

给机器人的最短说明书

本课讲一本 2017 年出版的数学书。作者是 Shen、Uspensky、Vereshchagin 三位数学家，他们来自 Kolmogorov 在莫斯科的学派。书里讲：一串东西有多少信息，就看写出它的最短说明书有多长。

适合 10 岁左右

约 40 分钟

需要：铅笔、一枚硬币、两位同伴

1 热身：给机器人传话

有一个死板的机器人，只认两种写法：① 直接写 0 和 1，它照着抄；② $(\text{一段}) \times \text{几}$ ，它把括号里的抄几遍。比如 $(01) \times 3$ 就是 010101。数长度时，**每个符号算 1 个字**，括号和 $\times$ 也算。 $(01) \times 3$ 一共 6 个字。

请给下面每串写一份说明书，越短越好。每串原来都是 16 个字。

0000000000000000 $\rightarrow$ _____ (____ 字)

0101010101010101 $\rightarrow$ _____ (____ 字)

0011001100110011 $\rightarrow$ _____ (____ 字)

1111000010100111 $\rightarrow$ _____ (____ 字)

前 3 串很好办，比如第 1 串写成 $(0) \times 16$ ，只要 6 个字。最后一串是掷硬币得来的（示意）。把开头写成 $(1) \times 4$ ，反而多了 1 个字！它好像只能整串照抄。这就是今天的主题。

2 最短说明书有多长

电脑**程序**就是写给电脑的一步一步的命令，这一课叫它“给机器人的说明书”。机器人照着做，必须**一字不差**地写出原来那串。同一串东西，可以有很多份说明书。**最短的那份有多长**，就叫这串东西的**Kolmogorov 复杂度**，中文简称**柯氏复杂度**。它量的是：这串东西里到底装了多少信息。

一串东西（都是 1000 个字）	一份短说明书	说明书几个字
“01”重复 500 次	$(01) \times 500$	8
掷 1000 次硬币，记下 0 和 1	找不到规律，每一位都是新消息，只能照抄	差不多 1000

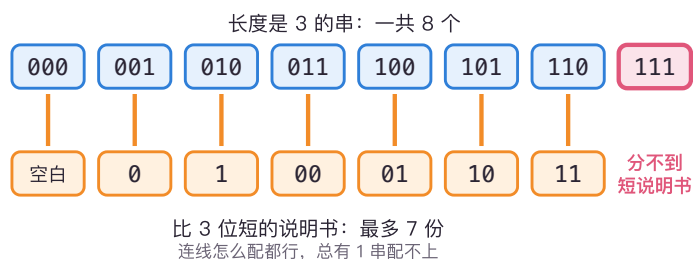
和前面几课连起来：第 4、5、6、7、23 课都在讲压缩。这一课是“压缩”路线的终点：**最短到底有多短？** 这个想法在 1960 年代提出，Kolmogorov、Solomonoff 和 Chaitin 各自想到了它。这本书大约 500 页，作者之一 Uspensky 是 Kolmogorov 的学生。

3 写不短，就是随机

第 5 课讲过：掷硬币的结果没有规律，压缩不了。这本书把这句话变成了“随机”的定义：

如果一串东西**没有比它自己短多少的说明书**，它就是**随机的**。

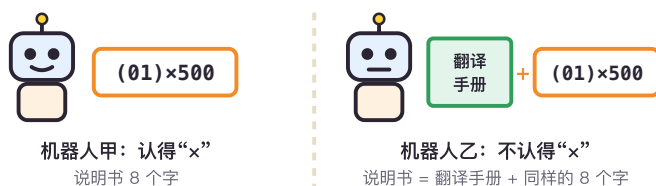
这个定义只看串本身，不管它怎么来的。写不短的串多不多？电脑程序本来就是一串 0 和 1，我们假设说明书也只用 0 和 1 写，数一数。



一份说明书只能写出一串。8 个串只有 7 份更短的说明书，所以**至少有 1 串写不短**。串越长，差距越大：1000 位长的串里，能缩短 10 个字以上的**不到千分之一**。所以大多数串都是随机的。

4 换一种语言，差别不大

换一种电脑语言写说明书，长度会变吗？会变，可是**变不了多少**。



机器人乙不认得“x”。没关系，在说明书前面附一本**翻译手册**，教它“x 几就是抄几遍”，后面照样写 (01)×500。关键是：**手册长度固定**，不管写哪一串，附的都是同一本。所以给乙的最短说明书，最多比给甲的长一本手册。数学家把这叫**不变性定理**：换一种电脑语言，柯氏复杂度最多差一个固定的数。这个数只看两种语言，和要写的串无关。

小心：手册可能比短串本身还长，所以只比很短的串没有意义。比长串时，手册那点长度才可以不管。

5 惊人的事实：最短说明书算不出来

想象一场“最短说明书比赛”：第 1 天，小明写了 30 个字。第 2 天，小红找到规律，写了 20 个字。第 3 天，小刚找到更深的规律，写了 12 个字。第 4 天呢？

裁判能说“12 个字比 20 个字短”，却**没法宣布冠军**：也许明天有人找到更短的。书里证明了：**没有任何一个程序，能对所有的串都算出最短说明书有多长**。数学家说柯氏复杂度**不可计算**。我们只能越找越短，却不能确定哪份最短。

为什么不把短说明书一份份试过去？有的说明书会让机器人一直算，你不知道它是快算完了，还是永远不停。还有一个更深的原因。读一读这句话：“**用二十个字说不出的最小的数**”（数指 1、2、3 这样的数）。它只有 14 个字，却说出了一个“二十个字说不出来”的数，自己跟自己打架！这叫**贝里悖论**。“悖论”就是怎么想都说不通的话。

同样的道理：测量机不可能存在

假如有一台**测量机**，给它任何一串，它都能说出最短说明书有多长。那我就能写一份说明书：

“用测量机按顺序一串一串地查，找出第一串最短说明书超过一亿字的，把它写出来。”

第 3 节讲过，这样的串一定有。测量机也写进去，整份说明书也远远不到一亿字。可它写出了那一串，那一串的最短说明书就不到一亿字了。说不通！所以测量机不可能存在。

6 实际中：用压缩软件估一估

算不出最短，怎么办？第 6、7 课的科学家用了**压缩软件**，就是电脑里把文件变小、还能一字不差还原的工具。小文件加上“解压程序”，就是一份说明书。真正的最短**只会更短或一样长**，所以压缩后的大小可以粗略估计信息有多少。

压缩软件也会看走眼。圆周率 3.14159…… 的前 100 万位，压缩软件几乎压不动。可是一个很短的程序就能把它们一位一位算出来。所以“压不动”不等于“真的随机”。

7 这和今天有什么关系

第 5 课讲过：朋友猜得越准，你传话用的字越少。**会预测，就会压缩**。柯氏复杂度就是压缩能达到的“最好成绩”。

ChatGPT、Claude 这些聊天 AI 靠的是大语言模型，它们靠“猜下一个字”练出来（第 1 课）。2023 年，DeepMind 公司的论文发现：大语言模型配上一种编码方法，就能当压缩软件。这些模型只读过文字，可它们压缩图片小块和声音，比 PNG（常用的图片格式）和 FLAC（常用的声音格式）还压得小。第 3、8 课的作者之一 Sutskever 多次说过：**好的预测器就是好的压缩器**。这本书就在他推荐的阅读清单里。

也要说清楚：柯氏复杂度算不出来，没人直接拿它训练 AI。模型压出的长度也只是“目前找到的一份说明书”。

8 动手：给机器人写最短说明书

玩法（2—4 个人，约 12 分钟）

1. 用热身的机器人规则，各自给每串写说明书。括号里可以再套括号。“×几”后面不要紧跟 0 或 1，不然看不清。也可以用珠子摆：红是 0，蓝是 1。
2. 同伴照着说明书写回原串，一字不差才打 ✓。谁最短，谁赢这一串。最后一行填自己掷 20 次硬币的结果：正面 1，反面 0。

串（每串 20 个字）	我的说明书	几个字	✓
00000000000000000000			
00110011001100110011			
01101101100110110110			
01000010110010000010			
我掷的硬币：			

你会发现：前 3 串能写得很短。第 3 串要套括号：“0，再加 3 遍 110”这一段重复 2 次。第 4 串和你掷的那串，怎么写都短不了几个字。就算你赢了，也不能保证没人写得更短。

9 动手：谁是随机的

玩法（3 个人，约 12 分钟）

A 真掷 20 次硬币，正面写 1，反面写 0。B 不碰硬币，凭脑子编 20 个 0 和 1，尽量像真掷的。两人把串抄到纸条上，不写名字，交给 C。C 用 3 把尺子检查，猜哪张是真掷的。揭晓后换人再玩。

	尺子 1：最多连着几个一样	尺子 2：0、1 换了几次	尺子 3：能写短吗	猜真的
纸条甲				
纸条乙				

你会发现：真掷 20 次硬币，大约 4 次里有 3 次会出现连着 4 个或更多一样的。0 和 1 平均换 9 到 10 次。人编的串常常换得太勤，很少连着 4 个一样。“几乎不连着一样”也是一条规律：想装随机，反而露出了规律。

10 词语小卡片

柯氏复杂度

写出一串东西的最短说明书有多长，量的是信息有多少。

随机的串

没有比它自己短多少的说明书。大多数串都是。

不变性定理

换一种电脑语言，最短说明书最多差一本“翻译手册”的长度。

不可计算

没有程序能对所有串算出最短说明书的长度。

11 想一想

1. 小明真的掷了 20 次硬币，结果全是正面。按这一课的定义，这串算不算“随机的串”？为什么？

2. 小刚说：“我给这串写了 15 个字的说明书。我敢保证，世界上没有更短的。”你相信吗？为什么？

3. 聊天 AI 很会猜下一个字。为什么说它也能当压缩软件用？

一句话记住：最短说明书有多长，信息就有多少。写不短的，就是随机的。

给家长和老师：本课讲的是 A. Shen、V. A. Uspensky、N. Vereshchagin 的书 [Kolmogorov Complexity and Algorithmic Randomness](#) (AMS Mathematical Surveys and Monographs 第 220 卷，2017)。讲义只涉及前几章： $C(x)$ 的定义、不变性定理、大多数串不可压缩、 $C(x)$ 不可计算（贝里悖论式证明）和“不可压缩即随机”的直觉；无限序列的 Martin-Löf 随机与前缀复杂度未展开。

“机器人语言”是给孩子设计的简化描述方式，不是通用编程语言；书中程序是固定通用机器上的二进制串。各串、比赛、翻译手册和测量机的说法都是示意。“掷 20 次约 77% 出现 ≥ 4 的连串、平均换 9.5 次”是我们按公平硬币算的；“人编的串换得太勤”是常见课堂现象，非书中内容。DeepMind 论文为 Language Modeling Is Compression (2023)。想深入可看本站的 [第 26 天中文导读](#)，代码见开源项目 [30u30 第 26 天目录](#) (英文)。