

基于内分泌调节机制的行为自组织算法¹⁾

黄国锐 曹先彬 徐敏 王煦法

(中国科学技术大学计算机科学技术系 合肥 230026)

(E-mail: huangdav@mail.ustc.edu.cn; huangdava@sina.com)

摘要 借鉴内分泌系统的高层调节机制,提出了一种新的自主行为自组织算法.此算法用神经系统接受环境信息,通过一种情感学习模型来产生情感因子(即生物激素),再由情感因子来调节神经系统的记忆和行为决策,最后神经系统的记忆与行为模式由遗传系统得以继承.其优点是避免了神经系统复杂的自学习过程,同时保证系统的行为决策具有较高的自组织、自适应能力.为了验证算法的有效性,倒立摆控制的仿真实验表明该算法具有很强的自适应求解能力.

关键词 内分泌系统,行为自组织,情感学习,神经网络,遗传算法,倒立摆

中图分类号 TP13

Self-Organization Algorithm of Behaviors Based on Endocrine Regulation Mechanism

HUANG Guo-Rui CAO Xian-Bin XU Min WANG Xu-Fa

(Department of Computer Science and Technology, University of Science and Technology of China, Hefei 230026)

(E-mail: huangdav@mail.ustc.edu.cn; huangdava@sina.com)

Abstract Motivated by the high-level regulation mechanism of endocrine system, an agent self-organization algorithm of behavior is put forward. This algorithm receives environmental information using a neural system, produces emotional factors (hormones) through an emotion learning model, then uses these factors to regulate the neural system. At last, the memorization and behavior mode of the neural system are exported to the genetic environment. So it avoids the self-study of neural system and gets better self-organization and self-adaptation power. In order to prove the validity of this algorithm, it is used to control an inverted pendulum. The result shows strong self-adaptation performance of the algorithm.

Key words Endocrine system, self-organization of behaviors, emotion learning, neural network, genetic algorithm, inverted pendulum

1) 国家自然科学基金项目(69971022)资助

Supported by National Natural Science Foundation of P. R. China(69971022)

收稿日期 2003-03-04 收修改稿日期 2003-08-20

Received March 4, 2003; in revised form August 20, 2003

1 引言

复杂环境下 agent 行为的自组织、自适应是当前智能控制领域的一个研究热点. 生物信息处理机制有很好的借鉴意义, 如基于人工神经网络和遗传算法的研究工作现在已大量涌现. 近几年来, 相关研究者意识到内分泌调节机制作为生物信息处理的一个重要组成部分, 在开发新的智能学习方法时是不应忽视的^[1~6]. 实际上, 生物的行为除了受神经和遗传系统的影响外, 还受内分泌系统的调节. 我们借鉴内分泌系统的调节机制, 提出了一种新的行为自组织算法. 为了验证该算法的有效性, 针对倒立摆控制问题进行了详细实验, 实验结果表明该算法具有很好的性能.

2 生物内分泌系统简介

生物内分泌系统由内分泌细胞、内分泌细胞所释放的激素和内分泌腺体所组成. 适宜的刺激可刺激生物体内分泌细胞产生适当种类和数量的激素, 这些激素和神经系统的共同作用, 维持着机体内环境的相对稳定, 进而影响生物体的行为. 图 1 说明了内分泌系统和神经系统对生物机体的共同调节作用. 图 1 中各种内分泌腺体形成一个闭合回路, 它们之间通过复杂的相互作用维持着机体内环境的相对稳定. 在闭合回路的基础上, 中枢神经系统可接受外环境中的各种信息(声、光、温度、味等), 通过下丘脑把内分泌系统与外环境联系起来, 形成开口回路. 本文正是借鉴神经系统和内分泌系统的相互作用机制, 提出了一种新的自主体行为自组织算法.

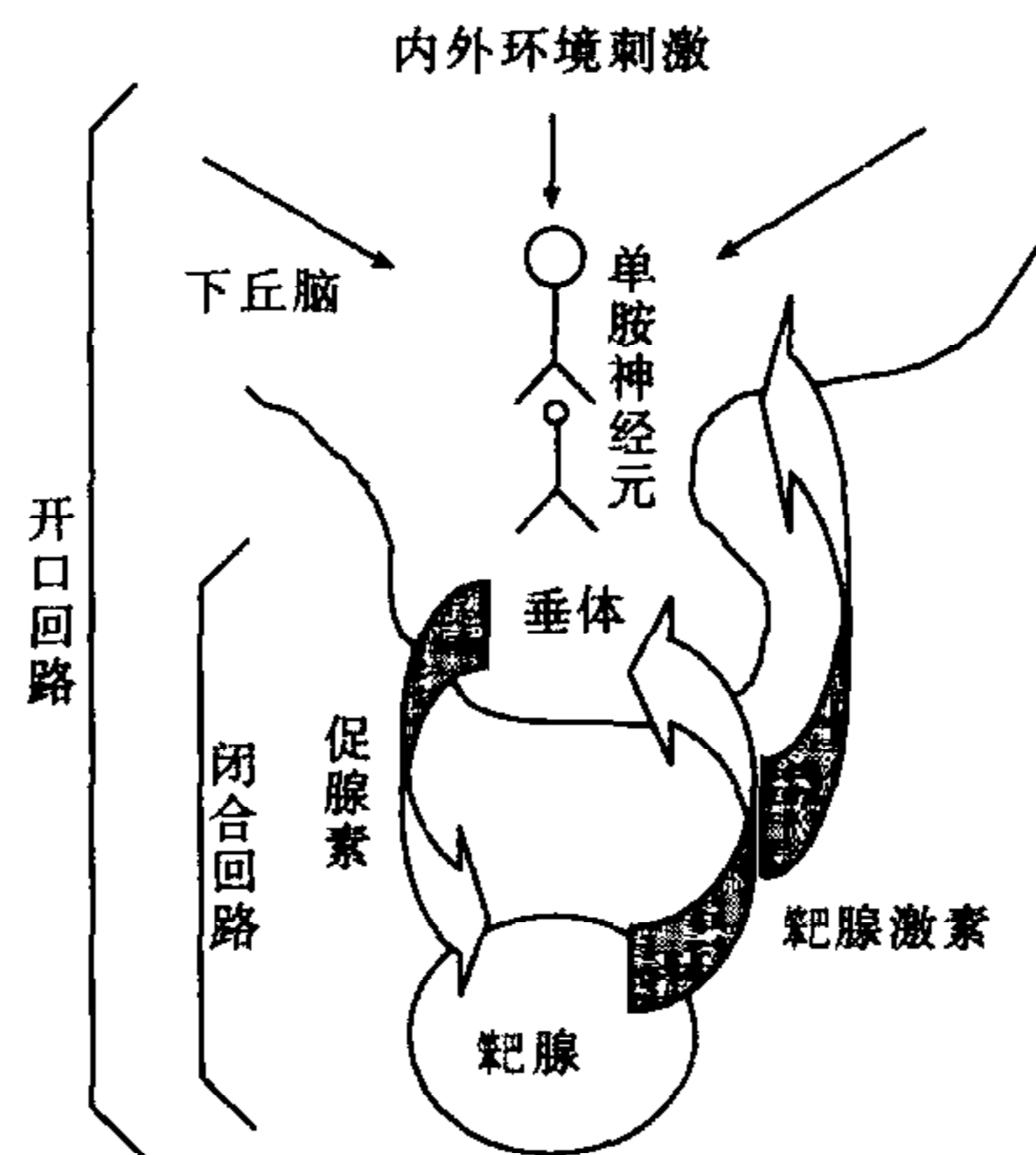


图 1 内分泌系统和神经系统的共同调节作用

Fig. 1 Collective regulation of endocrine system and neural system

3 基于内分泌调节机制的行为自组织算法

3.1 行为自组织算法的结构

自主体行为自组织算法的工作结构如图 2 所示. 自主体从遗传环境中通过染色体的输

人得到神经网络的初始记忆. 神经网络作为系统的一个主要组成部分, 它从行为环境中接收信息, 并控制效应器的动作和内分泌系统的状态变化. 内分泌系统在接收到相应的神经信号后, 分泌相应的激素, 这些激素通过激素回路而返回给神经网络, 对神经网络起着重要的调节作用, 进而影响自主体的行为决策. 神经网络中保存有自主体对行为环境的适应结果, 这些适应结果通过染色体的输出而传给下一代个体.

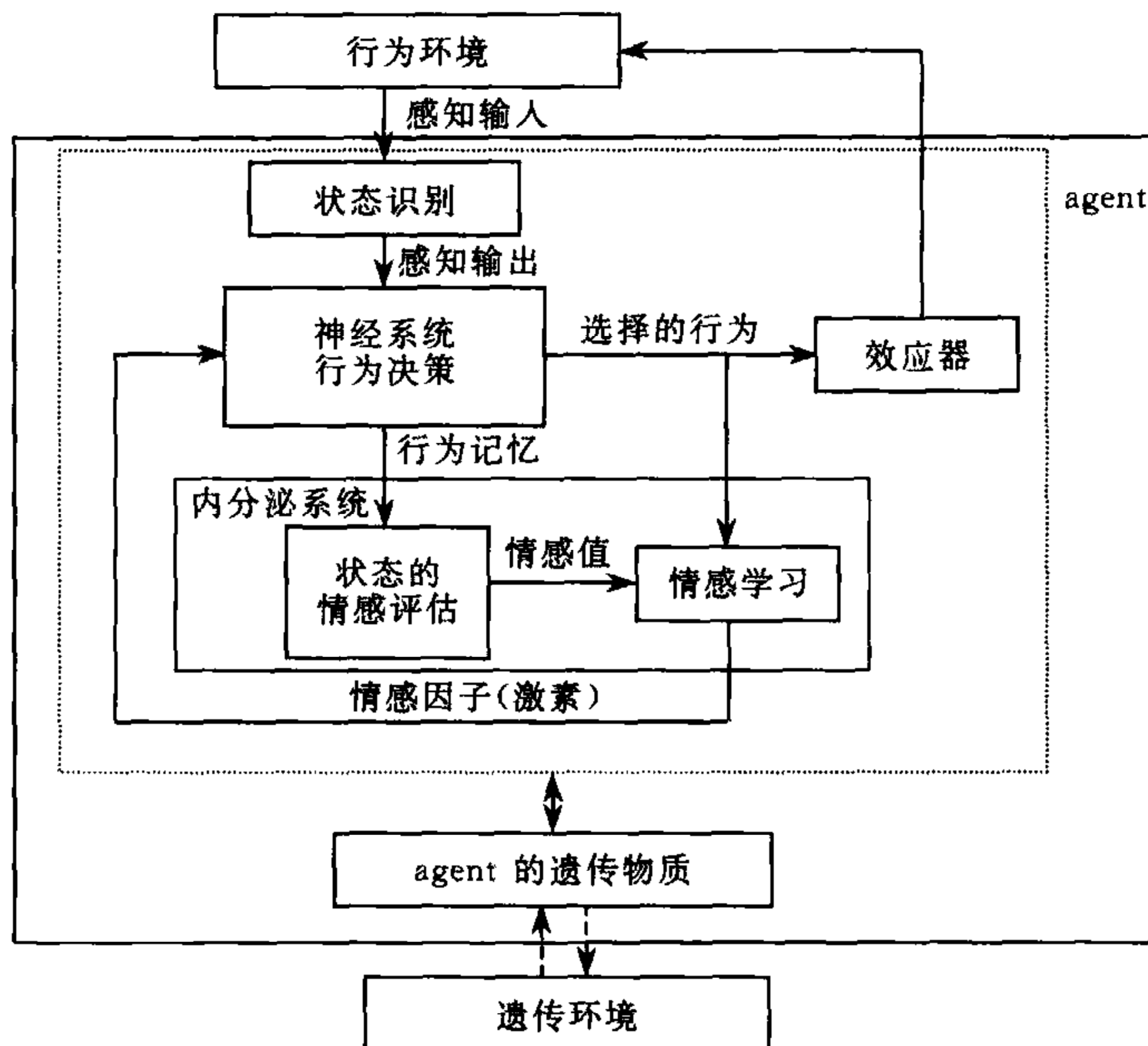


图 2 行为自组织结构
Fig. 2 Architecture for behavior self-organization

3.2 行为自组织算法的实现

本文提出的行为自组织算法实际上由三个相互作用的子部分所组成: 神经网络的记忆与决策、遗传的继承与变异、内分泌的高层调节. 对前两部分的论述本文从略, 下面将重点论述第三部分.

假设系统的状态是离散的, 神经网络用〈状态、动作、情感值〉这样三维序列的集合来表述. 集合的起始值是从遗传环境中通过染色体的输入而得到的. 为了算法描述的方便, 引入三个记号: E, E_{ai} 和 E_i . E 表示的是三维序列〈状态、动作、情感值〉的集合, 可以用一个矩阵来表示, 矩阵的行号表示动作的序号, 矩阵的列号表示状态的序号, 矩阵的元素值表示相应状态动作对的情感值; E_{ai} 表示的是在状态 i 下执行动作 a 的情感值; E_i 表示状态 i 的情感因子. 内分泌的高层调节是通过对系统状态的评估及根据评估结果的反馈学习来实现的, 下面是其算法步骤:

1) 在状态 j 下选择一个动作, 设所选动作为 a , 在状态 j 下执行动作 a 后返回状态 i . 动作选择规则用公式(1)进行描述, 其中 * 号表示状态 j 下所有可能的动作.

$$a = Afunc\{E_{*,j}\} = \max\{E_{*,j}\} \tag{1}$$

2) 对状态 i 进行情感评价, 得到其情感因子 E_i . 评价规则用公式(2)描述.

$$E_i = Efunc\{E_{*,i}\} = \max\{E_{*,i}\} \tag{2}$$

3) 情感学习. 情感学习规则用公式(3)描述, 其中 β 表示学习率.

$$E_{a_j}(t+1) = Ufunc(E_i, E_{a_j}(t)) = (1 - \beta)E_{a_j}(t) + \beta E_i \quad (3)$$

4) 令 $j=i$, 返回 1).

上面步骤中的 3 个函数关系可以针对实际问题自行设计, 只要能保证学习过程的收敛就可以. 另外在实际应用中, 我们都要定义一个情感学习的结束条件. 情感学习结束后, 神经网络中保存有自主体对行为环境的自适应结果, 这些自适应结果通过染色体的输出而传给下一代个体. 因而下一代与上一代相比, 下一代有更加适应行为环境的趋势.

4 仿真实验及结果分析

本实验中控制系统将决定各时刻加在小车上的力的作用方向(左还是右), 力大小不变, 为 10N, 所以本实验中可选的动作只有两个, 分别是左推和右推. 定义左推对应的动作序号为 0, 右推对应的序号为 1. 关于倒立摆系统的详细介绍请参考文献[7].

4.1 实验设计

4.1.1 编码

这里采用先对状态空间划分, 然后对各个子空间用 4 位二进制串进行表示, 二进制串值表示该状态对应的一个情感因子, 情感因子要归一化为 -1 到 $+1$ 之间的一个数. 详细编码情况如下: 角度 θ 的论域为 $[-12, +12]$, 对 θ 作 6 个划分: $-12 \sim -8$, $-8 \sim -4$, $-4 \sim 0$, $0 \sim +4$, $+4 \sim +8$, $+8 \sim +12$ ($^\circ$). 小车的位置 x 的论域为 $[-2.4, +2.4]$, 对 x 作 3 个划分: $-2.4 \sim -0.8$, $-0.8 \sim +0.8$, $+0.8 \sim 2.4$ (m). 角速度 θ' 的单位为 rad/s, 论域为 $[-\infty, +\infty]$, 对 θ' 作 6 个划分: $-\infty \sim -n/3$, $-n/3 \sim -n/6$, $-n/6 \sim 0$, $0 \sim +n/6$, $+n/6 \sim +n/3$, $+n/3 \sim +\infty$ (rad/s). 对速度 x' 作 3 个划分: $-\infty \sim -0.5$, $-0.5 \sim +0.5$, $+0.5 \sim +\infty$ (m/s). 这样把摆立摆系统的状态空间离散化为 $6 \times 6 \times 3 \times 3 = 324$ 个子空间, 各个子空间对应的序号为 0 到 323 之间的一个数.

4.1.2 算法步骤

1) 随机生成初始群体.

2) 如果进化代数超过 50, 则结束运算; 否则用群体中每个个体所包含的信息去控制倒立倒的运动. 在控制过程中, 首先将摆的初始角速度、小车的速度、摆的角度和小车的位置赋 0, 然后采用前面所描述的内分泌调节算法去控制倒立倒的运动. 在本实验中情感学习的结束条件是摆倒下或者摆能保持 1000000 步不倒. 如果后一种情况出现, 则在记录下此次的运行步数后, 终止运行, 同时输出 E 矩阵. 对每个个体连续进行 100 次试探, 在这 100 次试探中取最大的运行步数, 作为该个体的适应度, 同时记录对应该适应度的 E 矩阵. E 矩阵经归一化处理后转化为一个新的二进制串, 并用新的二进制串去替换原来的个体, 这样新一代群体应有更好的适应相同行为环境的能力. 在此实验中学习率 β 取值为 0.2.

3) 遗传操作. 选择、交叉与变异, 转向 2).

4.2 实验结果与分析

基于前面的设计, 本文做了 10 次独立实验, 群体规模为 50, 实验结果如表 1 所示.

表 1 仿真实验结果
Table 1 The simulation result

| 实验序号 | 进化代数 | 最大运行步数 | 实验序号 | 进化代数 | 最大运行步数 |
|------|------|---------|------|------|---------|
| 1 | 10 | 1577598 | 2 | 13 | 1568316 |
| 3 | 10 | 1583238 | 4 | 10 | 1580871 |
| 5 | 9 | 5006282 | 6 | 20 | 1035626 |
| 7 | 10 | 1577222 | 8 | 12 | 1579750 |
| 9 | 15 | 1499557 | 10 | 11 | 1400562 |

图 3 是以第 8 次实验结果绘制的求解曲线图,平均进化代数为 12. 表 2 是本文所采用的行为自组织算法和其他学习算法的一个对比. 表 2 中 AHC 的结果是由文献[7]提供的, FACL 的结果是由文献[8]提供的. 对比表明本文所采用的行为自组织算法的学习速度明显比其他的学习算法快. 对算法稳定性的理论分析,我们将进一步进行探讨.

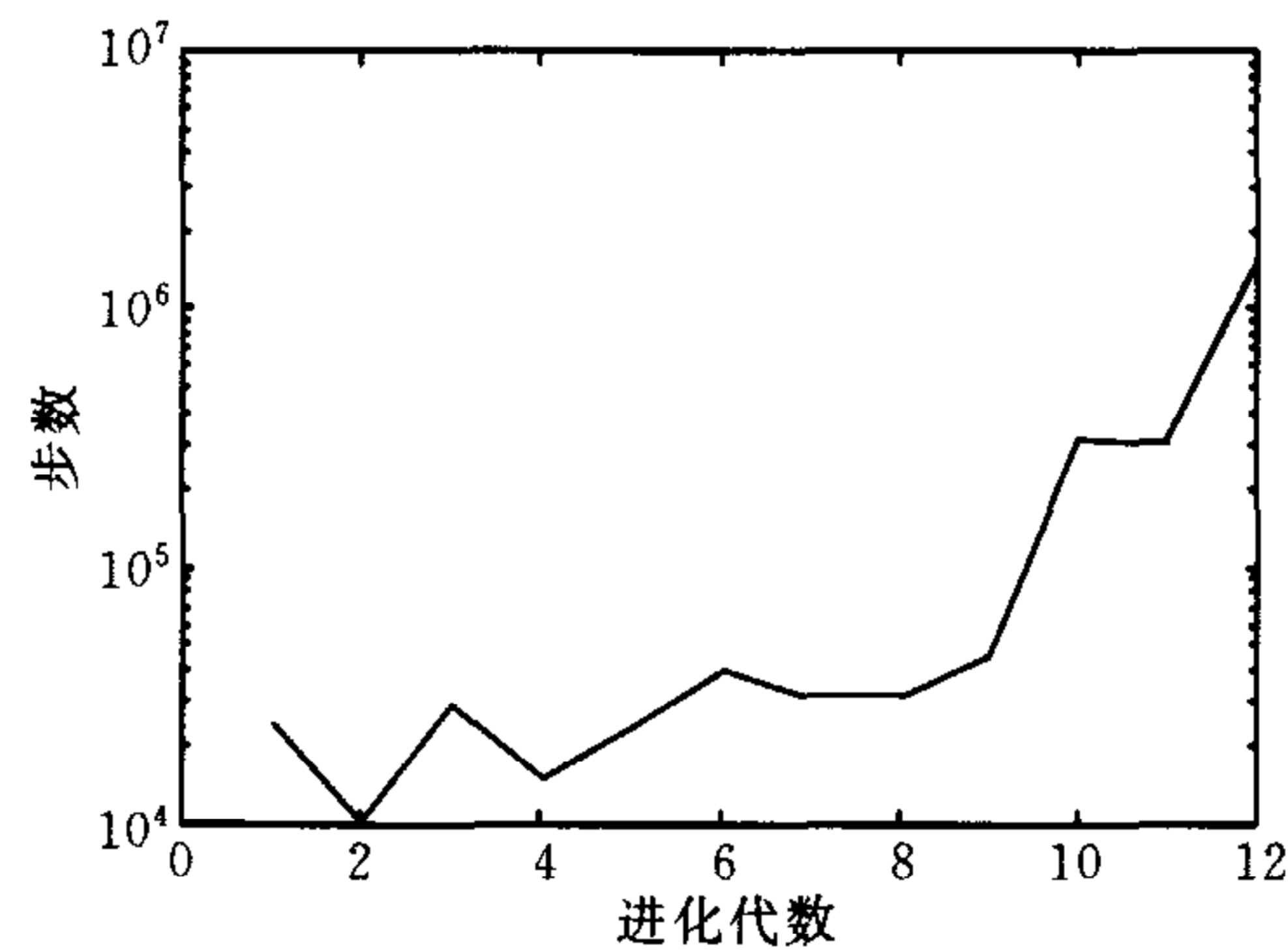


图 3 算法求解曲线图
Fig. 3 The graph of algorithm's olution

表 2 几种学习算法对比
Table 2 Comparison of several algorithm

| 学习算法 | 设置的最大运行步数 | 进化代数(代) |
|---------|-----------|---------|
| AHC[7] | 500000 | 75 |
| FACL[8] | 500000 | 12.59 |
| 本文算法 | 1000000 | 12 |

5 结束语

借鉴神经系统和内分泌系统的相互作用机制,提出了一种新的行为自组织算法. 在该算法中,内分泌系统通过分泌激素(情感因子)来调节神经系统的作用,进而调整自主体的行为;同时在内分泌系统中,内分泌腺体之间通过相互促进与抑制,来对环境的变化做出反应,以达到逐渐适应行为环境的能力. 最后本文针对倒立摆控制问题进行了详细实验,实验结果表明该算法具有很好的求解性能.

References

- 1 Will P, Castaño A, Shen W M. Robot modularity for self-reconfiguration. In: Proceedings of SPIE Sensor Fusion and Decentralized Control II, Boston, 1999, 236~245
- 2 Shen W M, Lu Y, Will P. Hormone-based control for self-reconfigurable robots. In: Proceedings of International Conference Autonomous Agents, Barcelona, Spain, 2000, 1~8
- 3 Shen W M, Salemi B, Will P. Hormone for self-reconfigurable robots. In: Proceedings of International Conference Intelligent Autonomous Systems, Vinece, Italy; IOS Press, 2000, 918~925
- 4 Salemi B, Shen W. M; Will P. Hormone controlled metamorphic robots. In: Proceedings of International Conference Robotics and Automation, 2001, 4194~4199
- 5 Canamero D. A hormonal model of emotions for behavior control. In: Proceedings of European Conference on Artificial Life, 1997, 28~31
- 6 Ogata T, Sugano S. Emotional communication between humans and the autonomous robot which has the emotion model. *IEEE International Conference Robotics and Automation*, 1999, 4(2):3177~3182
- 7 Chen Guo-Liang, Wang Xu-Fa. Genetic Algorithm and Its Applications. Beijing: People's Post and Telecommunications Publishing House, 1996(in Chinese)
- 8 Jouffe L. Fuzzy inference system learning by reinforcement methods. *IEEE Transactions on Systems, Man and Cybernetics*, 1998, 28(3):338~355

黄国锐 博士研究生. 主要研究领域为计算智能、人工内分泌.

(HUANG Guo-Rui Ph. D. candidate. His research interests include computational intelligence and artificial endocrine system.)

曹先彬 博士, 副教授. 主要研究领域为计算智能、人工免疫.

(CAO Xian-Bin Ph. D., associate professor. His research interests include computational intelligence and artificial immune system.)

徐敏 硕士研究生. 主要研究领域为机器人足球、智能计算.

(XU Min Master student. His research interests include computational intelligence and robot soccer.)

王煦法 教授, 博士生导师. 主要研究领域为计算智能、人工生命、网络安全.

(WANG Xu-Fa Professor. His research interests include computational intelligence, artificial life, and network security.)