

·科学论坛·

## 与信息科学的结合为生命科学的研究开辟新的前景

李衍达\*

(清华大学信息学院,北京 100084)

**[摘要]** 本文通过对生命科学和信息科学的阐述,指出生物体并不是细胞的简单堆积,而是有多种信号在多种通道中传递、调控的复杂网络系统,认为信息科学与生命科学的交叉将为生命科学的研究开辟一个新的前景。

**[关键词]** 生物信息,分子生物学,调控机制,信息科学,控制论,复杂系统

生命科学的研究已经发展到一个新的阶段,人类基因组计划的实施标志着生物学的研究在分子水平上有新的突破,另一方面,生物体并不是细胞的简单堆积,它是有多多种信号在多种通道中传递、调控的复杂网络系统,因此,必须在系统分析水平上有新的突破。无论在分子水平上还是在系统水平上进行进一步的研究,都需要与信息科学的密切结合。

信息科学,尤其是控制论,最早还是来自对生物的研究。控制论的创始者维纳教授就是通过与哈佛医科学学校的许多年轻的医学专家经过多次讨论而提出控制论的,《控制论》这一书的副标题就是“或关于在动物和机器中控制和通讯的科学”。通过对动物和机器的通讯与控制系统的分析,他们发现,有一个共同的、核心的特点,就是对信息的传递和加工过程。因此,对信息的传递与加工的研究就抓住了动物与机器的控制系统研究的实质。信息科学因而成为通讯系统与控制系统研究的最重要的理论与方法。随着微电子、计算机的飞速发展,信息科学已成为发展最为迅速、影响极其广泛的学科,几乎所有领域的发展都离不开信息技术。可惜的是,生命科学与信息科学之间的结合尚未得到应有的重视与发展。现在已经很清楚,无论是机器还是社会,或是生物,它们的控制过程,其核心都是信息的传送与加工过程。信息理论的许多原理、方法对生物系统的研究都是适用的,不仅如此,应用信息理论对生物控制

系统(无论分子水平还是整个生物体)的研究,将使人们从生物体多种纷繁的现象中抽出其本质,不仅使问题更加简明,而且使研究者能从一个更高的视角去分析问题,从而可能取得新的突破。而各种信息技术与手段的应用也极大地帮助研究者从浩渺的数据中找出其中的规律。

例如,我们以编码理论来分析基因的结构时,可以发现,遗传密码采用三联体编码是很符合编码原理的,它既可充分描述 20 多种氨基酸(4 个碱基三联体的全部排列有 64 种可能),又是最省的组合。又如计算启动子区的信息熵,发现其可描述的信息量与随后的编码区的大小有对应的关系,这也是符合信息论原理的。有理由可以假设,为了根据输入信号按规定的程序进行转录、复制与调控等,除了三联体编码外,还存在其他的编码方式,而新的编码方法也应符合信息论的要求,即既能满足被编码事件的容量要求,有一定的容错能力,同时也是符合节省原则的。

如果我们把 DNA、细胞、各种组织间的相互作用看作一个信息系统,就可以看到,代表生物体的信息系统将是一个多信号、多通道、相互交联的复杂的信息网络。因此,我们可以用描述、分析复杂信息网络的方法对它进行分析。例如,近年国际上兴起的国际互联网络 Internet(因特网)就是传递多种信号(有数据、语音、文件、图形、图象等)的复杂的信息网

\* 中国科学院院士。

本文于 1999 年 7 月 29 日收到。

络,我们对其业务流量进行建模时,发现与过去单纯的语音网络有很大的不同。其业务显示出有长时相关特性,即显示出有复杂系统的分形特性,而且具有多重分形的性质。应该采用 Hurst 指数来描述才更为合理,由此,我们联想到 DNA 序列也应具有长时相关特性,即具有复杂信息网络的性质。对现有的 DNA 长序列的计算表明,的确具有长时相关的性质。这进一步说明可以应用复杂信息网络的观点来分析生物体的调控作用,可以充分利用已有的信息网络各种理论成果来分析生物体的信息系统。

由于计算机等信息技术的飞速发展,利用计算机对 DNA 序列及其三维结构等进行仿真及建模对研究来说十分重要。计算机建模与实验技术相配合已成为当前重要的研究手段,而计算机建模技术也有重大发展,例如人工生命的出现,计算机仿真软件可以使屏幕上出现具有生命特征的模型,如人工鱼,可具有觅食、避害、对环境的自适应、繁殖等能力。可以预期,计算机人工生命的技术将对生命科学的

研究起着不可缺少的作用。尤其对生命现象的建模会有重大影响,它将大大加快我们对生命现象的各种试验和推测的验证。

近年大量出现基因组数据,各种有关生物数据数量正以指数式增长,如何对大量数据进行存储、管理、对比、检索,寻找其相关性,挖掘数据中所隐藏的规律,综合多个数据库所提供的信息以及对海量数据进行自动分析等等都需应用信息技术。

从历史来看,信息科学技术在它的发展过程中大量吸收来自生物学的知识,例如,人工智能、人工神经网络、遗传算法、进化策略、启发式搜索算法、人工生命等等。我们相信,信息科学将继续从生命科学的研究中吸取新的营养,发展出新的分支。反之,现在到了大力提倡,认真地将生命科学与信息科学结合起来,用信息理论与信息系统的观点对生物分子与生物组织进行新的深入的分析的时刻,这种学科的交叉必将为生命科学的研究开辟一个新的前景,从而可能取得新的突破。

## THE PROSPECT OF COMBINATION INFORMATION SCIENCE WITH LIFE SCIENCE

Li Yanda

(The School of Information Science Technology, Tsinghua University, Beijing 100084)

**Abstract** The essence of living things is identified by its information system. The mutual operation among biological molecules, no matter in macroscopical sense, viz. the whole body of the organism, or in microcosmic sense, viz. in the genomic level, can be described and analyzed by cybernetics and the theory of complex systems, therefore, the combination information science with life science will open up a new prospects for life science.

**Key words** bioinformatics, molecular biology, regulation mechanism, information science, cybernetics, complex system