第51卷第3期 1915年创刊 Vol.51 No.3

科學

ISSN 0368-6396

SCIENCE

1999 3 (双月刊)

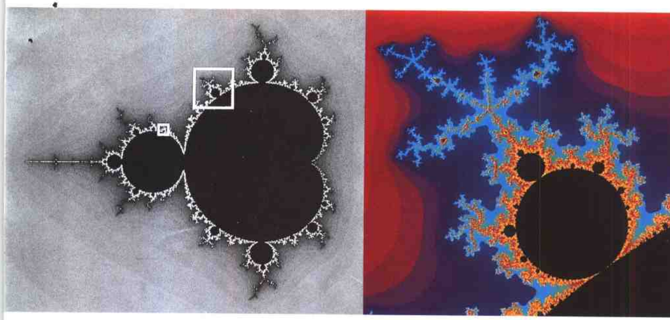
第51卷第3期 1915年创刊 Vol.51 No.3

科學

ISSN 0368-6396

SCIENCE

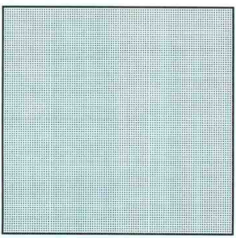
1999 3 (双月刊)



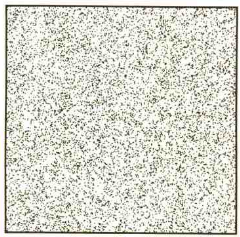
曼德勃罗集合是复数迭代 $z_{n+1}=c-z_n^2$ 中复参数 c 的平面。黑色表示迭代收敛到某种周期的参数区，不同灰度或色彩代表各种发散速率的参数区域。左图是整个参数平面的中心部分，右图是左图中小方框的放大。后者还可以继续放大下去，层出不穷的图案只是似曾相识，并没有自相似性。(原作者为H.-O. Peitgen和P.H. Richter)

图1 曼德勃罗集合一例

曼德勃罗集合与“复杂”图案

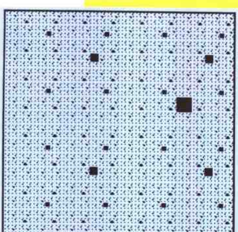


(a)101 × 101 个周期排列的点。

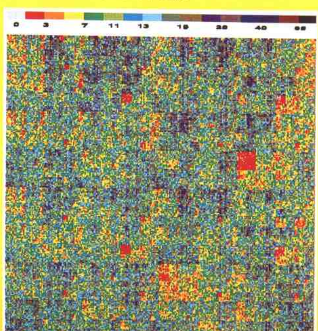


(b)10000 个随机分布的点。

图2 周期、随机、“自相似”和“复杂”图案



(c)一种稍有自我重叠的自相似图案，只画出最初的几个层次，原则上可以无穷细化，也可以有严格的数学描述。(取自郝柏林、谢惠民等的论文)



(d)大肠杆菌DNA序列中长度为8的字母串出现频度的一种表示。图中左上角为gggggggg，右上角为cccccccc，左下角为aaaaaaaa，右下角为tttttttt。频度0到774只用由白到黑的16种颜色表示，这也是一种粗粒化。图中隐约可见c图所示的图案。这是因为大肠杆菌基因组中包含c、t、a、g的字母串特别稀少。(取自郝柏林、李弘谦等的论文)

郝柏林

复杂性的刻画与“复杂性科学”

近些年来，关于“复杂性”、“复杂系统”乃至“复杂性科学”的议论在国内外颇为流行。出现了大批讨论“复杂性科学”的专著和文集，召开了许多以“复杂性”为论题的会议。中国的“复杂性热”固然有偏好“新”提法的科学管理人士起了推波助澜的作用，但这也确实反映出科学发展的一种趋势。

自然科学在宇观、宏观和微观三个前沿发展着。探索宇宙演化奥秘和微观世界结构的两个尖端，涉及愈益庞大的投资和设备，往往成为单个国家难以独立支持的项目。美国前几年下马的超级超导对撞机和现仍在巡天的哈勃望远镜是两个实例。当今世界上，有幸直接从事这类研究的科学家数目逐年减少，多数人只能从旁欣赏他们的成果与发现。这两方面的研究结果，有重大的认识论意义，在历史尺度上而不是计日程功地改变着人类的生产方式。与此形成尖锐对比，宏观层次的自然科学，包括广义的物理科学、生命科学、地球与环境科学等等，集中了众多的人力与物力，也同人类的生产和生活更加息息相关。这里的研究对象斑驳陆离、五花八门，有没有共同的理论线索呢？复杂性的刻画看来是一个贯穿了不少领域的课题。然而，这又是一个容易使人望文生义，引发种种伪科学议论的概念。本文试图分析复杂性研究的现状，并陈述笔者的观点。

什么是复杂性

首先，复杂性并不是新问题。在美国国会图书馆1975年至1999年2月15日的入藏书目中，标题里含复杂性(complexity)一词的就有489种。其中涉及算法复杂性、计算复杂性、生物复杂性、生态复杂性、演化复杂性、发育复杂性、语法复杂性，乃至经济复杂性、社会复杂性，凡此种种，不一而足。其实，现代电子计算机之父冯·诺伊曼早就说过，“阐明复杂性和复杂化概念应当是20世纪科学的任务，就像19世纪的熵和能量概念一样。”看来，20世纪的科学没有完成这个任务，要把它传递到新的千年。

普里高津(1977年诺贝尔化学奖获得者)等[1]曾建议，把1811年作为“复杂性科学”的起点。这一年，傅里叶因发现热传导定律而获法国科学院的大奖。诚然，作为最早的普适定律之一，傅里叶的贡献具有重大历史意义；然而，它只是一个线性律，而现代复杂性研究要面对的是非线性现象。在普里高津与人合著的《探索复杂性》一书[2]中，首章首节开宗明义地以“什么是复杂性”为标题，后文又罗列了许多自然界中的复杂现象和复杂过程，然而却没有回答这个问题，只是说“提复杂行为比复杂系统更自然些”。

安德森（1977年诺贝尔物理奖获得者）写过几篇短文[3]，介绍以研究复杂性为宗旨的圣菲(Santa Fe，亦有人译作圣达菲)研究所。他把诸如对称破缺、局域化、分形和奇怪吸引子等“各种新性质怎样冒出来”的种种思想贯穿起来，作为复杂性科学的研究对象。安德森指出，复杂性研究不应像一般语义学或一般系统论那样，“早熟和轻率地”企图建立包罗万象的构架；而应当注重特定的、可以检验的机制和概念。

几本90年代初出版的标题类似的通俗读物[4-6]，都把复杂性研究作为介于有序和混沌边缘的科学。书中宣传的基本上是圣菲研究所那批人的观点，虽然该所创始人之一、诺贝尔物理奖获得者盖尔曼曾把湍流和混沌列为该所不予研究的“经典”问题。

圣非研究所地处美国新墨西哥州首府圣菲，1984年由盖尔曼、安德森，以及阿罗(K. Arrow，1972年诺贝尔经济学奖获得者)等人倡议建立。它是一个专门研究复杂性的机构，只有少数固定编制、多数人员是流动

郝柏林：研究员，中国科学院理论物理研究所，北京 100080；中国科学院院士；第三世界科学院院士。

Hao Bailin: Professor, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100080; Member, Chinese Academy of Sciences; Member, Third World Academy of Science.

的，目前的年度预算约为500万美元（由私人基金和美国科学基金会支持）。其研究课题涉及地球上出现生命之前的化学演化和之后的生物演化、哺乳动物的免疫系统理论、人类与动物个体的学习和思维、人类文化和语言的演变、全球经济作为复杂的演化系统、计算机和程序设计的全新战略，等等。该所的计算机条件很好，从核酸大分子、人工生命，到社会组织、股票市场，似乎模拟什么像什么。然而，在那里，透过海量的数据和图象，洞察所得的概念却相对贫乏；在一定意义上，还是回到了霍兰德早就提出的自适应系统的概念[7]。1994年，该所召开了建所十年的评估会，人们在会上讲了许多好话，会下却不乏微辞。有人甚至写了题为"从复杂性到困惑性"的文章[8]。

然而，什么是复杂性呢？复杂性能成为一门科学吗？能不能对复杂性作定量测度？复杂性是状态的函数还是过程的函数？只有先思索一些自然科学中从简单到复杂的事例，才能试图寻求这些问题的答案。

简单性、随机性和复杂性

简单性一向是现代自然科学、特别是物理学的一条指导原则。许多科学家相信自然界的基本规律是简单的。爱因斯坦曾是这种观点的突出代表。虽然复杂现象比比皆是，人们还是努力要把它们还原成更简单的组分或过程。事实上不少复杂的事物或现象，其背后确实存在简单的规律或过程。应当学会比较和刻画来自简单机理的复杂性，否则很难期望会正确分析那些机理不明的复杂事物或现象。下面考察几种由简单变复杂的途径。

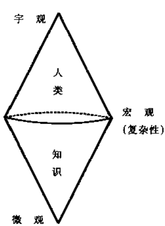
首先，重复使用简单的规则，可能形成极为复杂的行为或图形。

一维非线性函数的迭代导致混沌[9]，是一个熟知的例子。像 $f(x) = a - x^2$ 这样的抛物线函数，只要相当任意地选取一个初始值 x_0，不断进行迭代，即计算 $x_{n+1} = f(x_n)$，在参数 a 的某些取值范围，所得到的"轨道" x_0、x_1、x_2…可能具有极其复杂的结构。如果把上面的实数 x 扩展成复数，在整个二维复数平面上进行迭代，那末就有可能得到层出不穷的花纹图样。曼德勃罗(B. B. Mandelbrot)集合（参见封二图1），就是一例。有人称它为数学所知的最复杂的对象，因为在精益求精、小而又小的无穷层次上，都要出现新的、与已有结构并不严格相似的花纹图案。

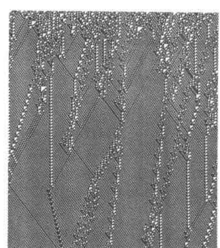
简单规则导致复杂行为的另一个例子，是所谓一维元胞自动机。取多枚硬币排列成一条直线，每个硬币可能正面向上，也可能反面向上。根据每一枚硬币自身和左右两邻的状态，确定下一时刻是否翻动。这样，使用只涉及最近邻的简单生成规则，却可能得到各种各样的图案，包括只能由万能的图灵计算机模拟的复杂行为。

其次，把物理过程从高维空间投影到低维，会使它们看起来更复杂。或者倒过来说，增加新的参数或变量，扩大参数空间或相空间，往往可使事情简化。

随意放置在三维空间中的一条曲线，一般不会自己相交；即使发生自交，也可以用小扰动排除，因而自交是非实质的。然而，如果把它投影到二维平面，一般说来定会产生自交，而且不可能由小扰动排除，看起来也比三维空间中的曲线复杂。某些非线性问题可以嵌



人类知识的三个发展前沿 纺锤体体积代表人类知识的总和。



一维元胞自动机所产生的图案 这是第110规则产生的花样，时间由上往下发展。（原作者为S. Wolfram）

入更高维的空间，成为线性问题。某些非马尔可夫过程可以靠增加新的随机变量，成为马尔可夫过程。许多连续模型的离散化方案，也可看成从无穷维空间投影而来，它们一般具有比原来更复杂的性质。

还有，错误的参考系可能带来不必要的复杂化。

历史上托勒密的地心系就是如此。为了描述当时对太阳和六大行星运行的观测结果，曾经不得不引入80多个“本轮”和“均轮”。一旦换成哥白尼的日心系，天体运行的图象就变得简单多了。

显然，考察种种事物由简单变复杂的途径，颇具启发意义。不过，在这种考察中，不要混淆描述体系的复杂性和刻画客观的复杂性。

客观地定义和量度复杂性，与人们对自然界描述体系的复杂性是两回事。这很像是美和美感的关系。前者应有客观定义，而后者涉及接受者的主观条件。一个突出的例子是描述电磁场的麦氏方程组。麦克斯韦在其1864年的论文中引入了20个变量和决定这些变量的20个方程，即电磁场的基本方程组。麦克斯韦当时甚至没有使用矢量记法，因此这些式子看起来特别复杂。要从这些式子中看出电磁波的存在，确实需要一位天才。麦克斯韦之后一百年，电磁场的基本方程组可以写成“简单”的形式：$\delta d\Theta = J$。然而，要看懂这个方程，必须明白现代微分几何学中外微分 d、上微分 δ 和微分 1－形式 Θ 的意义。只有站得高，事物才显得简单。原始人心目中的复杂事物，现代人看来未必复杂。

如何刻画客观的复杂性，这实际上是个如何脱离主观评价定义某种可测量对象的问题。其实，科学发展史上早就遇到过这类问题。信息量的定义是一个著名实例。1928年，哈特利(R. V. Hartley)按对事件概率估计的改变量定义信息的对数测度，它对于无关事件具有相加性。1948年，香农把它推广成现在广为人知的信息公式，信息论开始成为一门科学。香农的信息定义，使人联想到当年玻尔兹曼对熵的统计解释，而熵是宏观系统中无序度、随机性的量度。纵观宏观物理学史上各种概念的发展过程，熵、信息和复杂性这三个概念，所引发的深刻研究之多、无稽议论之不着边际，确实很少有其他概念能与之相比。

为了对简单性、随机性和复杂性有更多的直觉，不妨考察一下封二图2中的四种花纹图案。试问何者简单，何者复杂？通常人们会立即指出，图2a是周期图案，因而是简单的。图2b一下说不清楚，如果告知这是由随机数产生的1万个点，也不会觉得有多么复杂了。图2c要稍加思索，当明白它具有可以精确描述的有一定重叠的“自相似”结构以后，也不会被认为十分复杂。图2d像是在某种随机背景上叠加了一定的结构，这些结构既非周期也不完全自相似。看来这是四个图案中最复杂的一个。

$$
\begin{aligned}
a &= \frac{dH}{dy} - \frac{dG}{dz} \\
b &= \frac{dF}{dz} - \frac{dH}{dx} \qquad (A) \\
c &= \frac{dG}{dx} - \frac{dF}{dy}
\end{aligned}
$$

$$
\begin{aligned}
P &= c\frac{dy}{dt} - b\frac{dz}{dt} - \frac{dF}{dt} - \frac{d\Psi}{dx} \\
Q &= a\frac{dz}{dt} - c\frac{dx}{dt} - \frac{dG}{dt} - \frac{d\Psi}{dy} \qquad (B) \\
R &= b\frac{dx}{dt} - a\frac{dy}{dt} - \frac{dH}{dt} - \frac{d\Psi}{dz}
\end{aligned}
$$

$$
\begin{aligned}
X &= vc - wb & \qquad a &= \alpha + 4\pi A \\
Y &= wa - uc \quad (C) & \qquad b &= \beta + 4\pi B \quad (D) \\
Z &= ub - va & \qquad c &= \gamma + 4\pi C
\end{aligned}
$$

$$
\begin{aligned}
4\pi u &= \frac{d\gamma}{dy} - \frac{d\beta}{dz} \\
4\pi v &= \frac{d\alpha}{dz} - \frac{d\gamma}{dx} \qquad (E) \qquad & \mathcal{D} &= \frac{1}{4\pi}K\mathcal{E} \quad (F) \\
4\pi w &= \frac{d\beta}{dx} - \frac{d\alpha}{dy} & \mathcal{R} &= C\mathcal{E} \quad (G)
\end{aligned}
$$

$$
\begin{aligned}
u &= p + \frac{df}{dt} & \qquad u &= CP + \frac{1}{4\pi}K\frac{dP}{dt} \\
v &= q + \frac{dg}{dt} \quad (H) & \qquad v &= CQ + \frac{1}{4\pi}K\frac{dQ}{dt} \quad (I) \\
w &= r + \frac{dh}{dt} & \qquad w &= CR + \frac{1}{4\pi}K\frac{dR}{dt}
\end{aligned}
$$

$$
\rho = \frac{df}{dx} + \frac{dg}{dy} + \frac{dh}{dz} \qquad (J)
$$

$$
\sigma = lf + mg + nh + l'f' + m'g' + n'h' \qquad (K)
$$

麦克斯韦著作中的电磁场基本方程组 这是从麦克斯韦经典著作《电学与磁学专论》(1873年初版)中抄来的式子，其中(A)、(B)等是原来的公式标记。麦克斯韦深知这一方程组的重要性，用大写字母标示出有关公式，而其他式子用数字编号。

以上直观考察揭示了近年复杂性研究的一条重要成果：随机性并不复杂（虽然也有人说随机性是最大的复杂性[6]），历史上不少复杂性的定义其实针对的是随机性，复杂性介于随机和有序之间，是随机背景上无规地组合起来的某种结构和序。只有考察一些复杂性的定义，才能领会这几句话的具体涵义。

复杂性的各种“定义”

同信息论的奠基过程相比，复杂性的刻画已经走过更长的道路，却远远未臻完备。根本的原因在于不存在复杂性的绝对尺度。不界定基本的框架，根本不能就复杂性问题对话，更休提进行科学分析。人们已经提出过不下30种各有其适用范围的复杂性“定义”，这里举示若干。

历史上较早提出的是计算复杂性概念，它源于20世纪30年代数学逻辑的一些深刻命题。每个问题都有其特定的规模 N，如货郎担问题有必须访问的村庄数目，或是需要求逆的矩阵的阶。解决问题所需的代价

（如计算时间），如何随问题规模 N 变大而增长？若代价的增长不超过 N 的某个幂次或多项式，该问题是简单的，属于 P(即多项式)类。若增长速率超过 N 的任何多项式，则问题是艰难的，属于 NP 类。严格地说，若已知某 NP 问题的解，可付出 P 类代价加以演示；但若要找到这个解，则须付出 NP 代价。若两个 NP 问题可用 P 代价彼此转换，则它们是同等艰难的。这些等价的 NP 问题构成所谓 NP 完备类。目前已经知道上千个 NP 完备问题。同属于 P 类的问题，其代价比例于 N 的对数或平方，仍是一种实际差别。把一种算法从 N 改进到 $\log N$，或者证明对某个问题不存在比 N^2 更好的算法，都是严肃认真的科学成果。计算复杂性的研究有专门期刊[10]，以及许多书籍和会议文集。

算法复杂性是 1964－1966 年由索洛莫诺夫(R. J. Solomonoff)、科尔莫戈罗夫(A. N. Kolmogorov)和柴廷(G. J. Chaitin)分别独立提出的。粗略地说，算法复杂性就是产生特定的图形花纹（或符号序列）的最短程序的长度与图形花纹（或符号序列）本身的大小之比的极限——当后者趋向无穷时的极限。这里“长度”和“大小”均按二进制位数计，而“程序”则是在普适的理论计算机上执行。

以一个二进制数，如

1101110010100110100001011001110011010111110 …

为例。不难明白，上述算法复杂性其实是随机数的一种判据。为写出实质上比给定数更短的程序，须利用该数包含的某些规律。若无规律可循，依定义该数就是随机的，相应的最短程序只能是在打印语句中照抄该数。程序长度比该数多出 PRINT 五个字母，取无穷极限后就没有差别了。算法复杂性等于 1，就是最随机的无法压缩的图形花样。这是一种深刻的思想，但一般说来，它是不可操作的：是否存在比该数字更短的程序，往往无法确定。不仅如此，下面将说明，算法复杂性并没有解决复杂性的比较问题。

回到产生曼德勃罗集合的复数迭代。它的程序是很简短的。要看到复杂的细节，必须长时间迭代。迭代愈久，细节的复杂层次愈多。因此，有人引人“逻辑深度”的概念，即以在普适计算机上执行前述最短程序的指令步数或时钟拍数，来比较不同图形的复杂程度。这是把计算复杂性和算法复杂性结合起来的度量。不过，它也还没有解决问题。

从复数退回到实数，回到前面提到的抛物线迭代。对于参数 a 和初值 x_0 的无穷多种选择，这种迭代会导致混沌轨道。用相同的程序，执行同样多的步数，在两个不同参数下得到的同样长的混沌轨道，其复杂性有差别吗？如果不能回答这样简单明确的问题，就不能奢望去刻画更普遍的复杂性。所幸的是，可以通过符号动力学建立运动轨道和形式语言的联系，然后借助语法复杂性理论确切地回答一部分这类问题。

在抛物线映射中，忽略轨道点的具体数值，只注意它是落在抛物线最高点的左面还是右面，相应地把数值换成字母 L 或 R。于是，数字轨道变成符号序列，迭代动力学成为符号动力学。这实际上就是物理学里经常实行的粗粒化描述。不同程度的粗粒化，舍去更小层次上的细节，有利于突出更根本的行为。适当地粗粒化，有助于得出更严格的结论。事实上，粗粒化考虑是把人们引向复杂性问题研究的一项深刻纲领。

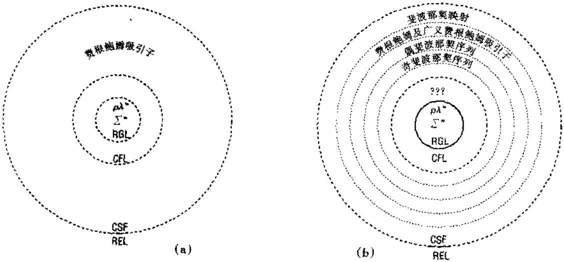
粗粒化描述与符号序列

科学研究不能从定义而要从对事实的分析出发。直观地看，从原子、分子、晶态和非晶态固体，到高分子、液晶这些“软”物质，到组成生物的“活”物质，乃至人脑、意识和思维，科学研究的对象确实是在沿着复杂性的阶梯上升，同时也变得愈益特殊。唯其特殊，才有丰富的内容。这再次提醒人们，复杂性的研究切忌泛泛议论。下面的观察，会有助于构造出分析复杂性的一种有效框架。

研究基本粒子结构的人，看见 u、d、c、s、b、t 这六个小写字母，会立即认出它们是六种夸克的名字，并且从字母联想到质量、电荷和其他性质。更多的学者用 p、n、e 这几个符号，表示质子、中子和电子，知道它们的电荷、质量、自旋、磁矩等特征，但并不关心质子和中子分别由那三个夸克组成。

化学家们看到 H、C、N、O、P、S 这些元素符号时，会想到它们的原子序数、离子半径、化学价和亲和力等等。看到用元素符号写出的化合物的分子式，如 H_2O、NO、CO_2 等，会联想到它们具有一定的分子量，是透明液体或无色无嗅气体。然而，即使在遇到还不算太大的核苷酸分子时，如果每次都把三四十个原子的元素符号写出来，则既不方便也无必要。

人们用 a、t、c、g 代表四种不同的核苷酸，更注意到它们在组成 DNA 双螺旋时，a 和 t 由两个氢键相连，而 c 和 g 由三个氢键相连，分别称为弱偶合和强偶合。地球上各种各样生物的遗传基因都是由这四个核苷酸以不同的顺序排列“编码”的。小小的大肠杆菌的遗传信息乃是由 4639221 个符号排成的一个长字，其中只有 a、c、g、t 四个字母，而人的全部基因分别组织在 23 对染色体中，估计有 30 亿个字母。一切蛋白质都是根据编码于基因中的信息在细胞中合成的。每个蛋白质是



对抛物线映射中轨道复杂性的认识 RGL、CFL、CSF、REL分别代表正规语言、上下文无关语言、上下文有关语言和递归可数语言。Σ^*和$\rho\lambda^*$代表周期轨道和最终周期轨道。(a)1991年的认识水平：只知道费根鲍姆吸引子属于CSF。(b)现在的认识水平：CSF有了更细的层次，偶和奇斐波那契序列、费根鲍姆及广义费根鲍姆吸引子、斐波那契映射，一个比一个复杂；RGL圈的虚线变成实线，因为正规语言问题完全解决；三个问号表示谢惠民的"不存在上下文无关语言"猜测。

一条由20种氨基酸按特定的顺序排列成的大分子。单个氨基酸比核苷酸小一些，有10到27个原子，又可以用一个字母表示。胰岛素是一种相当小的蛋白质，研究其三维立体结构时需要知道上千个原子的位置，比较人、牛、猪的胰岛素时可以考察51个符号组成的大同小异的序列，讨论葡萄糖代谢过程时胰岛素又往往用一个记号代表。

这类叙述还可以继续延伸几个层次。倒过来说，"从头算起"，企图由夸克出发阐明生物圈，或者把宇宙称为一个"复杂巨系统"，都并未真正深化对自然界的认识，丰富人类的科学知识。

以上观察启示了一个深刻的研究纲领：

(1) 研究自然现象时必须瞄准一定层次进行粗粒化描述。更精细层次上的差异，在所关注的层次中表现为某些特征量，例如热传导系数。

(2) 粗粒化描述不可避免地要使用符号和符号序列。很多情形下这导致一维符号序列。其实，对于有限个符号组成的序列，"高维"可以归结为具有远邻关系的一维序列。

(3) 符号序列可以自然地纳入"形式语言"的框架。形式语言有别与我们日常使用的"自然"语言(后面再稍详细地介绍)。

(4) 形式语言可按语法复杂性的阶梯分类，各种符号序列的具体研究又能丰富语法复杂性的内容。

应该强调，以上所述只是一个研究纲领，而非已经完成的建筑。迄今为止，这个纲领只是在动力学问题中实现得较好，在与生命有关的符号序列上有所尝试。

形式语言和语法复杂性

先简单介绍一下语言和语法复杂性的概念。首先要有一个特定的字母集，例如前面提到的R和L两个字母，或者0和1两个数字，都可以构成两个符号的字母集。26个大写和小写罗马字母，加上13个标点符号，可以构成65个符号的字母集。字母集里面的符号，本身不再具有任何含义。通常考虑由有限个符号组成的字母集。取给定字母集里的符号，组成一切可能的符号串，包括不含符号的空串，它们构成有无穷多元素的大集合。这个集合的任何一个子集合，称为一种语言。这就是形式语言的定义。

从如此普遍而抽象的定义出发，走不了多远。必须指出，所关心的语言是由那些字符串或"字"组成的。有限个字组成的语言原则上可以用穷举法描述。有些无穷语言也很容易表述，如集合$\{R^nL^n, n\geq 1\}$就是一种基于字母集{R, L}的语言。更有效的办法是，指定一个或几个初始字母和一组"生成规则"，把生成规则反复使用到初始字母和新生成的字上，产生出整个语言。这就是由"生成语法"定义的语言。

生成规则可以是串行或并行的。乔姆斯基(N. Chomsky)在20世纪50年代给出了串行生成语法的完全的分类。原来一切串行生成的语言，包括人们熟悉的许多程序设计语言，可以分成四大类，每类语言可由一类自动机来接受(或者叫执行、识别)：

（1）正规语言（RGL），由不需存储器的有限自动机接受。

（2）上下文无关语言（CFL），由带堆栈（又称后进先出区）存储器的计算机接受。常用的算法语言，例如FORTRAN，就属于这一类。

（3）上下文有关语言（CSF），由存储器容量比例于输入量的“线性有界自动机”接受。

（4）递归可数语言（REL），由存储器无限大的图灵计算机接受。

从上往下的分类，给出了由简单到复杂的乔姆斯基阶梯。由物理观察经粗粒化所得的符号序列，如果可以阐明其生成语法，就有了一个有严格数学背景的复杂性尺度。乔姆斯基阶梯并不是一成不变的标准。对具体问题的具体分析，往往可以丰富其内容，在某些阶梯中作出更细的划分。

不难看出，一维抛物线映射中的周期和最终周期轨道，对应最简单的正规语言。最近，我国学者谢惠民等证明了更困难的逆命题，即对应正规语言的轨道只有这两大类。许多对应正规语言的轨道族的极限，例如倍周期分岔序列的无穷极限，跳过上下文无关语言，达到上下文有关语言的复杂程度。于是出现猜测：一维抛物线映射导致的符号序列中没有上下文无关语言。还可以引用语法或自动机的更细致的性质（例如，允许字或禁止字的数目、有限自动机的结点数目，特别是它们的增长速率），来比较符号序列的复杂程度。符号序列的演化可以用转移矩阵描述。无穷大的转移矩阵的块结构，提供考察超出正规语言复杂性的另一种途径。于是，关于抛物线映射中轨道复杂性分类的知识，已经大为超过90年代初期的水平。对这些进展感兴趣的读者，可参阅文献[11]和[12]。

20世纪60年代，荷兰发育生物学家林登梅耶（L. Lindenmayer）观察了某些多细胞藻类植物的发育过程。其中有种项圈藻属的串珠藻（*Anabaena catenula*），它的细胞具有分裂方向的极性。小细胞长大后进行分裂，把极性传给新出生的小细胞，自己的极性反向，下一次在另一头分裂出一个小细胞。这样，从一个具有某种极性的小细胞开始，若干次分裂后就成为一株大小和极性具有特定排列图案的项圈藻。林登梅耶用4个字母代表极性向左向右的大小细胞，写下它们之间的变化规则。反复使用这些规则，得到的结果同显微镜观察完全一致。不久，人们认识到这种做法可以推广成一套并行生成语法，并对它进行了分类，得到不同于乔姆斯基的另一种形式语言分类体系，特称为L系统。L系统很可能对生物学问题有更多应用。

林登梅耶分类中较为简单的一个层次，可能跨越乔姆斯基分类的几个复杂性阶梯。有些看起来“简单”的语言，可能处于相当高的复杂性层次。例如，$\{R^nL^n, n\geq 1\}$不是正规语言，而是上下文无关语言，$\{R^nL^nM^n, n\geq 1\}$进而达到上下文有关语言。这从另一个侧面说明前文的论断，即不存在复杂性的绝对度量。形式语言的分类，当然也不限于这里介绍的两个体系。

对具体问题作具体分析，是科学知识的源泉。许多普适性的概念，例如分形、混沌、信息，复杂性，有助于整理研究结果，却不能代替具体研究[13]。用“新”名词包装，造不出科学成果，何况许多“时髦”名词并不新鲜。逻辑学中关于概念外延与内涵成反比的定律，说明创立大而无当的空洞体系的种种企图，不会导致具体和有益的科学知识。在现代科学和科学方法论基础都十分薄弱的中国，人们应该警惕不要受这样的“新学科”的诱感而误入伪科学的泥沼。

（文中提及的本文作者的研究受到国家自然科学基金和攀登计划的支持。）

[1] Prigogine I, Stengers I. *Order Out of Chaos*. London: Heinemann, 1984. 104；中译本：普里高津，斯腾热尔．从混沌到有序，人与自然的新对话．上海：上海译文出版社，1987

[2] Nicolis G, Prigogine I. *Exploring Complexity*. New York: W H Freeman & Co, 1989；中译本：尼科里斯，普里高津．探索复杂性．成都：四川教育出版社，1986

[3] Anderson P W. *Physics Today*, 1991(7): 9；1992(6): 9

[4] Waldrop M M. *Complexity: The Emerging Science at the Edge of Order and Chaos*. New York: Simon and Shuster, 1992；中文缩译本：复杂，诞生于秩序与混沌边缘的科学．北京：生活·读书·新知三联书店，1997

[5] Lewin R. *Complexity, Life on the Edge of Chaos*. Macmillan, 1993

[6] Casti J L. *Complexification*. Harper Perennial, 1994

[7] Holland J H. *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley, 1995

[8] Horgen J. *Scientific American*, 1995, **272**(6): 104

[9] 可参看：郝柏林．从抛物线谈起——混沌动力学引论．上海：上海科技教育出版社，1993

[10] 1985年创刊的 *Journal of Complexity* 基本上是一份讨论计算复杂性的刊物．

[11] 关于乔姆斯基和林登梅耶分类的介绍和引文，参见：Huimin Xie（谢惠民）. *Grammatical Complexity and One-Dimensional Dynamical Systems*. Singapore: World Scientific, 1996

[12] Bai-lin Hao（郝柏林），Wei-mou Zheng（郑伟谋）. *Applied Symbolic Dynamics and Chaos*. Singapore: World Scientific, 1998. Chapter 8

[13] 读本文校样时，见到 *Science* 的一个关于复杂系统的专辑（1999年4月2日出版），其中涉及的正是这类“具体研究”．

关键词：复杂性　简单性　随机性　粗粒化　符号序列　形式语言　计算复杂性　语法复杂性　□