

基于生长神经网络的进化机器人行为研究

王洪燕, 杨敬安, 蒋 培

(合肥工业大学人工智能研究所, 合肥 230009)

摘 要: 提出了一个基于生长神经网络的进化机器人行为算法, 新算法的主要特点是: 1、通过自然选择对神经网络进行进化, 并能自主实现机器人避碰、移动、复制和攻击等行为; 2、开发了一个自主机器人模拟环境, 对所提出的算法进行运行测试。模拟结果证明, 生长神经网络系统是研究渐增进化的有效工具, 新算法能够有效地实现机器人创新行为。

关键词: 自然选择; 生长神经网络; L-系统; 进化算法; 创新行为

中图分类号: TP24 **文献标识码:** A **文章编号:** 0372-2112 (2000) 12-0041-04

Research on Evolutionary Robot Behavior Using Developmental Network

WANG Hong-yan, YANG Jing-an, JIANG Pei

(Artificial Intelligence Institute, Hefei University of Technology, Hefei 230009, China)

Abstract: This paper propose an evolutionary robot behavior algorithm which is based on developmental network. The new algorithm use nature selection to evolve neural network and can autonomously avoid obstacles, move, reproduce and attack. A simulation environment is developed to test the new algorithm. The experimental results demonstrate that the developed system is well suited to long-term incremental evolution and original behaviors evolution.

Key words: natural selection; developmental network; L-system; evolutionary algorithm; original behavior

1 引言

人工生命的一个重要课题是如何在机器人上重现自然生命, 但由于对自然生命理解的局限性, 人们还不能很好地做到这一点。因此, 当前的任务是, 增进对自然生命的理解并更好地实现它或者产生一个能超越其预定义的系统。现有的研究方法有两类: 第一类为传统的“自顶向下”的方法, 这种方法已证明不适合于 AI, 显然也不适于人工生命; 第二类方法是渐增地自底向上构造自主系统, 目的是增加对生命的理解, 并通过建模, 在目标模拟环境中对其进行测试, 这种方法所产生的系统能超越其预定义, 并可产生能与简单自然生命相类比的行为。在自然界中, 进化是没有显式评价函数的, 通过个体与环境间的交互(包括相似个体间的交互), 某些行为会明显优于另外一些。通过这种渐增式的交互, 能够产生新的结构和行为, 这种机制也被称为突现(emergence)。

由于人工进化所面临的问题愈来愈复杂, 因而选取合适的评价函数就越来越困难^[1]。同时, 诸如 Tierra^[2]和 Poly-World^[3]等原型系统却越来越呈现出自然选择的优越性。人工选择意指通过人工引入评价函数指导进化选择, 而自然选择并不需要如此。采用自然选择是进化突现的必要条件, 为产生突现, 还必须解决“进化对象设计”问题^[4]。

因为神经网络具有优良的容错性能, 成为进化机器人研究的当然选择。但是网络结构的确是一个困难问题, 创新行为的进化需要新的网络结构。用于创新行为进化的网络结构应更为自然, 而不是通常的层次分明的标准结构。现有的研究工作主要包括如下几个方面: (1) 胞元编码: Gruau 使用遗传编程^[5]进化胞元编程语言 (cellular programming language)^[6], 并进一步进化模块人工神经网络。(2) 格构自动机: 许多研究者已开始利用传统的格构自动机 (CA) 来构造神经网络。但是, 这方面的工作大部分集中在神经元生长的层次上, 而不是开发整个神经网络。(3) 传播模型: 尽管有几项有关进化神经网络的研究工作采用了使用沿浓度 (concentration) 梯度的传播来决定网络的生长, 但所产生的网络结构的规模较小。(4) Lindenmayer 系统: Kitano 利用一个上下文无关的 L-系统^[7]来进化网络连接矩阵^[8], 基因型所含的规则个数是可变的。Boers 和 Kuiper 利用一个上下文敏感的 L-系统进化前馈神经网络的结构^[9]。所得结果似乎更证明了反向传播而不是进化的有效性。

关于长期进化的“种族自适应遗传算法”(SAGA) 理论^[10]认为, 基因长度改变的速率应远低于染色体操作的交叉算子的速率。群体应以种族为单位进化, 群体的最优状态为准收敛 (nearly-converged) 状态。SAGA 理论也为我们进行渐增长期

进化研究提供了理论依据和方法。

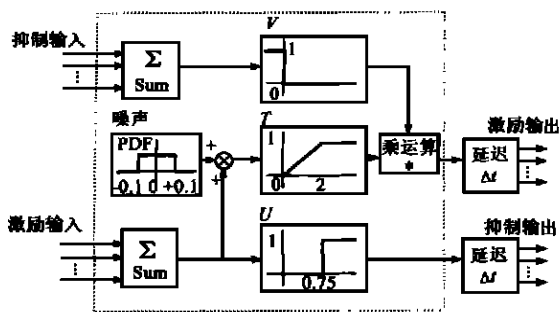


图1 神经元结构框图

本文在现有进化目标模型基础上,提出了生长神经网络的详细生长和解码算法,然后给出了神经网络进化方法和一个较复杂的模拟环境,模拟环境对避碰、移动、复制和攻击等行为进行了仿真,最后对模拟结果进行了分析。结果证明,所提出的方法能有效地实现非预指定的机器人创新行为并能成功实现渐增进化。

2 生长神经网络

基因理论认为,在任一个体生长的任何阶段,仅有基因的一小部分被利用,基因的效用则依赖于个体的局部环境,因此我们考虑用模块神经网络控制机器人的行为。对于创新功能的神经网络模块,在进一步的进化过程中,为了不失去其已有的功能,就要求要么产生一个新的模块,要么存在一个空闲模块能被利用和改进,这可以通过基因复制或网络生长来完成。因为一个非模块编码的网络连接所需存储容量巨大,网络结构的直接复制方法是不可行的。为了有效地进化复杂的实现行为,需要采用神经网络模块化生长方法。

本文所用的神经网络为反馈神经网络,其神经元结构框图如图1。为进化神经网络,我们设计了一个上下文无关的L-系统。所产生的系统的子网络与父网络是相似的,每一个节点都有一个位串标签(label)与之相关,这些标签在节点产生时初始化,并随生长的进程而改变,其长度可以是任一非零值。一个节点可以是网络输入节点或网络输出节点,也可能为中间节点,节点类型可在生长过程中确定。

网络的生长是依据产生式规则的匹配来进行的。若存在一产生式规则的前件(Predessor)与节点的标签起始部分字符相同,则认为此规则与此节点相匹配。匹配的字符越长表示匹配度越高,通常总是选择应用匹配度最高的规则。

产生式规则包括一个前件、两个后件(Successor)和连接位串编码,其一般形式如下:

$$P \rightarrow S_r, S_n; b_1, b_2, b_3, b_4, b_5, b_6$$

上式中,

P 前件:初始节点的字符位串

S_r 后件1:替换初始节点的字符

S_n 后件2:新节点的字符

(b_1, b_2) S_n 的输入输出连接的类型/0表示无,1表示有/

(b_3, b_4) S_r 到 S_n 的连接

(b_5, b_6) S_n 到 S_r 的连接

若一后件的字符为空(长度为0),则就不产生对应节点。若存在替换后件1,则与原前件具有相同的连接权,只是节点字符不同。而新后件2为一新产生的节点,若不存在由旧节点到新节点的输入(b_3 或 b_4)连接,则新节点将继承旧节点的输入权;若不存在由新节点到旧节点的输出(b_5 或 b_6)连接,它将继续旧节点的输出权。新的网络输入节点仅产生于网络的输入节点,新的输出节点仅产生于输出节点。基于字符的网络输入与输出匹配保证了在生长或进化过程中增加或删除节点不会破坏先前修改的网络输入与输出间的关系。

产生式规则的遗传解码总是试图读取一条始于当前起始位的产生式规则。一有效规则总是由11开始,并需要足够的长度来保证规则的完整性。为了读取一条规则,引入了段(segment)的概念,段是一个奇数位全为零的位串。读取一个段的过程如下:读取当前位,若为1则停止并获得一个段;否则读下一位,此位即为本段的下一信息位;然后依次跟踪段的信息位。若一个段不含任何信息位,则为空串。

要读取一个规则,首先为每个前件、后件1(替换节点)和后件2(新节点)读入一个段;然后再依次读入6个连接信息位。例如:

Genotype: 1 1 1 0 1 1 0 0 1 0 1 1 1 0 1 1 0 0

Decoding: + + + → -1- * -0- * 0 1 1 1 0 1

+ + + -1- → -0- * 0 1 1 1 0 1 1 0 0

Rules: 1. $P \rightarrow S_r, S_n, \text{link bits}$

Any → 1, 0, 0 1 1 1 0 1

2. $P \rightarrow S_r, S_n, \text{link bits}$

1 → 0, 1, 1 0 1 1 0 0

3 模拟仿真

为实现并验证所提出的方法,本文开发了一个机器人模拟系统“Life”。“Life”为二维的模拟机器人环境,每个个体由神经网络控制,网络则由生长系统进化。进化采用自然选择,进化的整个过程完全由个体自主控制。

Life环境(图2)

由单位方格组成,大小为20×20。环境中包含个体和障碍物,在任一时刻,一个方格内仅可容纳一个个体,而一个障碍物最多可以占用4个单元。障碍物的上限为10。个体可以在世界中的方格内及方

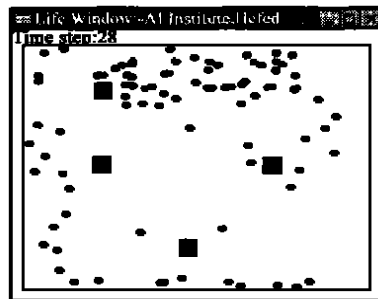


图2 模拟环境

格间自由移动,也可以进入障碍区,但会因此受到惩罚。每个个体除了具有位置参数,还具有一个头方向,其初始设置是随机的。个体由实心圆弧表示,圆弧的缺口部分为头部,黑色代表普通个体,浅色代表优化个体,障碍物用黑色实心矩形表

示.

3.1 算法流程

初始化 在模拟环境中生成随机障碍物,在环境的单元中对个体进行初始化,初始个体的基因型仅含一个位'0'.

主循环 在每个时间步(循环),所有存活的个体都被处理一次,个体的处理次序则是随机的.个体的处理流程包括如下几步:(1)更新网络输入;(2)网络生长;(3)更新所有的神经元激励,得到网络输出;(4)根据网络输出,执行相关个体行为;这些行为包括避碰、复制、攻击、顺时针转动、逆时针转动、向前移动.

3.2 神经网络结构

原子网络包括三个节点和两个连接:网络输入 0010-000-010 网络输出.网络输出节点的字符(010)与复制行为相匹配;网络输入节点的字符(左端输入的 010 部分)也与此相匹配;隐节点不含匹配的节点或行为字符.尽管生长过程中对节点和连接的数目进行了限制,自主生长过程贯穿个体的整个生命期.

3.3 个体与环境间的交互

每个个体含六个内置的行为,每一个行为都由网络输出节点生成,节点的起始位串为:

- 010 * 试图复制位于当前个体前方的个体
- 011 * 避开前方的障碍物(向后移动或静止)
- 100 * 攻击并消灭位于当前个体前方的个体(若存在)
- 101 * 逆时针转动
- 110 * 顺时针转动
- 111 * 向前运动(若前方无个体挡路)

例如,若一个网络输出节点含字符 1101001,个体将沿顺时针旋转,其转角与节点的激励输出成比例.如果一种行为有两个输出节点与之匹配,则相关输出是这些节点激励输出的和的标准平均值.若一个网络不含与各种行为匹配的输出节点,则认为网络输出为噪声.

复制与攻击都是二元行为,若与之相关的网络输出超出一定的阈值,这些行为将被激活,若前方的邻单元为空时,这些行为无效.转动和向前移动也与输出成正比.障碍物则被视为输出为 100 的个体处理,若个体进入障碍物 3 次则自然死亡.

当一个个体试图复制其前方的个体时,若前方个体下方的单元为空时,新个体将在此空单元中产生;若不为空,新个体将取代前方个体的位置;若前方个体正处在战斗状态且复制可能导致前方个体被取代时,则取消复制.与复制相关的两个算子为交叉和变异,交叉算子也是基因长度可变的原因,变异算子则为单位变异操作.

个体控制网络的输入节点的输入值为相邻个体的“匹配”输出节点的激励输出值的加权和.若一输入节点字符的首位为 1,则其输入来源于其右方的个体(包括右前、右后),否则来源于左方的个体.一个输入节点和输出节点匹配的充分条件是:输入节点的从第二位开始的字符与输出节点的起始部分字符相同.例如,字符为 10011 的节点仅匹配于右方个体网络的以 0011 开始的输出节点.输入值加权的权系数反比于个

体间的欧氏距离,当前的输入邻域为以当前个体为中心的 5×5 的区域.

4 结果分析

图 3 给出系统运行 2412 个时间步后的仿真结果.使用本文算法,两个个体交配所产生的子个体的网络结构与父个体的网络结构总存在一定的相似性.在进化的任一刻检查群体,可以发现总有许多个体具有一定的相似性.群体保持一种准-收敛状态,并贯穿于整个进化过程.生长系统总存在一个充分相关的(隐含的)适应值曲面,这一点对长期进化是非常必要的.

进化开始,存在一段时间进行基因长度的生长直到基因至少包含一个产生式规则.在以下的 1000 到 2000 时间步内,网络会产生非常简明的策略如“执行所有的行为”或“一直向前走并攻击前方的个体”,这

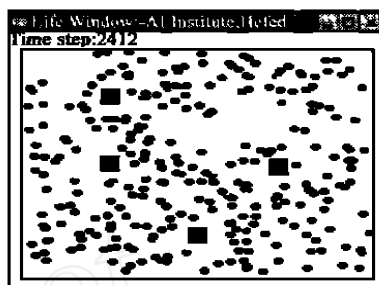


图 3 模拟结果(运行 2412 个时间步)

种个体在群体中占主导地位,还有一些个体比其它的个体要优化一些,但并不占主导地位.

在每一次运行中,首先突现产生的占主导地位的种族个体具有如下特征:避开障碍物,单向转动并试图同时执行攻击和复制行为.图 4 给出了一个这样个体的例子,其输出 $o101$, $o011$, $o010$ 和 $o100$ 代表逆时针转动、避碰、复制和战斗等行为.网络中存在大量的权连通输入和输出节点,输入节点字符与同一网络中“无任何行动”行为输出字符相匹配($o000$, $o00$),这类个体使用相邻个体作为激励源.

L Rule: <0>->0110*1001*000000

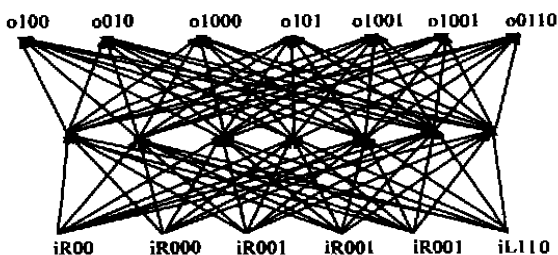


图 4 个体-1 控制网络

正是因为这一占主导地位的个体的存在,模拟世界中总存在足够的空间容纳其他的个体.环境中其它个体一般寿命不长,它们随时都有可能撞到占主导地位的个体而被消灭.但是它们也有可能发展成为一类优秀的个体并具有一直前行并试图攻击或复制行为,这种个体可见图 5.

图 6 给出了运行结果平均性能曲线,从个体复制与攻击次数的平均值上来看,系统可以突现产生更复杂的种族.

5 小结

本文提出了有效创新行为的进化算法,并建立了仿真系

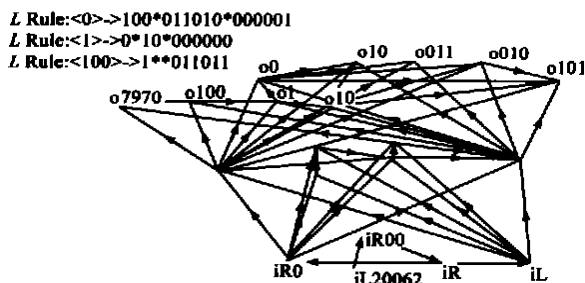


图5 个体-2 控制网络

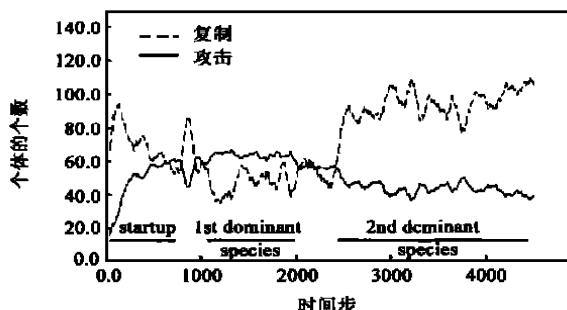


图6 运行结果平均性能曲线

统,系统成功地产生了多机器人控制网络并实现了机器人避碰、移动、复制和攻击等行为.由于采用了生长系统,这些网络在复制过程中克服了传统进化网络的基因型相同而表现型不同的固有缺点.模拟结果证明系统能在长期渐增进化中处于准-收敛状态.进一步的研究方向是,从效率上对算法进行分析,优化进化算法.

参考文献:

- [1] Zaera N., Cliff D. and Bruten J. Evolving collective behaviours in synthetic fish [J]. In Maes P., Mataric M., Meyer J. A. eds., Proceedings of SAB96. MIT Press Bradford Books., 1996:635 - 644.
- [2] Ray T. S. An approach to the synthesis of life [J]. In Langton C., Taylor C., Farmer J. and Rasmussen S. eds., Artificial Life II. Redwood City, CA: Addison-Wesley., 1991:371 - 408.

(上接第47页)

参考文献:

- [1] Y. Bar-Shalom, E. Tes., Tracking in a cluttered environment with probability data association [J]. Automatica, 1975, 9(11):451 - 460.
- [2] H. A. P. Blom and Y. Bar-Shalom. The interacting multiplemodel algorithm for system markovian switching coefficients [J]. IEEE Trans. 1988, AC-8(33):780 - 783.
- [3] M. R. Garey and D. S. Johnson. Computers and Intractability: A Guide to the Theory of NP-Completeness [M]. San Francisco, CA: Freeman 1979.
- [4] C. H. Papadimitriou and K. Steiglitz. Combinatorial optimization: algorithms and complexity [Z]. Englewood Cliffs, NJ: Prentice-Hall, 1982.

- [3] Yaeger L. Computational genetics, physiology, metabolism, neural systems, learning, vision, and behavior or polyworld: Life in a new context [J]. In Langton C. G. ed., Artificial Life III., 1993:263 - 298.
- [4] Channon A. and Damper R. Perpetuating evolutionary emergence [J]. In Proceedings of SAB98. MIT Press. 1998.
- [5] Koza J. R. Genetic Programming [J]. Cambridge, MA: MIT Press/Bradford Books. 1992. Gruau F. Artificial cellular development in optimization and compilation [R]. Technical report, Psychology department, Stanford University, Palo Alto, CA. 1997.
- [6] Lindenmayer A. Mathematical models for cellular interaction in development [J]. Journal of Theoretical Biology 18. Parts I and II., 1968:280 - 315.
- [7] Kitano H. Designing neural networks using genetic algorithms with graph generation system [J]. Complex Systems., 1990, 4:461 - 476.
- [8] Boers E. J. and Kuiper H. Biological metaphors and the design of modular artificial neural networks [J]. Master's thesis, Departments of Computer Science and Experimental Psychology, Leiden University. 1997.
- [9] Harvey I. Evolutionary robotics and SAGA: the case for hill crawling and tournament selection [J]. In Langton C. G. ed., Artificial Life III. 1993.

作者简介:



王洪燕 1971年生,1995年毕业于合肥工业大学,获硕士学位,现为合肥工业大学计算机学院博士生,主要研究兴趣为进化计算、人工智能。

杨敬安 合肥工业大学人工智能所教授,博士生导师,IEEE高级会员,纽约科学院 Fellow Academy Member. 主要研究兴趣为图像理解,模式识别,计算机视觉,人工智能与机器人以及多传感器融合与集成。

- [5] Bertsekas D. P. and Castanon D. A. The auction algorithm for the transportation problem [J]. Annals Operat. Res., 20, 1989:67 - 96.
- [6] C. L. Morefield. Application of 0-1 integer programming to multitarget tracking problems [J]. IEEE Trans. 1977, AC-6(22):302 - 312.
- [7] K. R. Pattipati, S. Deb, Y. Bar-Shalom and R. B. Washburn. A new relaxation algorithm and passive sensor data association [J]. IEEE Trans. 1992, AC-2(37):198 - 213.
- [8] Y. Bar-Shalom, and Li X. R. Multitarget-Multisensor Tracking: Principles and Techniques [M]. Storrs, CT: YBS Publishing, 1995.
- [9] S. Deb, M. Yeddanapudi, K. Pattipati and Y. Bar-Shalom. A generalized S-D assignment algorithm for Multisensor-Multitarget state estimation [J]. IEEE Trans. 1997, AES-2(33):523 - 536.