

文章编号: 1001-0920(2001)03-257-06

基于进化计算的神经网络设计与实现

何永勇¹, 褚福磊¹, 钟秉林²

(1. 清华大学 精仪系, 北京 100084; 2 教育部 高教司, 北京 100816)

摘 要: 基于进化算法可有效解决神经网络设计和实现中存在的一些问题, 使网络具有更优的性能。在此对基于进化计算的神经网络设计和实现的研究内容及进展情况进行综述, 讨论了网络实现的关键问题, 包括网络权重的进化训练, 网络结构进化设计, 学习规则进化选取以及进化操作算子设计等, 并分析了相关的研究和发展方向。

关键词: 进化计算; 神经网络; 优化算法

中图分类号: TP 18

文献标识码: A

Artificial Neural Networks Design and Implementation Based on Evolutionary Computation

H E Yong-yong¹, CHU Fu-lei¹, ZHON G B ing-lin²

(1. Department of Precision Instruments, Tsinghua University, Beijing 100084, China;

2 Bureau of Higher Education, Ministry of Education, Beijing 100816, China)

Abstract: Design and implementation of artificial neural networks based on evolutionary computation can lead to significantly better network performance. A review on evolutionary computation based neural networks design and implementation is presented in details and some related aspects are discussed emphatically. Among them are the evolving training of network weights, the evolving design of network architecture, the evolving design of learning rule, different search operators, and the possible future research directions in this area.

Key words: evolutionary computation; neural networks; optimization algorithm

1 引 言

人工神经网络^[1]可看作由一组内联的神经元组成的高度并联模型, 其性能由拓扑结构、节点传递函数和节点间联结权重唯一确定。因此, 神经网络实现的关键是网络结构构建、学习算法设计和学习规则选取。这些也是神经网络实现中难以满意解决的问题, 如何有效地解决这些问题一直是神经网络研究的主要内容和方向。

进化计算(EC)是一种基于群体的全局随机优化算法, 主要包括遗传算法(GA)^[2]、进化规划(EP)^[3]和进化策略(ES)^[4]3种算法。这3种算法在实现细节上各有侧重, 但算法实质相同, 即都直接借鉴了生物进化和自然遗传的基本思想, 通过群体中个体间相互竞争和信息交换进化出适应度高的个体。进化计算采用随机变换规则, 以目标函数的值信息指导进化搜索, 无需目标函数的导数信息或其它

收稿日期: 2000-04-17; 修回日期: 2000-07-06

作者简介: 何永勇(1969—), 男, 重庆梁平人, 讲师, 博士, 从事智能诊断、神经网络等研究; 钟秉林(1951—), 男, 北京人, 教授, 博士生导师, 从事智能诊断、神经网络等研究。

问题相关的特殊信息。与传统的优化算法相比,进化计算具有极强的“鲁棒性”,特别适合处理大规模、复杂、多态、非可微,尤其是梯度信息难以获取以及优化目标函数难以定义的优化问题^[5]。

将进化计算与神经网络相结合以解决神经网络设计和实现中存在的问题已受到广泛重视。结合的基本方式是将神经网络的联结权重初始化和训练、网络结构设计、学习规则调整等网络实现问题,视为某一优化问题,并应用进化计算完成其优化。

本文主要阐述基于进化计算的神经网络设计和实现的研究内容和进展情况,着重讨论神经网络与进化计算结合的方法和策略,即基于进化计算神经网络的实现途径,包括联结权重的进化训练、网络结构的进化设计、学习规则的进化设计等,并分析了相关的研究和发展方向。

2 网络联结权重的进化训练

传统的网络权重训练一般采用基于梯度下降的算法^[1],但这易于陷入误差函数的局部极小点,尤其当误差函数为多态或非可微时,这种训练算法往往难以得到满意的训练结果,甚至训练难以进行。

利用进化计算进行网络权重训练,就是将权重训练过程看成权重在网络结构和学习任务环境中进化,并应用进化计算实现其进化的全局优化过程。进化中可根据网络训练误差和网络复杂性等因素确定适应度函数,该函数无须可微甚至无须连续,因此可有效地克服传统训练方法的缺陷。

权重进化训练的实现主要包括两方面内容:一是权重个体染色体表达方式的确定;二是根据选取的染色体表达方式设计遗传操作算子。不同的表达方式和操作算子将产生不同的训练性能。

2.1 权重个体染色体的二进制位串表达

标准遗传算法用二进制位串进行个体染色体表达^[2]。在网络权重进化训练的早期研究中,多采用这种表达方式^[6,7],其基本方法是将权重以一定长度的二进制位串表示,并将所有位串按一定方式串接而成权重个体的染色体。位串串接的基本原则是与同一隐节点或输出节点相联的权重在串接中彼此相邻。

个体染色体的表达精度与复杂性是一对矛盾:染色体太短,会降低权重的进化质量;染色体太长,又会使权重的进化缓慢低效。因此,在应用中一般作折衷处理。

二进制位串表达的主要优点是其简单性和一般性,且利于直接采用经典成熟的遗传操作算子(如单点杂交、均匀杂交等),无须设计特定的问题相关的操作算子,也利于神经网络的硬件实现。

2.2 权重个体染色体的实值表达

权重个体染色体的实值表达是将网络权重直接组合成一实值向量来表达染色体。对于实值表达方式,EP和ES算法更为适合,因为它们主要是基于突变操作,可避免杂交操作造成的负面影响。目前,在这方面已有大量成功的研究事例^[8,9]。在这些事例中,遗传算子主要是 Gaussian 突变算子,也有采用 Cauchy 突变算子^[10]。利用 EP 进化权重的基本步骤可归纳如下:

1) 随机产生包含 u 个个体的初始群体,且设进化次数 $k = 1$,个体为一实值向量对 (w_i, η_i) , $\forall i \in \{1, 2, \dots, u\}$, 其中 w_i 为权重向量, η_i 为参数向量(即自适应 ES 中的策略参数)。

2) 每个个体 (w_i, η_i) , $\forall i \in \{1, 2, \dots, u\}$ 分别产生自己的子代 $(w_{i(j)}, \eta_{i(j)})$, $\forall i \in \{1, 2, \dots, u\}$, 即

$$w_{i(j)} = w_i(j) + \eta_i(j)N_j(0, 1) \quad (1)$$

$$\eta_{i(j)} = \eta_i(j) \exp\{\tau N(0, 1) + \tau N_j(0, 1)\} \quad (2)$$

式中, $w_{i(j)}$, $w_i(j)$ 和 $\eta_i(j)$, $\eta_i(j)$ 分别为向量 w_i , w_i 和 η_i , η_i 的向量分量; $N(0, 1)$ 为一维正态分布的随机数; $N_j(0, 1)$ 为与向量分量相关的正态随机数,也可为 Cauchy 函数^[10]; 参数 τ 和 τ 一般设为^[11]

$$\left(\sqrt{2\sqrt{n}}\right)^{-1} \text{ 和 } \left(\sqrt{2n}\right)^{-1}, j \in \{1, 2, \dots, n\}。$$

3) 根据网络训练误差确定父代和子代个体的适应度,误差函数可依应用要求定义。

4) 随机均匀地从父代、子代中选取 q 个竞争个体,分别与该个体做适应度比较,适应度高者“获胜”;最后从父代、子代中选取“获胜”多的 u 个个体作为下一代进化个体。

5) 判断进化目标,若满足则停止进化;否则设 $k = k + 1$,并转步骤 2) 继续进化。

在应用进化计算对网络权重进化训练中,存在的主要问题是“排列”问题,亦即所谓的“协定竞争”问题,原因是权重基因型与网络显型之间是多对一的映射关系,完全相同的网络,由于隐层节点的排列不同而造成权重个体染色体的不同,这使得在进化过程中杂交操作变得低效甚至无效。

3 网络结构的进化设计

神经网络结构由其拓扑结构和节点传递函数

确定。对于一给定任务, 合理的网络结构是网络信息处理能力和网络性能的保障。因此, 网络结构的合理设计一直是神经网络研究领域的一项重要内容。

目前, 对于网络结构设计已提出了一些方法, 这些方法均可归结为结构“增构法”和“减构法”两类^[12~14]。所谓增构法, 就是网络训练从一可能的最小规模的网络开始, 训练过程中根据一定判据增加网络层数、节点及联结; 减构法则与之相反。但是, 正如文献^[15]所指出的: 这种结构“爬山”或“下山”法极易陷入结构的局部极小点, 且其搜索空间也只是整个结构空间中一个极小的子空间。因此, 由这两类方法难以得到优化合理的网络结构。

网络结构优化设计可形式化为一结构空间的优化问题。该空间中每一点代表一结构, 给定结构的性能(优化)判据, 如最低网络训练误差、最低网络复杂性等, 则在结构空间中形成一结构性能曲面, 结构的优化设计即为探索该曲面最高点的优化过程。该性能曲面一般具有无界性、离散性、多态性、非可微、复杂、含噪声、伪真等特征^[16]。显然, 进化计算正适合这类优化问题的优化求解。

网络结构的进化设计包括结构个体的染色体表达方式设计和进化方法的选取两方面内容。结构个体染色体表达主要有直接编码策略和间接编码策略两种。

3.1 直接编码策略

网络的拓扑结构可由一联结矩阵表示, 如一 $N \times N$ 矩阵 $C = (c_{ij})_{N \times N}$ 表达一个有 N 个节点的网络结构, 其中 c_{ij} 值表示从节点 i 到节点 j 的联结。 $c_{ij} = 1$, 表示节点间有联结; $c_{ij} = 0$, 表示节点间无联结。可见, 联结矩阵与网络结构间是一一对应关系。

网络结构的直接编码, 就是将联结矩阵的行或列串接形成结构个体的染色体^[16, 17]。图 1 给出一前馈网络的网络结构个体染色体直接编码的例子。网络为前馈网络, 故联结矩阵左下角全为 0 (见图 1(b))。据此可对染色体进行简化以减小其长度, 如图 1(c) 所示。

网络结构的直接编码在实现上直观, 且染色体中包含的结构信息较完备, 结构的进化更加精确细微, 易于进化出优质紧凑的网络结构。由于适应度函数的定义不受可微或连续等条件限制, 因此可基于信息理论或统计理论引入各种性能判据^[18], 以提高结构进化质量。

直接编码策略存在的一个潜在问题是: 随着网络规模的增加, 网络的联结矩阵和相应的染色体长

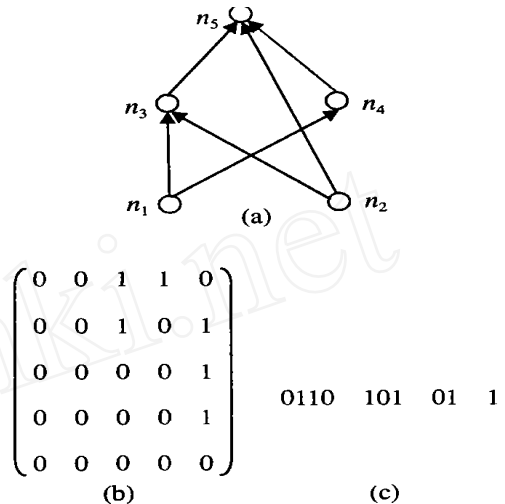


图 1 前馈网络的网络结构直接编码

(a) 网络结构

(b) 网络联结矩阵 (c) 染色体表达

度会随之增加, 从而使计算量指数增加。虽然可根据领域知识对联结矩阵施加约束以控制规模, 但实际应用中领域知识难以获取, 且搜索空间的人为限制可能丢失一些好的进化解; 同时, 结构进化设计也存在“排序”问题。因此, 许多研究者在结构进化中采用突变操作而避免杂交操作^[15, 19]。

3.2 间接编码策略

为了控制个体染色体长度, 提高进化效率, 许多研究者采用了间接编码策略^[20, 21], 即只对有关结构的重要特征进行编码, 至于其它结构细节则根据先验知识预定或通过结构“生成”规则确定。间接编码策略一般有两种实现方式, 即结构参数表达和结构生成规则表达。

3.2.1 结构参数表达

网络结构可由其结构参数唯一决定, 如隐层数、隐层节点数、层间联结等。因此, 可将这些参数进行编码, 以表达结构个体染色体^[21]。尽管参数表达方法可减小结构表达中染色体长度, 但也给搜索空间强加了约束。如若只将隐层节点编码, 则其潜在地假设了网络为单隐层前馈网络, 且两邻层间为全联结。所以, 参数表达方法适合于所需网络结构的基本类型为已知的结构设计场合。

3.2.2 结构生成规则表达

所谓结构生成规则, 就是用以构建网络结构的规则^[20]。它是一系列从粗略到精细描述结构的迭代等式, 类似于产生式系统中的产生式规则。通过这些迭代等式由粗到精的迭代, 可生成网络结构的联结

矩阵。图2给出了一个由结构生成规则构建网络结构的例子。

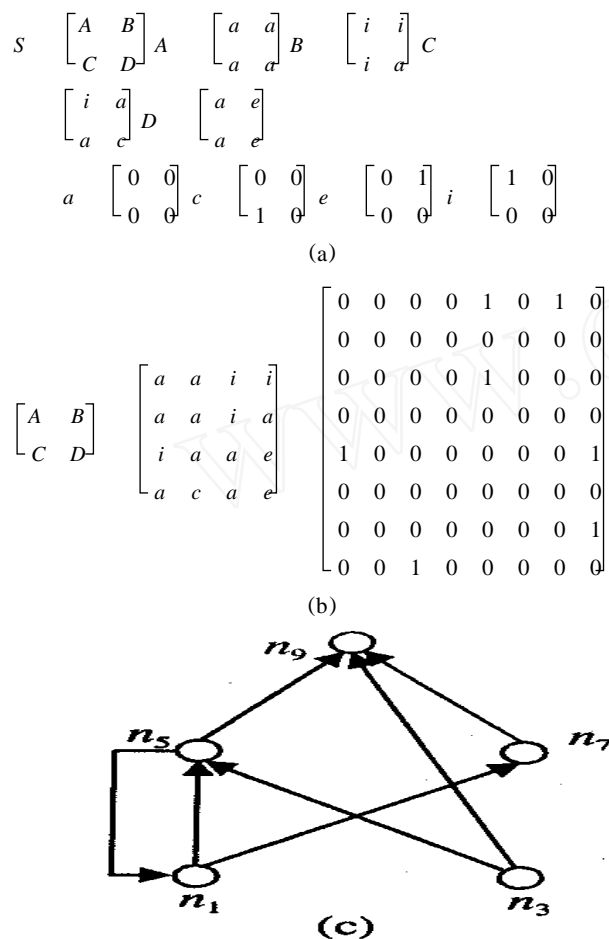


图2 结构生成规则生成及结构构建

(a) 结构生成规则

(b) 由生成规则生成结构联结矩阵

(c) 生成的网络结构

结构生成规则表达就是直接对结构生成规则进行编码。这样,结构设计便从结构优化转化为结构生成规则的优化。具体编码时可将所有规则编码为一个个体(即所谓Pitt方法),或将每条规则分别编码为个体(即所谓Michigan方法)。其基本思路是将规则中矩阵的元素顺序串接。

尽管结构生成规则表达方式在网络结构优化设计中已得到广泛应用,且取得了良好效果,但也存在一些局限性^[20]。如需预先定义迭代步数,在进化中对节点间的联结细节缺乏有力表达和进化等,且生成规则表达方法一般是将结构进化与权重进化分开处理,难以实现结构和权重的同时进化。

3.3 节点传递函数的进化

在传统的网络设计中,网络所有节点或至少同层节点采用相同的传递函数。上述结构进化设计的

讨论也只涉及其拓扑结构。大量的研究^[22]表明,节点传递函数的不同选取将直接影响网络性能,因此网络结构的进化设计应考虑节点传递函数的选取。文献[23]将节点传递函数信息融入网络拓扑结构的编码表达中,实现了节点传递函数的进化选取;文献[24]则在初始群体中设定每个个体80%的节点为Sigmoid函数,20%的节点为Gaussian函数,传递函数的进化即为两种传递函数的节点比例的优化。

3.4 网络结构和联结权重的同时进化

以上讨论的网络结构进化设计未涉及联结权重的进化学习。这种网络结构与权重分离的进化策略存在的主要问题是适应度评价时存在噪声,所得出的适应度不能真实反映个体的进化质量。适应度评价存在噪声的根本原因是基因型结构个体与显型网络个体间是一对多的映射关系。文献[15]对上述问题给出了详尽的讨论,指出无权重信息的结构进化设计中个体适应度的准确评价是困难的。克服这一问题的有效方法是对网络结构和网络权重同时进化,即将结构信息和权重信息同时编码,使基因型网络个体与显型网络个体间呈一一对应关系,以消除适应度评价噪声,从而提高个体进化效率和进化质量。

4 训练算法的进化设计

大量研究已揭示出网络结构与训练算法间存在内在联系。因此,网络要得到更优性能,须对其训练算法进行进化调整。训练算法的进化设计主要包括训练算法参数进化和学习规则进化两方面。

4.1 训练算法参数的进化

训练算法进化设计的初期研究主要集中在对算法参数(如学习率和惯性系数等)的进化选取。其基本方法是将算法参数与网络结构一起编码进行进化^[25]。通过这种进化可使网络训练算法与网络结构更好地匹配,从而改善网络性能。对于BP算法的学习参数也提出了一些非进化方式的参数调整方法,但这两者间的优缺点比较还有待于进一步研究。

4.2 学习规则的进化

学习参数的进化可在一定程度上提高网络的学习效率,但并未触及训练算法的核心部分,即学习规则。

网络训练算法的学习规则一般隐含了两个基本假设:一是权重的调整只依赖于节点的局部信息,如输入节点的激励、输出节点的激励和当前联结权重

等;二是训练中网络所有权重的调整均采用相同的学习规则。学习规则的一般形式可描述为^[26]

$$\Delta w(t) = \sum_{k=1}^n \sum_{i_1, i_2, \dots, i_k=1}^n \left(\theta_{i_1 i_2 \dots i_k} x_{i_j}(t-1) \right) \quad (3)$$

其中, Δw 为权重变化量, t 为迭代步数, x_1, x_2, \dots, x_n 为节点局部变量, θ_s 为待定实值系数, 它的不同取值将产生不同的学习规则。因此, 学习规则的进化即为 θ_s 实值参数向量的进化。

文献[27]根据式(3)将学习规则定义为 4 个节点局部变量及其 2 阶生成项的线性组合, 忽略了 3 阶项和 4 阶项, 并将 10 个 θ 系数和 1 个尺度系数编码为二进制串, 个体适应度计算采用单隐层网络。进化从一随机产生的学习规则群体开始, 经过 1 000 代的进化, 得到了著名的 delta 规则和 delta 规则的派生规则。文献[28]采用[27]的进化策略对感知器的学习规则进行进化, 但用了 7 个高阶生成项, 即 1 个 1 阶项 3 个 2 阶项和 3 个 3 阶项, 也取得了满意的结果。文献[29]进一步将网络结构、联结权重和学习规则同时进化。为了控制染色体的复杂性, 网络节点设定为二进制门限节点, 但网络节点数固定不变, 学习规则中也只考虑了 2 个 Boolean 变量。

5 神经网络与进化计算结合的其它方式

进化计算为神经网络本身的设计和实现提供了一条有效的途径, 使神经网络的设计和实现更具客观性和科学性, 也使网络具有更优的性能。许多学者对神经网络与进化计算结合的其它方式进行了探讨, 这项工作主要集中在以下几方面^[30, 31]:

- 1) 进化计算与基于梯度下降算法相结合的“混合”算法, 即由进化计算进化生成落在全局最优区域的初始权重, 由 BP 算法进行局部搜索;
- 2) 应用进化计算对网络的输入特征参量进行进化选取, 以获得最优的输入特征向量, 达到控制网络规模, 提高网络性能的目的。
- 3) 基于进化计算的训练样本优化选取, 即网络根据训练情况选择适合的训练样本;
- 4) 将神经网络与进化计算结合起来对控制系统进行进化, 即用神经网络对控制系统进行模拟建模, 用进化计算进化控制系统的参数;
- 5) 基于进化计算从神经网络中获取知识和提取规则。

基于进化计算的神经网络设计和实现已在众多

领域得到应用, 如模式识别^[32]、机器人控制^[33]、财政预测^[34]等, 并取得了较传统神经网络更好的性能和结果。但从目前应用看, 还主要限于小规模问题。随着网络规模的扩大和复杂度的提高, 基于进化计算的神经网络将显示出更强的优越性。

6 结 语

神经网络和进化计算都是借鉴生物个体或生物界的某些行为特征和结构属性而发展起来的人工智能理论方法, 神经网络侧重于对生物个体学习智能的描述, 进化计算则是对生物界进化特征的模拟。因此, 神经网络与进化计算相结合, 必将表现出更加完备的智能特性。即这种结合不仅使神经网络同时具备了学习和进化性能, 表现出更强的智能, 而且可以解决神经网络设计和实现中存在的一些问题, 使神经网络具有更优的性能。

目前, 基于进化计算的神经网络设计和实现已成为神经网络领域一个重要的研究方向, 国内外学者在这方面已做了大量工作。但从总体上看, 这方面研究还处于初期阶段, 理论方法有待于完善规范, 应用研究有待于加强提高。如目前的研究多基于具体事例, 还没形成一般性的方法体系; 计算量较大, 进化进程需进一步提高, 尤其是并行进化计算方法应受到重视; 进化计算理论本身(如编码表示方法、遗传算子、种群收敛性和多样性等)有待于进一步完善和发展; 神经网络与进化计算相结合的其它方式也有待于进一步研究和挖掘。

参考文献:

- [1] Hertz J. Introduction to the theory of neural computation[M]. MA: Addison-Wesley Press, 1991.
- [2] Holland J H. A adaptation in natural and artificial system [M]. M I Univ Michigan Press, 1975.
- [3] Fogel L J. Artificial intelligence through simulated evolution[M]. New York: Wiley, 1996.
- [4] Schwefel H P. Numerical optimization of computer models[M]. Chichester: Wiley, 1981.
- [5] Fogel D B. An introduction to simulated evolutionary optimization [J]. IEEE Trans on Neural Networks, 1994, 5(2): 3-14.
- [6] Whitley D, Starkweather T, Bogart C. Genetic algorithms and neural networks: Optimizing connections and connectivity[J]. Parallel Comput, 1990, 4(3): 347-361.
- [7] Janson D J, Frenzel J F. Application of genetic algo-

- rithms to the training of higher order neural networks [J] J Syst Eng, 1992, 2(1): 272-276
- [8] Greenwood G W. Training partially recurrent neural networks using evolutionary strategies[J] IEEE Trans on Speech Audio Proc, 1997, 5(1): 192-194
- [9] Fogel D B, Wasson E C, Boughton E M. Evolving neural networks for detecting breast cancer [J] Cancer Lett, 1995, 96(1): 49-53
- [10] Yao X, Liu Y. Fast evolutionary programming[A] Proc 5th Annu Conf Evol Prog[C]. Cambridge: MIT Press, 1996 451-460
- [11] Back T, Schwefel H P. An overview of evolutionary algorithms for parameter optimization [J] Evol Comput, 1993, 1(1): 1-23
- [12] Fream M. The upstart algorithm: A method for constructing and training feedforward neural networks [J] Neural Comput, 1990, 2(2): 198-209
- [13] Sietsma J, Dow R J F. Creating artificial neural networks that generalize [J] Neural Networks, 1991, 4(1): 67-79
- [14] Roy A, Kim L S, Mukhopaduyay S. A polynomial time algorithm for the construction and training of a class of multilayer perceptrons[J] Neural Networks, 1993, 6(4): 535-545
- [15] Angeline P J. An evolutionary algorithm that constructs recurrent neural networks[J] IEEE Trans on Neural Networks, 1994, 5(2): 54-65
- [16] Miller G F. Designing neural networks using genetic algorithms[A] Proc 3rd Int Conf Genetic Algorithm & Their Appl[C]. CA: Morgan Kaufmann, 1989 379-384
- [17] Olliker S, Furst M, Maimon O. A distributed genetic algorithm for neural network design and training[J] Complex Syst, 1992, 6(5): 459-477
- [18] Bichsel M, Seitz P. Minimum class entropy: A maximum information approach to layered networks [J] Neural Networks, 1989, 2(2): 133-141
- [19] Bornholdt S, Graudenz D. General asymmetric neural networks and structure design by genetic algorithms Neural Networks, 1992, 5(2): 327-334
- [20] Kitano H. Designing neural networks using genetic algorithms with graph generation system [J] Complex Syst, 1990, 4(4): 461-476
- [21] Vonk E, Jain L C. Using genetic algorithm grammar encoding to generate neural networks[A] Proc IEEE Int Conf Neural Networks[C] 1995, 4(6): 1928-1931
- [22] Mani G. Learning by gradient descent in function space[A] Proc IEEE Int Conf Syst, Man & Cyb[C] 1990 242-247.
- [23] Stork D G. Preadaptation in neural circuits[A] Proc Int Joint Conf Neural Networks[C] 1990 202-205
- [24] White D, Ligomenide P. GANNet: A genetic algorithm for optimizing topology and weights in neural network design [A] Proc Int Workshop Artificial Neural Networks Lecture Note in Computer Science [C] Berlin: Springer-Verlag Press, 1993, 686: 322-327.
- [25] Harp S A, Samad T, Guha A. Toward the genetic synthesis of neural networks[A] Proc 3rd Int Conf Genetic Algorithms & Their Appl[C] 1989 360-369
- [26] Liu Y, Yao X. Evolutionary design of artificial neural networks with different nodes[A] Proc IEEE Int Conf Evol Comput[C] 1996 670-675
- [27] Chalmers D J, Cruana R A, Esheleman L J. The evolution of learning: An experiment in genetic connectionism [A] Proc Connectionist Models Summer School [C] CA: Morgan Kaufmann, 1990 81-90
- [28] Fontanari J F, Meir R. Evolving a learning algorithm for the binary perceptron[J] Neural Networks, 1991, 2(4): 353-359
- [29] Baxter J. The evolution of learning algorithms for artificial neural networks[A] Complex Syst[C] 1992 313-326
- [30] Zhang B T, Veenker G. Neural networks that teach themselves through genetic discovery of novel examples[A] Proc IEEE Int Joint Conf Neural Networks [C] 1991 690-695
- [31] Chao S, Cha K. Evolution of neural network training set through addition of virtual samples[A] Proc IEEE Int Conf Evol Comput[C] 1996 685-688
- [32] 方建安, 邵世煌. 采用遗传算法学习的神经网络控制器 [J] 控制与决策, 1993, 8(3): 208-221.
- [33] Maniezzo V. Genetic evolution of the topology and weight distribution of neural networks [J] IEEE Trans on Neural Networks, 1994, 5(1): 39-53
- [34] Harrauld P G, Kamstra M. Evolving artificial neural networks to combine financial forecast [J] IEEE Trans on Evol Comput, 1997, 1(1): 40-52