

计算机人工生命

Artificial Life

胡瑞安 胡纪阳 徐树公

(中科院计算所 CAD 开放实验室, 华中理工大学 武汉 420074)

一、什么是人工生命

人工生命(Artificial Life, 简称 A—Life)是 80 年代后期国外开始兴起的新的学科领域,也是计算机科学继人工智能之后出现的新的发展方向之一。目前美国、西欧各国和日本均十分重视人工生命的研究,推出了一系列成果。人工生命作为新的计算模型和智能模型,是当前计算智能领域的关键技术之一。因此,在我国该项目的研究也受到国家 863 高科技智能计算机项目的重视与支持。

人工生命的研究,最早可追溯到 50 年代几位科学家的开创性研究,如薛定谔的《生命是什么》、冯·诺意曼的《细胞自动机》、维纳的《控制论》。随后,60 年代霍兰的《基因算法》、林登迈尔的《L 系统(植物形态学)理论》,70 年代曼德勃罗的《混沌学》和《分形理论》、康韦的《生命游戏程序》,以及 80 年代的“人工神经网络”等研究,促进了人工生命学科的确立。

首先提出“人工生命”概念的是美国洛斯·阿莫斯国家实验室非线性理论组的兰顿(C. Langton)。他在 1987 年召开的首届人工生命国际学术会议上发表其研究成果时指出,生命的特征在于具有自繁殖、进化等功能,地球上的生物只不过是生命的一种特例。只有用人工的方法,用计算机或其它智能机械制造出具有生命特征的行为并加以研究,才能揭示生命的全貌。因此,人工生命是关于再现自然界生物行为特征的人造系统的研究。人工生命以计算机为研究工具模拟生命现象,在计算机上生成类似于生命的行为,以生物分析为主对现有的生物学作出必要的补充。在人工生命的研究中,其建模是以地球上已进化的生命现象为依据的。在计算机上研制并实现那些能进行自我复制并能进化的程序,用来阐明生命的进化、生态系统的机理及生命现象所蕴含的复杂性。

为了便于比照,国外常把生物学范畴的生命称为 B—Life(Biological Life),而将人工生命称为 A—Life。

人工生命的研究是现代科学中一系列学科相互交叉与渗透的结果,其中包括计算机图形学、智能机器人技术、混沌与分形理论、生物形态学(L 系统)等。目前开展的人工生命的研究,是人类探讨“生命”这一复杂现象,阐明“复杂性”的需要。在人工生命会议上,兰顿指出,今后的科学,面临的是一个复杂的适应系统(如生态系统、脑、社会现象等)。在这种复杂系统中,科学研究中过去认为行之有效的方法,如“分治”(Devide and Conquer),将无法有效运用。“分治”是将研究对象分为组成要素,经简化后再逐一分析、各个击破,而复杂的事物与对象,只能按其复杂的面貌来认识与理解。这是人们研究人工生命的更本质的一个原因。

二、人工生命的内容和研究方法

目前,国外对人工生命的研究,主要是计算机模拟,研究工作从科学和工程两个不同的侧面开展。

1. 在计算机上,以能够自我复制并能进化的算法为工具,探讨生命进化、生态系统的复杂性和人工生命的计算理论与模型。

2. 将研究的成果用于解决科学与工程实际问题。在生命科学领域,用以阐明生物的进化和生态系统的形成;在工程领域,用以研制自适应机器人(适应环境)和进行工程设计。还可用于研究社会现象和经济行为,探讨其复杂性。

以上提到“复杂性”与“生命进化”等术语。在科学上,对于什么是“生命”,目前还没有明确的定义。迄今,人们在解释什么是生命时,经常列举生命的一系列特征,如新陈代谢、运动、繁殖、环境适应、进化、信息存贮,以及时间、空间的模式识别,等等。另外,为了理解什么是“生命”,还有必要首先弄清楚

“复杂性”的概念。研究复杂性是现代非线性科学的任务之一。在生物的群体行为和人类社会中,微观行为与宏观秩序之间,存在着既互相影响又互相制约的关系。人类大脑接受与理解知识的过程也是如此。可以把神经细胞视为群体的成员,而把神经网络视为这些成员组成的社会。人脑获取外界信息时,一方面,大脑动员已存的信息;另一方面,还吸取新的信息,自律地生成新的信息结构。每个神经细胞都对神经网络产生一定的影响,同时后者又控制各种细胞的行为。从以上的信息系统的动态过程可以看出,无论是人类大脑、进化,乃至人类社会,它们在本质上是存在着共性的。

人工生命的研究领域非常宽广,它涉及生命的起源,分子的自聚合,胚胎的发生、成长、发育、分化,自复制,动物的行为,昆虫集团的动力学,生物进化、种的形成,生态动力学和语言—社会—文化的进化等生物现象。从中找出具有共性的东西,并在计算机上人工地再现。通过研究,不仅可以再现地球上现存的生命体,而且可以进一步探索未知及可能的生命存在的形式。

为了更具体地说明人工生命的研究内容和方法,下面列举 1987、1990、1992 年召开的三届人工生命会议以及 1991 和 1993 年召开的欧洲人工生命会议上发表的研究课题和进展:

1. 用计算机模拟生物的进化,阐明生物的多样性和生物的智能行为。
2. 用计算机探讨并再现那些在简单的生存竞争法则下形成和保持平衡的生态系统。
3. 研究生物的群体行为,并用之于计算机图形学及智能机器人的群体生成与控制。
4. 研究真实环境中的某些智能行为,以提高机器人的功能,使之适应复杂条件下的作业要求。
5. 研究生物型生产系统的建模理论和自适应环境的柔性机械。
6. 人工生命方法与灵境技术或虚拟真实(Virtual Reality)结合,用以模拟某些生物的生态。
7. 人工生命的美学研究,将生物进化与繁殖的机制渗入计算机图形学算法中,使生成的计算机图形更具活力,创作出富有生命气息的计算机艺术作品,或用于空间音乐和可视音乐的研究。

国外研究人工生命的近期目标是:在 90 年代“脑的十年”(Decade of the Brain)中,主要是模拟人工脑。为了实现这一目标,日本生体信息技术研究所在一项研究计划中提出了一系列任务,包括研制能进行自我复制、学习与适应以及自律生成的理论和模型,并实现相应的算法;研制并行、分布式的软硬件,以实现预期目标。

综上所述,可将人工生命所需解决的核心问题归纳为以下几点:

1. 生物的形态生成模型,主要是 L 系统(或 DiL 系统)。
2. 生物体的进化模型,主要发展基因(遗传)算法(Genetic Algorithm)。
3. 学习模型,如人工神经网络(ANN)与增强学习等。
4. 自增殖细胞自动机模型(Cellular Automata)。
5. 生物的免疫系统模型以及细胞之间的通讯通道研究。
6. 生物体的自修复模型及相应的算法。
7. 生物体内部的并行、分布式处理模型。

在研究方法上,可将人工生命的研究分为两类。一类是生物体的内部系统,如脑神经系统、免疫系统、遗传系统、酶与代谢系统等;另一类是生物的群体及其宏观行为等外部系统,包括生物的环境适应系统、进化系统等。

人工生命的研究离不开计算机,其生成算法尽管种类很多,但不外乎以下两种方法:

1. 构形法:根据生物的内部或外部系统显示的生命行为构造其计算机模型并在计算机上模拟实现。作为生物内部系统的构形法有:免疫网络、ANN;作为外部系统(种群)的构形法,其典型代表就是基因算法。
2. 动作原理法:由于生命行为经常呈现自律分布式的非线性特征,作为内部系统的有生物分形与混沌及自组织现象;作为外部系统的有自组织及协调分布式处理等。

三、人工生命的计算模型与关键算法

人工生命是一门迅速发展的学科,通过研究出了一系列建模理论与方法,其中最重要的有以下几种:

1. 自增殖细胞自动机理论与模型

细胞自动机理论在人工生命中可用来模拟生物系统中的某些自组织现象,这是冯·诺意曼的另一重要贡献。尽管这一理论是 50 年代提出的,但其作用却经久不衰,后来又有很大发展。它也是研究现代非线性科学,探讨复杂性的重要工具之一。1983 年,理论物理学家沃尔弗拉姆(S. Wolfram)注意到,植物的生长和湍流等都是由一些简单的组元所构成。这些组元虽然简单,但它们的组合形态与系统的行为则非常复杂,以至于不能简化为数学描述。因此,有必要采用新的理论模型,即自增殖细胞自动机模型。细胞自动机理论目前大多是借助于计算机模拟方法来进行研究,采用的是一种布尔型的离散化状态描述。它可以用来分析与求解各种偏微分方程,特别是那些科学与工程中难以求解或数值

求解工作量非常大的问题。细胞自动机利用自增殖将一个复杂问题用简单的模型来描述。在自增殖细胞自动机中,给出的是一个离散的网格,其结点称为细胞(或单元)。每个细胞的状态可取0或1,也可以设置其它有限种状态。随着时间的推移,每个细胞(网格点)的取值,遵循同一决定性的规则随时间而发生变化,其变化一般只取决于该点邻近的几个细胞的状态。这样,可使整个单元阵列由原来的某种静寂(死的)状态变成活的状态,从而实现“活的”细胞单元的繁衍与增殖。在模拟中,计算机可以从生物的遗传密码开始,跟踪其发育过程中的每一步,模拟产生多种多样的生物组织形式。

利用自增殖细胞自动机能否制造出复杂的组织构造呢?这个问题目前已经引起人们的广泛兴趣,因为任何生物,不外乎是由几十根DNA通过反复复制而形成。冯·诺意曼自增殖细胞自动机的一个实验性系统中,采用29种不同的状态,由20万个细胞单元组成。

本期封面和封底彩图为我们关于细胞自动机研究的部分成果照片。封底右上两幅图为从一个处于某种状态的“种子”按一定规则生长而成的复杂结构(类似于分形中的Sierpinski三角形及其各种变型)。封面和封底右下角图形是从一种随机初始状态出发,当 $t=0$ 时给出一列随机状态种子按一定规则生长而成。可以生成四种结构,即均匀结构、稳定或周期结构、复杂的局域结构和混沌结构。细胞机生成的复杂图形可用作随机数源在信息加密处理中应用。封底左图是不同初始状态下细胞自动机的演化过程。当种子数为1时,自增殖过程很快就收敛了。当种子数大于1时,生成三串图形。其中一串在有限步生长后收敛,停止增殖过程。其余两串呈现任意形状。计算具有不可约性,不论种子的初始状态如何,均无法预测其收敛性,只能通过计算机模拟进行精确研究。用计算机模拟复杂性,通常称为Automata,自增殖细胞自动机在人工生命研究中可用以模拟复杂的自组织过程。计算机模拟表明,当改变初始种子数目时所得结果呈现多样性和随机性。反之,在细胞自动机中引入随机成分,结果往往可以收敛于确定的终态。

2. 人工生命的形态生成模型——L系统

L系统是1968年丹麦植物学家林登迈尔在研究藻类植物(红藻)的生长过程时提出的一种模型,是以自动机为基础而建立的理论。上述诺意曼细胞自动机主要是描述细胞增殖的理论模型,而L系统的增殖主要着眼于生物形态的生成。L系统的创造性在于,它形式化地描述了植物或其它生物的结构特征和生长过程。林登迈尔的开创性工作也奠定了形式语言学的基础。1984年史密斯(A. R. Smith)、1986年普鲁斯米凯维奇(P. Prusmikevicz)将L系

统的理论和计算机图形学相结合,在计算机上模拟生成各种形态的植物的几何与拓扑结构。

L系统的内容很丰富,对于用计算机生成植物形态来说,比较适用且简单的有 i 阶L系统又称 DiL 系统,其中 i 取0或2,相应的有 $D0L$ 和 $D2L$ 系统。2阶L系统($D2L$ 系统)是可增殖的、前后相关的L系统。而0阶L系统是字符前后不相关的L系统,其生成规则所产生的链语言表示植物分枝形态之间的关系比较直观,常用来产生灌木和藻类等低级植物。 i 阶L系统可用下面的三元式表示:

$$DiL = \langle G_+, W, p_i \rangle$$

其中, $G = \{F, T\}$ 为基本字符集; $G_+ = \{[,]\}UG$ 为扩充字符集; $W=F$ 为起始字符元; p_i 为 DiL 的生成规则集。

3. 人工生命的进化模型——基因算法

基因(遗传)算法(GA)是美国密歇根大学心理学与计算机科学系教授霍兰德(J. H. Holland)首先提出的,它是人工生命研究中的重要算法之一,是建立在达尔文进化论基础上的一套理论与方法。该算法将生物的进化过程加以抽象,利用进化的机理,通过培育程序,用以解决那些结构尚未完全了解的问题。基因算法以自然界中活的有机体为模仿对象,通过进化和自然选择解决各类复杂问题,其结果有可能揭示生命及智能在自然界进化的细节。

根据达尔文的进化论,地球上的生物,在漫长的岁月中,通过自然选择、突变、杂交、隔离等过程,逐渐地由简单、低级生物发展成高级生物。生命是一场适者生存的斗争。基因算法就是仿效自然界的进化,模拟生物在进化过程中的自然选择。基因算法具有以下特点:

1. 具有隐含的并行性,因此,它易于采用并行处理以提高算法的速度。
2. 具有潜在的学习能力,适用于具有自适应与学习能力的系统。
3. 能把注意力集中于解空间中期望值最高的部分,发掘出目标区域。
4. 适于处理各类非线性问题,并能有效地解决传统方法难以解决的某些问题。

基因算法在下列领域中具有潜在的应用前景:

1. 用于模糊图象的复原。如日本理光公司的一项研究中,把图象的一部分作为遗传基因,通过基因算法,使它产生突变,将不需要的图象进行淘汰,并将所需图象复原。
2. 模拟生态系统的形成、生存竞争与生物进化。如日本富士通公司和美国Apple公司研究了一种类似多角体的人工生物,用基因算法使之进化,最后得到能趋利避害和既能互相攻击也能合作的群体。
3. 用以优化现有的程序。如日本创价大学在服

装设计(衣饰)和图案设计中,用基因算法将计算机图形程序,由最初的简单设计,通过杂交、突变等过程,进化为具有最佳审美配合的设计方案。

4. 人工生命的其它模型,如人工神经网络和混沌分形模型等均有广泛的应用。限于篇幅,此处从略。

四、人工生命的意义及应用发展前景

人工生命是以计算机为媒介模拟生命或具有生命特征的行为,包括那些自组织、自学习以及信息的复制与传播行为。因此,人工生命的研究是用计算机来构造生命特有的自律行为的人工系统的研究。其意义在于:

1. 由于人工生命是计算机与生命科学及有关学科相互渗透与交叉形成的,因此研究人工生命的计算理论与算法,有助于揭示生命的全貌,及探索生命的起源,为生物学研究提供新的途径。

2. 在人工生命的研究中,许多算法是仿效自然界生物体的信息处理方法(如基因算法、自增殖细胞自动机理论、人工神经网络等)。它们不仅效率高,而且可应用于科学和工程的许多实际领域,产生巨大的经济效益。

3. 在90年代“脑的十年”中,研究人工生命的理论与算法,有助于揭示脑的机制与本质并发展相应的领域。

在展示的人工生命成果中,令人注目的是美国Lits Compute公司在Macintosh计算机上研制的称之为AQUAZONE的软件系统。它再现了人工鱼及其生态环境。屏幕上显示的是一个饲养热带鱼的水槽,鱼的成长、产卵、孵化,鱼在水中的来回游动等,都可以实时地、以图像/声音等方式动态地展示在眼前。参观者可以对“鱼缸”进行“投饵”,“鱼群”立即游过来吞食、戏耍,一切都和真实的一样。一旦切断计算机鱼缸的电源,生态系统的模拟仍可维持1~2天。接通电源后,则水槽中的生态系统又恢复

如初。如果“鱼”生了病,则显出无精打彩,呆在槽底不动。如果几天不“投饵”,饿死的“鱼”便漂浮在水槽上部。尽管这是用计算机再现的生态系统,然而给人以活灵活现的真实感觉。从这一意义上说,人工生命的研究,可以认为是灵境技术或虚拟真实的继续。

人工生命的应用,主要是将生物的进化和适应的机理用之于科学与工程的有关领域。目前,人工生命用于机器人和人工生命美学领域已展现出巨大潜力和广阔前景。

在机器人研究方面,日本索尼公司计算机科学研究所、日立研究所等,都在开展高效率信息处理及自律机器人的研制,研究多台自律机器人的配合,开发能进行分裂与繁殖的机械。美国MIT的媒体实验室研制了类似于昆虫的六足机器人,采用了类似于条件反射的控制方法,无需高性能的处理装置即能适应环境。在自然界,并不聪明的昆虫,却能适应复杂的环境。在下一代机器人的研制中,将研究如何效法昆虫的智慧,用简单功能的机器人来完成复杂的作业。对自律机器人的研究,开创了在真实环境下研究智能行为的先河。东京大学正在进行类似的研究。当给出几条指令时,多台机器人同时工作,而不相互干扰。这种群体行为可以认为是一种智能行为。

传统的生物学是一门描述性的学科,它通过分类、列举性状并在宏观与显微镜的水平上观察其结构。在研究中,科学家接触到的主要是生命过程的结果。分子生物学的发展使人类对生命现象有了更深层的认识,它们描述、控制分子的能力已使生物学研究发生了根本变革。描述生命现象、解释生命复杂性并弄清其机制,是许多科学家为之奋斗的目标。人工生命的研究提供了用计算机再现生态系统及其复杂性、模拟进化过程的有效方法,使之有可能对“未知、未来”的生命现象进行探讨。从这一方面考虑,计算机人工生命的研究也许具有更深远的意义。

(责任编辑 肖庆山)

(上接第54页)

欠发达地区投资比重逐步增到2/3。重大项目地方配套资金比重不应一刀切,原则上人均GDP水平低的地区,地方配套资金就应低一些。

(3)帮助贫困地区发展铁路、公路等基础设施。除了中央专项拨款和地方配套资金外,还可考虑开辟新的建设资金来源。如每年从财政部发行的上千亿元国库券中划出一定比例,设立欠发达地区交通建设专项基金,由受益的地方财政偿还本金,中央财政偿还利息。建议中央支援建设经过毕节联通成渝线、滇黔线与桂昆线的铁道干线,形成川、滇、黔

三省通向北部湾海域的大动脉。

(4)在新增教育财政拨款中,设立贫困地区基础教育专项基金,专用于补助基础教育设施、教师工资、学生学习必需品、失学儿童学费等。同时设立少数民族培训专项基金,用于培养少数民族干部和专业人员。还可参照世界银行援助项目管理办法,从扶贫专项款中划出一定比例的培训费,免费或低收费对农民进行技术培训和职业培训。

(中国科学院供稿)