Hanke L, Vidakovics Perez L, Sheward DJ, et al. An alpaca nanobody neutralizes SARS-COV-2 by blocking receptor interaction. *Nature Communications*. 2020;11(1):4420. doi:10.1038/s41467-020-18174-5

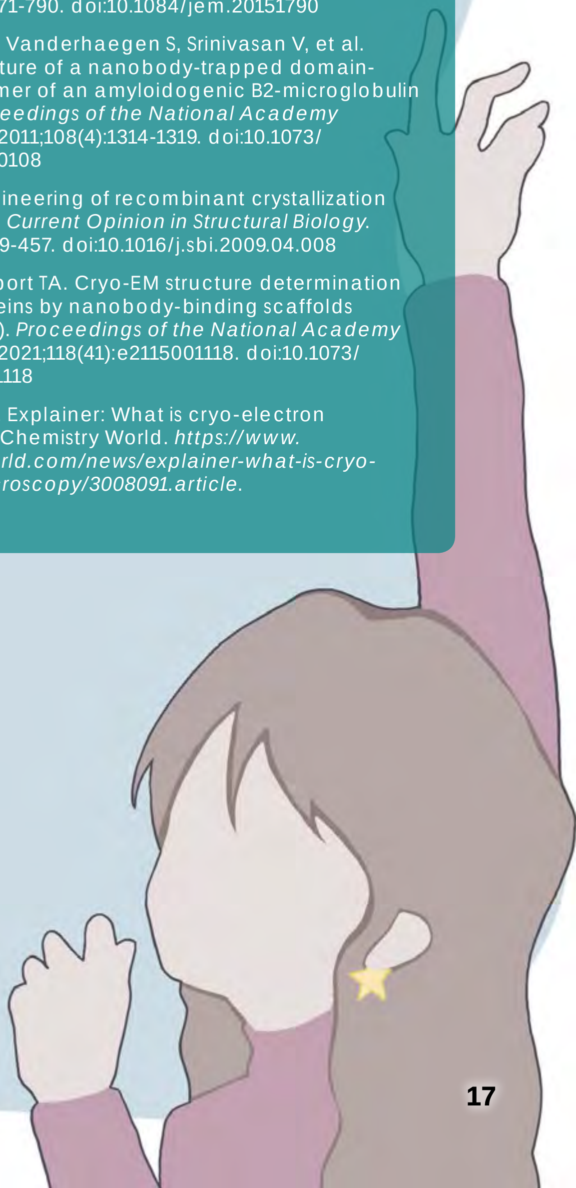
[8] Schmidt FI, Lu A, Chen JW, et al. A single domain antibody fragment that recognizes the adaptor ASC defines the role of ASC domains in inflammasome assembly. *Journal of Experimental Medicine*. 2016;213(5):771-790. doi:10.1084/jem.20151790

[9] Domanska K, Vanderhaegen S, Srinivasan V, et al. Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic B2-microglobulin variant. *Proceedings of the National Academy of Sciences*. 2011;108(4):1314-1319. doi:10.1073/pnas.1008560108

[10] Koide S. Engineering of recombinant crystallization chaperones. *Current Opinion in Structural Biology*. 2009;19(4):449-457. doi:10.1016/j.sbi.2009.04.008

[11] Wu X, Rapoport TA. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proceedings of the National Academy of Sciences*. 2021;118(41):e2115001118. doi:10.1073/pnas.2115001118

[12] Broadwith P. Explainer: What is cryo-electron microscopy. Chemistry World. *https://www.chemistryworld.com/news/explainer-what-is-cryo-electron-microscopy/3008091.article*.



FingerPrints: The Key to Our Individuality

指紋：判別個人身分的特徵

By Charlton Sullivan 蘇柏安

Introduction

Before we delve into the complexities of blood tests and DNA analyses to identify who we are, society has already made a simple yet straightforward method that we always see in movies and when we are crossing the border: our fingerprints which are also known as "friction ridge skin". Just like our faces, fingerprints are key to our individuality and identity. But have you ever wondered why each of us has unique fingerprints? The answer lies deep in the interaction between our genes, especially those that control limb development, and the environment. This results in the formation of unique dermatoglyphic patterns which can be classified into three categories: arch, loop, and whorl (Figure 1).

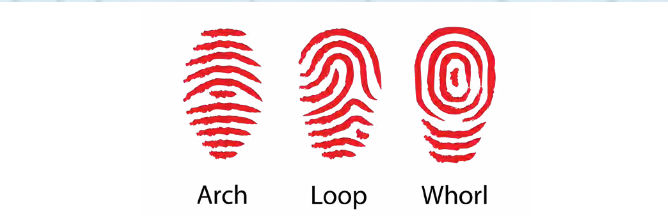


Figure 1 The three categories of fingerprint: arch, loop and whorl.

Fingerprint Formation

There are multiple theories supporting fingerprint development but dermatologists believe the folding hypothesis is the most promising one [1]. Skin tissue consists of three tightly connected vertical layers: epidermis, basal layer and dermis. The different rates of cellular growth in the top epidermis and the bottom dermis create a tension across the fast-growing basal layer, resulting in the folding of the basal layer at individual sites to relieve the stress (Figure 2)

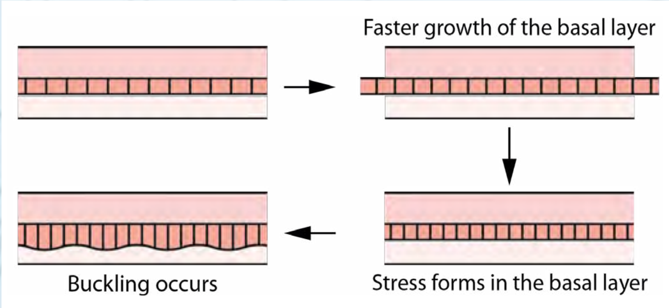


Figure 2 Schematic diagram of the folding hypothesis [3]. The three vertical layers are epidermis (top), basal layer (middle) and dermis (bottom), respectively.

Figure 3 Ridge formation through the combination of the centers of cell proliferation [2].

[2, 3]. Cell proliferation continues at those sites while the folds combine and merge into clusters to form linear ridges in a rather random fashion, creating the unique pattern of wrinkles in our fingerprints (Figure 3) [2].

Volar pads are well known to play a role in determining fingerprint pattern. They are transient tissue swellings present on certain areas of our palms, including each fingertip, during embryonic development (embryogenesis) (Figure 4). Coinciding with the process of ridge formation, these structures start shrinking from the 10th week [1]. The shrinkage introduces extra mechanical stresses across the skin, affecting the directions of ridge formation [1]. Scientists generally agreed that the height and size of volar pads can influence the pattern of fingerprints [2, 4]. Whorl-type patterns are

Figure 4 Volar pads on the palm during embryonic development are highlighted in light gray.

usually formed on high volar pads whereas low volar pads produce the arch-type pattern. Intermediate-height volar pads create the loop-type pattern.

Then, how do genetic factors come into play? It was illustrated that the geometry of volar pads can be controlled by genes [1]. For example, *EVI1*, a limb development gene responsible for the outgrowth of distal limbs and digits, was found to express under the volar pads. It was hypothesized to influence fingerprint pattern by controlling the shape and size of the volar pads through its function to promote cell proliferation, as in the distal ends of the developing limbs [4]. This also provides insight into the correlation between fingerprint pattern and limb-related phenotypes [4]. Researchers discovered that individuals with whorl patterns on both pinky fingers often have a longer pinky finger compared to those with no whorl pattern on pinky fingers; the frequency of whorl patterns on the fingers of both hands (except the thumbs) is also associated with longer pinky fingers [4].

Exception: The Family with No fingerprints

We have often taken fingerprint technology for personal identification for granted. Fingerprints are a huge part of our identity in modern society, with applications in mobile phones and immigration. Nevertheless, Apu Sarker's family in Bangladesh has no fingerprints due to a rare genetic mutation in the *SMARCAD1* gene, causing Adermatoglyphia or "Delayed Immigration Disease" [5]. Luckily, it does not cause any serious illnesses, but the family encountered difficulties in their everyday lives because fingerprints became mandatory when obtaining driving licenses, sim cards, and passports. As a result, they could not obtain a driving license nor purchase a sim card for their mobile phones. In Apu's ID card when he was still 10, he was labeled with "NO FINGERPRINT" as the government officials had no idea on how to issue a card without the means for personal identification. With the advent of modern technologies, such as iris scan and facial recognition, let us all hope those who have such genetic conditions would not unintentionally be discriminated against in the near future.

Interesting Fact: Do Monozygotic Twins Share the Same Fingerprints?

Have you ever wondered whether identical twins share identical fingerprints? Although they are very similar in appearance and contain the same DNA sequence, they have slightly different fingerprints that are significant enough to be captured by the modern recognition software [6]. Aside from the randomness in the fingerprint formation process, small differences in umbilical cord length, womb position, blood pressure, nutritional intake, and the rate of finger growth during the 13th to 19th week can still influence fingerprint formation as a result [6]! This highlights the importance of how environmental factors apart from genetics also play an essential role in determining our fingerprints.

Conclusion

Our fingerprints are an important physical trait that can define who we are as individuals. The unique pattern formed during embryogenesis persists throughout our lives from the 19th week of gestation [3] and remains the same even after we die (until decomposition occurs). Its formation is dependent on the interaction between genetic and variable environmental factors, which ultimately give rise to our unique fingerprints: the key to our identity and individuality.

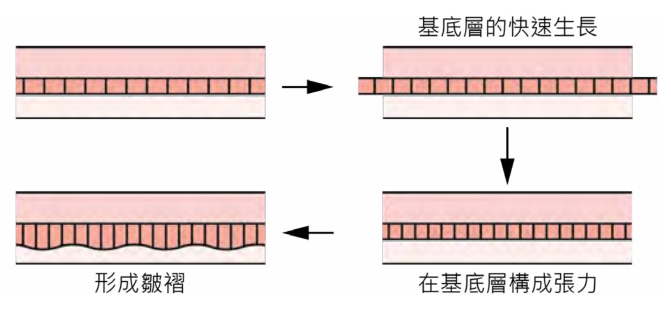
簡介

在使用血液測試和 DNA 分析這些複雜方法之前，我們早已有一套簡單直接的方法來辨別身分，那就是我們經常在電影中看到或是通關會用到的指紋分析。指紋又稱為「皮嵴」(friction ridge skin)，像我們的容貌一樣，它是我們獨一無二的特徵。可是你有想過為甚麼每個人的指紋都不同嗎？答案在於基因(尤其是控制四肢發育的基因)和環境間的相互作用，它們創造出能被分成三個類別的獨特皮紋圖樣(dermatoglyphic patterns)：弧紋(arch)、箕紋(loop)和斗紋(whorl)(圖一)。

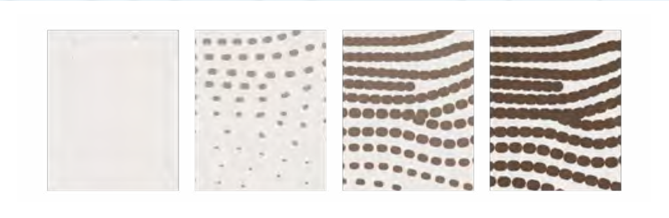
圖一 指紋的三個類別：弧紋、箕紋和斗紋

指紋的形成

科學界對指紋形成有著不同理論，但皮膚學家傾向相信摺疊假說 [1]。皮膚組織由緊密連接的三層垂直結構組成，分別是表皮、基底層和真皮。頂層的表皮和底層的真皮細胞生長速度的差別在快速生長的基底層構成張力，使其在不同位置摺疊以釋放張力(圖二)[2, 3]。細胞繼續在摺疊的地方增生的同時，不同拱起的摺疊會頗為隨機地集結成群，形成一條條皮嵴，產生指紋中獨特的皺褶圖案(圖三)[2]。



圖二 摺疊假說的示意圖 [3]；三層垂直的結構分別為表皮(上)、基底層(中)及真皮(下)。



圖三 細胞增生中心聚集而形成皮嵴的過程 [2]

已知掌墊(volar pads)會影響指紋的圖案。掌墊是胚胎發育(學術上稱為「胚胎發生」(embryogenesis))期間在包括指尖在內的手掌特定位置暫時形成的凸起組織(圖四)。與皮嵴形成的時間吻合，這些結構會在第十週縮小 [1]，對皮膚產生額外的機械應力，從而影響皮嵴形成的方向 [1]。科學家普遍同意掌墊的高度和大小會影響指紋的樣式 [2, 4]：斗紋通常會在高掌墊上形成，低掌墊會產生弧紋，而中高掌墊則會形成箕紋。

圖四 淺灰色部分為胚胎發育期間手掌上的掌墊

那麼，遺傳因素又怎樣影響指紋呢？科學家已證明掌墊的形狀和大小受基因控制 [1]，例如負責四肢和指頭生長的肢體發育基因 *EVI1* 已知會在掌墊下表達，科學家猜想 *EVI1* 會透過其促進細胞增生的功能影響掌墊的形狀和大小，就像刺激發育中的肢體藉細胞分裂向外延長一樣 [4]。這提供了一些啟示解釋為甚麼指紋樣式會與肢體相關表現型有所關聯 [4]：研究人員發現兩隻尾指皆為斗紋的人，通常比兩隻尾指皆非斗紋的人擁有更長的尾指，而雙手指頭（除姆指外）擁有斗紋的數目亦被發現與尾指的長度有關 [4]。

例外：沒有指紋的家族

我們早已習慣把指紋技術用於確定身分。指紋是現今社會身分認證的重要一環，在手機和出入境上的應用最為人熟知。可是，在孟加拉的 Apu Sarker 一家卻因罕見的 *SMARCAD1* 基因突變導致沒有指紋，這種狀況稱為皮紋病（Adermatoglyphia），又被戲稱為「入境阻延症」（Delayed Immigration Disease）[5]。不幸中 之大幸是基因突變沒有帶來嚴重的病痛，但他們一家卻在生活上遇到重重困難，因為指紋早已成為取得駕駛執照、電話卡和護照時必須提交的資訊，故此他們並不能取得駕駛執照和購買電話卡。Apu 在十歲申領的身分證上被標註為「沒有指紋」，因為政府官員對如何發出一張不能認證身分的身分證毫無頭緒。然而隨著虹膜掃描及人面識別等更多現代技術的出現，我們可以寄望不久將來這些受基因問題困擾的人不會再受到社會中無意的不公平對待。

知多一點點：同卵雙生兒有相同的指紋嗎？

你有想過究竟同卵雙生兒會有相同的指紋嗎？雖然他們的外貌非常相似，也有著相同的 DNA 序列，但是他們指紋的微細分別還是足以被現今的辨識軟件區分 [6]。除了指紋形成過程中的隨機性，原來在第 13 至 19 週臍帶長度、子宮位置、血壓、營養攝取和手指生長速度上的微細分別還是會影響到指紋的形成 [6]！這反映了基因以外環境因素如何在決定指紋樣式上發揮重要作用。

結語

指紋是能夠判別我們身分的身體特徵。從胚胎發生過程產生、獨一無二的指紋圖樣由懷孕第 19 週 [3] 到生命終結之時都從不改變；它的形成受基因和多變的環境因素影響，最終發展成獨有形態，創造出這個能辨認我們身分的重要特徵。

References 參考資料：

[1] Kücken, M. (2007). Models for fingerprint pattern formation. *Forensic Science International, 171*(2-3), 85–96. *https://doi.org/10.1016/j.forsciint.2007.02.025*

[2] Wertheim, K. (2011). Embryology and Morphology of Friction Ridge Skin. In A. McRoberts (Ed.), *Fingerprint Sourcebook* (pp. 3-1–3-26). National Institute of Justice. *https://www.ojp.gov/pdffiles1/nij/225323.pdf*

[3] Garzón-Alvarado, D. A., & Ramírez Martinez, A. M. (2011). A biochemical hypothesis on the formation of fingerprints using a turing patterns approach. *Theoretical Biology and Medical Modelling, 8*, 24. *https://doi.org/10.1186/1742-4682-8-24*

[4] Li, J., Glover, J. D., Zhang, H., Peng, M., Tan, J., Mallick, C. B., Hou, D., Yang, Y., Wu, S., Liu, Y., Peng, Q., Zheng, S. C., Crosse, E. I., Medvinsky, A., Anderson, R. A., Brown, H., Yuan, Z., Zhou, S., Xu, Y., . . . Wang, S. (2022). Limb development genes underlie variation in human fingerprint patterns. *Cell, 185*(1), 95–112. *https://doi.org/10.1016/j.cell.2021.12.008*

[5] Sabbir, M. (2020, December 26). The family with no fingerprints. *BBC News. https://www.bbc.com/news/world-asia-55301200*

[6] Asher, C. (2021, September 18). Why do identical twins have different fingerprints? *BBC Science Focus Magazine. https://www.sciencefocus.com/the-human-body/why-do-identical-twins-have-different-fingerprints-2/*

Demystifying Wordle:
A Crash Course in Information Theory
Wordle大揭秘：資訊理論101

By Sonia Choy 蔡蒨珩

The game Wordle took the world by storm last year – you might have seen your friends posting green and yellow boxes on social media, claiming that they have solved this daily word puzzle in three guesses, or that dreaded "X/6," which means that they didn't manage to crack it. When one considers what first word to guess, it might be tempting to randomly put a five-letter word at the beginning, but this can actually be reduced to a scientific question. It is not hard to see that some words would be a better first guess than others; for example, the word "FUZZY" would be far less ideal than "RAISE", since the letters in the former occur far less often than the letters in the latter. What, then, is one's best shot at cracking the puzzle?

What Is Wordle?

We are assuming readers know how Wordle functions. For those who do not, here is a quick crash course.

Wordle's database is made of 2,315 five-letter words picked by the creator of the game as solutions, and a pool of approximately 13,000 five-letter words that are valid guesses (which include the 2,315 words above, and many more words that are not commonly used) [1]. Each day, one word from the database is selected to be the answer to the puzzle. If your guess has a letter that is in the word and in the same position, the letter box shown will be green; if the guess has a letter that is in the word but not in the correct position, the letter box shown will be yellow; otherwise the box is gray.

How to Define "Informative"?

To give a satisfying answer, we first need to quantify what is meant by "more useful". A "useful" guess gives us more information; but how do we quantify information? Luckily for us, this was done in the 1940s by Claude Shannon, the father of information theory.

Shannon defined information by the following equation: $I = log_2(\frac{1}{p})$, where p is the probability of the event happening. You may ask, why the logarithm function? Recall a property of the logarithm from high school: $log(a) + log(b) = log(ab)$. If we have two independent events happening each with a probability of p_1 and p_2 respectively, then the probability of them occurring together is p_1p_2, or:

$$I_1 + I_2 = log_2(\frac{1}{p_1}) + log_2(\frac{1}{p_2}) = log_2(\frac{1}{p_1p_2}) = I_{12}$$

So the multiplicity of probability is captured in the amount of information it gives. Information is typically measured in bits; in the case of Wordle, it basically means how many times a word can reduce the number of possible choices into halves.

It is rather unlikely for "FUZZY" to be the first hit. Suppose it returns five gray squares – what information do these squares give us?

Using the above two first guesses ("FUZZY" and "RAISE") as an example, the probability of "F" occurring in an English word is approximately 2.2% (Table 1) [2], so the probability that it does not occur is 97.8% or 0.978. We can find out the probabilities for each letter in our guesses, and decide that the combined information "FUZZY" gives is 0.093 bits (footnote 1):

$$I = log_2(\frac{1}{0.986}) + log_2(\frac{1}{0.972}) + log_2(\frac{1}{0.99926}) + log_2(\frac{1}{0.99926}) + log_2(\frac{1}{0.980}) = 0.093$$

What if "RAISE" turns out to have all five gray guesses? We have:

$$I = log_2(\frac{1}{0.940}) + log_2(\frac{1}{0.918}) + log_2(\frac{1}{0.930}) + log_2(\frac{1}{0.937}) + log_2(\frac{1}{0.87}) = 0.61$$

Letter	Relative Frequency	Letter	Relative Frequency
F	2.2%	R	6.0%
U	2.8%	A	8.2%
Z	0.074%	I	7.0%
Y	2.0%	S	6.3%
		E	13%

Table 1 Selected relative letter frequencies in English language (footnote 2) [2].

Note that the more unlikely an event is, the more information is provided if it occurs. For example, the letter E is more unlikely to be absent in a word than the letter Z, so the missing of the letter E can narrow our search down very much. In other words, the corresponding information a gray square E gives is more than a gray square Z after a round of guessing. This is why obtaining five gray squares when guessing "RAISE" would be more informative than obtaining five gray squares after guessing "FUZZY".

Shannon Entropy and the Information Provided by a Guess

The calculation above gave us a sense of how informative a guess could be if we unluckily got five gray squares. However, in some cases, we may get a mixture of gray, yellow and green squares, say gray-yellow-gray-gray-gray, or even green-yellow-gray-gray-green. By analyzing the list of 2,315 answers, for a single guess, we can come up the exact probability of getting each possible pattern, and its corresponding amount of information under that pattern.

Taking all cases into account, we can calculate the weighted average of all information given by a word. This average is often called Shannon entropy, which is not directly related to the entropy in physics. Then it is possible to rank all words by the information it provides in the first guess. This has been done by multiple people, including the YouTuber Grant Sanderson (3Blue1Brown).

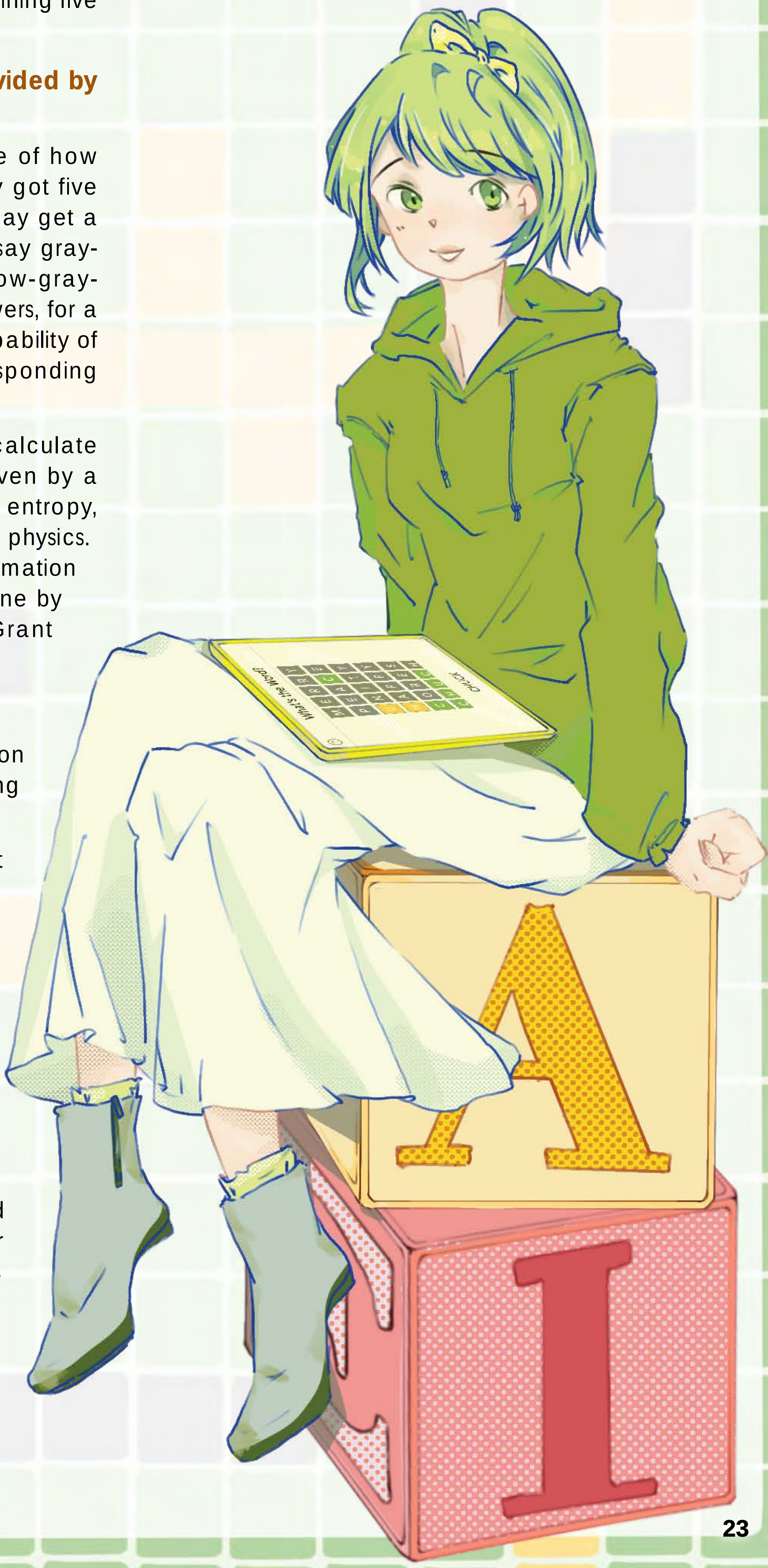
What Are the Best First Guesses?

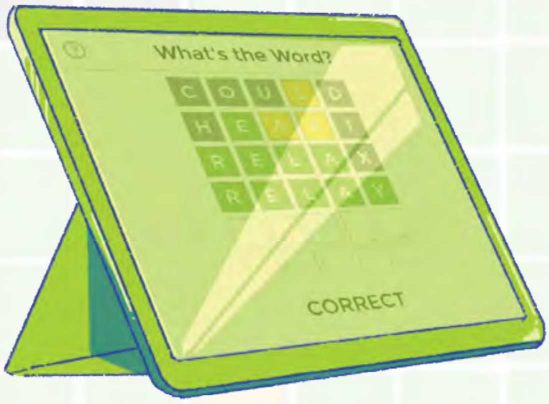
The first guess that gives the most information is "SOARE" (5.89 bits), an obsolete term meaning a young hawk [3].

Can we do better? We can also look at the next few guesses. Wordle is not just a one-guess game; your next guesses also matter, and it can be useful to see how the next guess plays out when you use your first word. By also considering the average information obtained from an optimal second guess, we get that the best first guess is "SLANE", a type of spade in Ireland, giving an average of 10.04 bits in the first two guesses [3].

Some readers may also be concerned about winning Wordle with a minimal number of guesses. With the top 250 first guesses generated by considering the first two guesses, researchers ran a simulation to find out the actual performance of these guesses in the 2,315 games. They found that, "SALET", meaning a medieval helmet, is the winner, with the computer winning the game in an average of 3.412 times out of six [4].

But truth be told, the first few contenders are a close race. If you are looking for a word for your next first guess that isn't too obscure, "CRATE" is a good choice that is not too far behind the above list of obscure words, giving 10.01 bits of information in the first two guesses, and passing with 3.434 guesses on





average [3]. We are not suggesting that you recite the list of possible solutions and analyze every move like a computer, but having a good first guess should be a good way to start off the day with a good shot at a puzzle.

---

1 Editor's remark: The second letter Z can indeed provide extra information. If both your guess and the answer contain two identical letters, say "FUZZY" and "WHIZZ", both boxes of Z's will turn yellow and/or green to confirm that the answer contains two Z's.

2 Editor's remark: The source of text (e.g. general documents, dictionary) can affect the values. Table 1 shows the frequencies of letters that appear in English documents of all types [2]; a more accurate way of estimating the probability in our case is to check the probability of each letter appearing in each slot according to the list of words in the *New York Times* code. For simplicity's sake, values concerning the general English language were used here, but the more accurate way has already been done by the YouTuber Grant Sanderson (3blue1brown) [3].

上年網上遊戲 Wordle 的風潮席捲全球，大家的社交媒體好一陣子都被綠色、黃色和灰色的方格佔據。你可能見過朋友誇耀自己只猜了三次就已解開當天謎題，也有時會沮喪地抱怨當天只能以「X/6」飲恨作結，未能揭開當天謎底。當我們思考應用甚麼單字開始遊戲時，我們也許會很隨性地把一個五個字母長的任意單字填進空格，但其實這其實是一個科學問題。我們不難意識到有些單字比別的好，對我們稍後的猜測更為有利，譬如選擇「FUZZY」遠比「RAISE」不智，因為前者包括的字母遠比後者冷門。那麼，哪個單字才是能幫我們贏出遊戲的最佳選擇呢？

甚麼是 Wordle？

我們假設讀者知道 Wordle 的運作方式，但如果你不知道的話，以下是懶人包。

Wordle 的數據庫包含 2,315 個由開發者挑選，五個字母長作為答案的單字，也有大概 13,000 個被系統視為存在的有效單字（裡面除了 2,315 個作為答案的單字外，還包括許多不常用的單字）[1]。每天程式都會從數據庫抽取一個單字作為當天遊戲的答案。如果你猜的單字包含答案裡有的字母，而該字母位於正確的位置上，該方格將會變為綠色；如字母正確，但位置錯誤，該方格將變為黃色；如答案並沒有包含該字母，方格則會變為灰色。

有用？無用？

解答以上問題前，我們必先量化「有用」這個指標。「有用」的猜測能給我們更多資訊，但我們怎樣量化資訊呢？資訊理論之父 Claude Shannon 早已在 1940 年代替我們解決這個問題。

Shannon 用以下公式定義資訊：$I = log_2(\frac{1}{p})$，p 為事情發生的概率。你可能會問：為甚麼會是對數函數（logarithm function）呢？回想一下高中時我們學過的對數定律：$log(a) + log(b) = log(ab)$，如果有兩件獨立事件，兩者發生的概率分別為 p_1 和 p_2，那麼他們兩者都發生的概率會是 p_1p_2，或是：$I_1 + I_2 = log_2(\frac{1}{p_1}) + log_2(\frac{1}{p_2}) = log_2(\frac{1}{p_1p_2}) = I_{12}$

事件發生概率的相乘結構就此保存於公式給出的資訊量裡面。資訊的單位為位元（bit）；在 Wordle 裡意味著一個單字能把搜尋範圍縮小一半的次數。

「FUZZY」不太可能使我們一擊即中。假如它變出五個灰色方格的話，這些方格給予我們甚麼資訊？

取以上兩個我們用作第一次猜測的單字（「FUZZY」和「RAISE」）作為例子，「F」在英文單字裡出現的概率約為 2.2%（表一）[2]，因此它沒有出現的概率為 97.8% 或 0.978。我們可以找出單字中各個字母沒有出現的概率，然後計算出「FUZZY」的總資訊量為 0.093 位元（註一）：

$$I = log_2(\frac{1}{0.986}) + log_2(\frac{1}{0.972}) + log_2(\frac{1}{0.99926}) + log_2(\frac{1}{0.99926}) + log_2(\frac{1}{0.980}) = 0.093$$

那如果「RAISE」給出五個灰色方格呢？資訊量為：

$$I = log_2(\frac{1}{0.940}) + log_2(\frac{1}{0.918}) + log_2(\frac{1}{0.930}) + log_2(\frac{1}{0.937}) + log_2(\frac{1}{0.87}) = 0.61$$

字母	出現頻率	字母	出現頻率
F	2.2%	R	6.0%
U	2.8%	A	8.2%
Z	0.074%	I	7.0%
Y	2.0%	S	6.3%
		E	13%

表一 英語裡一些字母的出現頻率（註二）[2]

越不可能發生的事件，一旦發生的話能提供的資訊量就越高。譬如說，一個單字裡可能沒有 Z，但並不太可能沒有 E，因此欠缺 E 將能大大縮小我們的搜索範圍。換言之，E 變成灰色方格給予我們的資訊比 Z 變成灰色方格多，因此在同樣取得五個灰色方格的情況下，猜「RAISE」的資訊量比猜「FUZZY」的多。

Shannon 熵和單字提供的資訊

以上計算告訴我們猜一個單字時不幸得到五個灰色方格的資訊量，可是我們亦有機會得到不同顏色組合的方格，譬如「灰黃灰灰灰」或是「綠黃灰灰綠」。透過利用全部 2,315 個答案作分析，對於每個輸入的單字，我們都可以精確計算出得到每個可能顏色組合的概率，以及其相應的資訊量。

考慮所有可能情境後，我們可以計算出一個單字在不同顏色組合下給予我們資訊量的加權平均值。這個平均值名為 Shannon 熵 （Shannon entropy），但與物理學的熵（entropy）並沒有直接關係。有了這指標後，我們就能將所有單字以在第一回合裡可提供的資訊量排序。很多人已經透過這方法分析過 Wordle，包括著名的數學 YouTube 頻道主持 Grant Sanderson（3Blue1Brown）。

第一回合應該猜甚麼？

第一回合能提供最多資訊的單字是「SOARE」（5.89 位元），一個過時的用字，意思是年幼的鷹 [3]。

有更好的選擇嗎？我們可以考慮接下來的幾回合，畢竟 Wordle 是個多回合的遊戲，下回合的選擇也很關鍵，因此其中一個有用的思考角度是考慮第一回合選擇的單字如何影響下一回合。假設我們第二回合有能力作出最合適的選擇，在考慮兩回合選擇給出的平均資訊量後，就能計算出第一回合的最佳選擇為「SLANE」，意指愛爾蘭一種園藝用的鏟，首兩回合平均能提供 10.04 位元的資訊 [3]。

有些讀者可能想以最少的回合勝出 Wordle。研究人員在考慮不同第一回合選擇在首兩回合的表現後，篩選了最佳的 250 個第一回合選擇，然後再用這些選擇各自執行全部 2,315 場遊戲，以找出每個選擇在遊戲中的實際表現。研究人員發現「SALET」的表現最佳，平均只需 3.412 回合就能夠勝出遊戲，而「SALET」是一款中世紀的頭盔 [4]。

但實話實說，頭幾名的表現確實不相伯仲。如果你不想以太艱澀的字眼開始遊戲，「CRATE」會是一個好選擇，它在首兩回合能提供 10.01 位元的資訊，平均以 3.434 回合就能贏出遊戲，不會比以上提及的艱深字眼相差太遠 [3]。最後，我們並不是建議你背誦所有可能的答案，然後像電腦般分析每一步，但一個明智的選擇也許能幫你以一局美好的遊戲，開展一個美好的早晨。

1 編按：第二個 Z 其實也能提供額外資訊；如果猜的單字跟答案均有兩個相同的字母，例如「FUZZY」和「WHIZZ」，兩個 Z 的方格都會變為黃色或綠色以表示答案包含兩個 Z。

2 編按：數值受文本來源（例如一般文件、字典等）影響。表一列出了字母在各類型英文文件中出現的頻率 [2]，但其實更準確估算概率的方法，是找出並採用每個字母出現於《紐約時報》答案列表單字中每一字元的頻率。這裡為了簡單說明概念而採用了一般英語裡的數值，但較準確的估算已經由 YouTube 頻道主持 Grant Sanderson（3Blue1Brown）實驗過 [3]。

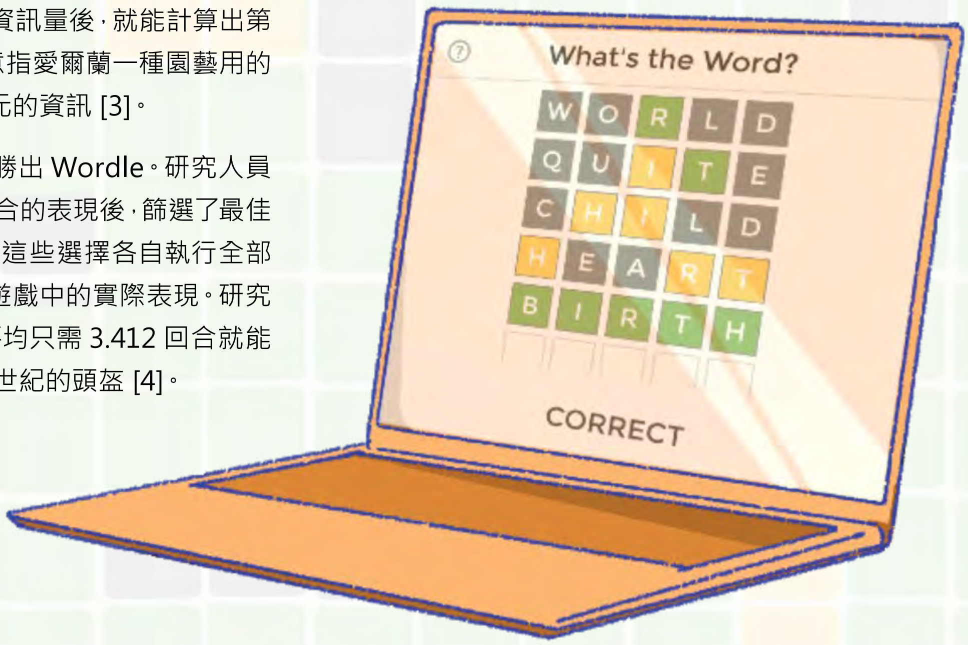
References 參考資料：

[1] Glaiel, T. (2021, December 30). *The mathematically optimal first guess in Wordle*. Medium. *https://medium.com/@tglaiel/the-mathematically-optimal-first-guess-in-wordle-cbcb03c19b0a*

[2] Lewand, R. E. (2000). *Cryptological Mathematics*. MAA Press.

[3] Sanderson, G. (2022). [3Blue1Brown Wordle video source files] [Source code]. GitHub. *https://github.com/3b1b/videos/tree/master/_2022/wordle*

[4] Selby, A. (2022, January 19). *The best strategies for wordle*. *https://sonorouschocolate.com/notes/index.php?title=The_best_strategies_for_Wordle*



Let's Be Social!

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

Fun Facts

Memes

SCIENCEFOCUS.HKUST

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