

給機器人的最短說明書

本課講一本 2017 年出版的數學書。作者是 Shen、Uspensky、Vereshchagin 三位數學家，他們來自 Kolmogorov 在莫斯科的學派。書裡講：一串東西有多少信息，就看寫出它的最短說明書有多長。

適合 10 歲左右

約 40 分鐘

需要：鉛筆、一枚硬幣、兩位同伴

1 熱身：給機器人傳話

有一個死板的機器人，只認兩種寫法：① 直接寫 0 和 1，它照著抄；② $(一段) \times 幾$ ，它把括號裡的抄幾遍。比如 $(01) \times 3$ 就是 010101。數長度時，**每個符號算 1 個字**，括號和 $\times$ 也算。 $(01) \times 3$ 一共 6 個字。

請給下面每串寫一份說明書，越短越好。每串原來都是 16 個字。

0000000000000000 $\rightarrow$ _____ (____ 字)

0101010101010101 $\rightarrow$ _____ (____ 字)

0011001100110011 $\rightarrow$ _____ (____ 字)

1111000010100111 $\rightarrow$ _____ (____ 字)

前 3 串很好辦，比如第 1 串寫成 $(0) \times 16$ ，只要 6 個字。最後一串是擲硬幣得來的（示意）。把開頭寫成 $(1) \times 4$ ，反而多了 1 個字！它好像只能整串照抄。這就是今天的主題。

2 最短說明書有多長

電腦**程序**就是寫給電腦的一步一步的命令，這一課叫它「給機器人的說明書」。機器人照著做，必須**一字不差**地寫出原來那串。同一串東西，可以有很多份說明書。**最短的那份有多長**，就叫這串東西的**Kolmogorov 複雜度**，中文簡稱**柯氏複雜度**。它量的是：這串東西里到底裝了多少信息。

一串東西（都是 1000 個字）	一份短說明書	說明書幾個字
「01」重複 500 次	$(01) \times 500$	8
擲 1000 次硬幣，記下 0 和 1	找不到規律，每一位都是新消息，只能照抄	差不多 1000

和前面幾課連起來：第 4、5、6、7、23 課都在講壓縮。這一課是「壓縮」路線的終點：**最短到底能有多短？**這個想法在 1960 年代提出，Kolmogorov、Solomonoff 和 Chaitin 各自想到了它。這本書大約 500 頁，作者之一 Uspensky 是 Kolmogorov 的學生。

3 寫不短，就是隨機

第 5 課講過：擲硬幣的結果沒有規律，壓縮不了。這本書把這句話變成了「隨機」的定義：

如果一串東西**沒有比它自己短多少的說明書**，它就是**隨機的**。

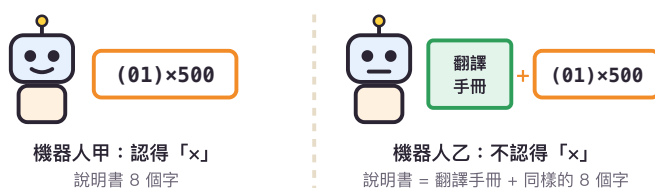
這個定義只看串本身，不管它怎麼來的。寫不短的串多不多？電腦程序本來就是一串 0 和 1，我們假設說明書也只用 0 和 1 寫，數一數。



一份說明書只能寫出一串。8 個串只有 7 份更短的說明書，所以**至少有 1 串寫不短**。串越長，差距越大：1000 位長的串裡，能縮短 10 個字以上的**不到千分之一**。所以大多數串都是隨機的。

4 換一種語言，差別不大

換一種電腦語言寫說明書，長度會變嗎？會變，可是**變不了多少**。



機器人乙不認得「x」。沒關係，在說明書前面附一本**翻譯手冊**，教它「x 幾就是抄幾遍」，後面照樣寫 $(01) \times 500$ 。關鍵是：**手冊長度固定**，不管寫哪一串，附的都是同一本。所以給乙的最短說明書，最多比給甲的長一本手冊。數學家把這叫**不變性定理**：換一種電腦語言，柯氏複雜度最多差一個固定的數。這個數只看兩種語言，和要寫的串無關。

小心：手冊可能比短串本身還長，所以只比很短的串沒有意義。比長串時，手冊那點長度才可以不管。

5 驚人的事實：最短說明書算不出來

想像一場「最短說明書比賽」：第 1 天，小明寫了 30 個字。第 2 天，小紅找到規律，寫了 20 個字。第 3 天，小剛找到更深的規律，寫了 12 個字。第 4 天呢？

裁判能說「12 個字比 20 個字短」，卻**沒法宣佈冠軍**：也許明天有人找到更短的。書裡證明了：**沒有任何一個程序，能對所有的串都算出最短說明書有多長**。數學家說柯氏複雜度**不可計算**。我們只能越找越短，卻不能確定哪份最短。

為什麼不把短說明書一份份試過去？有的說明書會讓機器人一直算，你不知道它是快算完了，還是永遠不停。還有一個更深的原因。讀一讀這句話：「**用二十個字說不出來的最小的數**」（數指 1、2、3 這樣的數）。它只有 14 個字，卻說出了一個「二十個字說不出來」的數，自己跟自己打架！這叫**貝里悖論**。

「悖論」就是怎麼想都說不通的話。

同樣的道理：測量機不可能存在

假如有一臺**測量機**，給它任何一串，它都能說出最短說明書有多長。那我就能寫一份說明書：

「用測量機按順序一串一串地查，找出第一串最短說明書超過一億字的，把它寫出來。」

第 3 節講過，這樣的串一定有。測量機也寫進去，整份說明書也遠遠不到一億字。可它寫出了那一串，那一串的最短說明書就不到一億字了。說不通！所以測量機不可能存在。

6 實際中：用壓縮軟件估一估

算不出最短，怎麼辦？第 6、7 課的科學家用了**壓縮軟件**，就是電腦裡把文件變小、還能一字不差還原的工具。小文件加上「解壓程序」，就是一份說明書。真正的最短**只會更短或一樣長**，所以壓縮後的大小可以粗略估計信息有多少。

壓縮軟件也會看走眼。圓周率 3.14159..... 的前 100 萬位，壓縮軟件幾乎壓不動。可是一個很短的程序就能把它們一位一位算出來。所以「壓不動」不等於「真的隨機」。

7 這和今天有什麼關係

第 5 課講過：朋友猜得越準，你傳話用的字越少。**會預測，就會壓縮**。柯氏複雜度就是壓縮能達到的「最好成績」。

ChatGPT、Claude 這些聊天 AI 靠的是大語言模型，它們靠「猜下一個字」練出來（第 1 課）。2023 年，DeepMind 公司的論文發現：大語言模型配上一種編碼方法，就能當壓縮軟件。這些模型只讀過文字，可它們壓縮圖片小塊和聲音，比 PNG（常用的圖片格式）和 FLAC（常用的聲音格式）還壓得小。第 3、8 課的作者之一 Sutskever 多次說過：**好的預測器就是好的壓縮器**。這本書就在他推薦的閱讀清單裡。

也要說清楚：柯氏複雜度算不出來，沒人直接拿它訓練 AI。模型壓出的長度也只是「目前找到的一份說明書」。

8 動手：給機器人寫最短說明書

玩法（2—4 個人，約 12 分鐘）

1. 用熱身的機器人規則，各自給每串寫說明書。括號裡可以再套括號。「×幾」後面不要緊跟 0 或 1，不然看不清。也可以用珠子擺：紅是 0，藍是 1。
2. 同伴照著說明書寫回原串，一字不差才打 ✓。誰最短，誰贏這一串。最後一行填自己擲 20 次硬幣的結果：正面 1，反面 0。

串（每串 20 個字）	我的說明書	幾個字	✓
00000000000000000000			
00110011001100110011			
01101101100110110110			
01000010110010000010			
我擲的硬幣：			

你會發現：前 3 串能寫得很短。第 3 串要套括號：「0，再加 3 遍 110」這一段重複 2 次。第 4 串和你擲的那串，怎麼寫都短不了幾個字。就算你贏了，也不能保證沒人寫得更短。

9 動手：誰是隨機的

玩法（3 個人，約 12 分鐘）

A 真擲 20 次硬幣，正面寫 1，反面寫 0。B 不碰硬幣，憑腦子編 20 個 0 和 1，儘量像真擲的。兩人把串抄到紙條上，不寫名字，交給 C。C 用 3 把尺子檢查，猜哪張是真擲的。揭曉後換人再玩。

	尺子 1：最多連著幾個一樣	尺子 2：0、1 換了幾次	尺子 3：能寫短嗎	猜真的
紙條甲				
紙條乙				

你會發現：真擲 20 次硬幣，大約 4 次裡有 3 次會出現連著 4 個或更多一樣的。0 和 1 平均換 9 到 10 次。人編的串常常換得太勤，很少連著 4 個一樣。「幾乎不連著一樣」也是一條規律：想裝隨機，反而露出了規律。

10 詞語小卡片

柯氏複雜度

寫出一串東西的最短說明書有多長，量的是信息有多少。

隨機的串

沒有比它自己短多少的說明書。大多數串都是。

不變性定理

換一種電腦語言，最短說明書最多差一本「翻譯手冊」的長度。

不可計算

沒有程序能對所有串算出最短說明書的長度。

11 想一想

1. 小明真的擲了 20 次硬幣，結果全是正面。按這一課的定義，這串算不算「隨機的串」？為什麼？

2. 小剛說：「我給這串寫了 15 個字的說明書。我敢保證，世界上沒有更短的。」你相信嗎？為什麼？

3. 聊天 AI 很會猜下一個字。為什麼說它也能當壓縮軟件用？

一句話記住：最短說明書有多長，信息就有多少。寫不短的，就是隨機的。

給家長和老師：本課講的是 A. Shen、V. A. Uspensky、N. Vereshchagin 的書 [Kolmogorov Complexity and Algorithmic Randomness](#) (AMS Mathematical Surveys and Monographs 第 220 卷，2017)。講義只涉及前幾章：C(x) 的定義、不變性定理、大多數串不可壓縮、C(x) 不可計算（貝里悖論式證明）和「不可壓縮即隨機」的直覺；無限序列的 Martin-Löf 隨機與前綴複雜度未展開。

「機器人語言」是為孩子設計的簡化描述方式，不是通用編程語言；書中程序是固定通用機器上的二進制串。各串、比賽、翻譯手冊和測量機的说法都是示意。「擲 20 次約 77% 出現 ≥ 4 的連串、平均換 9.5 次」是我們按公平硬幣算的；「人編的串換得太勤」是常見課堂現象，非書中內容。DeepMind 論文為 Language Modeling Is Compression (2023)。想深入可看本站的 [第 26 天中文導讀](#)，代碼見開源項目 [30u30 第 26 天目錄](#) (英文)。