

具有操作条件反射机能的人工感觉运动系统

黄 静^{1,2†}, 阮晓钢¹, 于乃功¹, 张晓平¹, 魏若岩¹, 范青武²

(1. 北京工业大学 电子信息与控制工程学院, 北京 100124; 2. 北京工业大学 实验学院, 北京 101101)

摘要: 为探索认知机理、模拟认知机能进而提高机器人的认知及智能水平, 提出了一种具有操作条件反射机能的人工感觉运动系统. 该系统以复现感觉运动系统的方式重现了生物的运动神经认知, 实现了从感知到运动的映射关系, 同时借鉴了斯金纳的操作条件反射理论, 使得该系统具有操作条件反射机能, 遵从“刺激-反应-强化”的逻辑形成了感知与运动之间的闭环系统. 为验证系统的正确有效性, 复现了行为心理学及《控制论》中的两个经典实验. 对比实验结果证明, 本系统成功地模拟了生物感觉运动系统, 使机器人具有类似生物的自学习能力, 能自主的认知环境, 实现对环境的自适应.

关键词: 感觉运动系统; 操作条件反射; 斯金纳鸽子实验; 维纳虫; 认知机器人学; 神经网络

中图分类号: TP273 文献标识码: A

Artificial sensorimotor system with operant conditioning function

HUANG Jing^{1,2†}, RUAN Xiao-gang¹, YU Nai-gong¹,
ZHANG Xiao-ping¹, WEI Ruo-yan¹, FAN Qing-wu²

(1. College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China;
2. Pilot College, Beijing University of Technology, Beijing 101101, China)

Abstract: In order to explore the cognitive mechanism and simulate the cognitive ability, we develop an artificial sensorimotor system with operant conditioning function (ASSOCF) to improve the cognitive and intelligent ability of robots. The system reproduces the biological motor cognition by simulating the sensorimotor system and realizes the mapping from perception to action. Meanwhile, the system is endowed with operant conditioning function based on Skinner's operant conditioning theory. It forms the close-loop system between perception and action in compliance with the logic of 'stimulus-response-reinforcement'. Two classic experiments in behavioral psychology and control theory, Skinner pigeon experiment and Wiener worm experiment, are performed to testify the validity of the system. Comparisons of the experiment results prove that the system successfully simulates the biological sensorimotor system. With such a system, the robots are endowed with the abilities of autonomous learning and self-adaptation to the environment.

Key words: sensorimotor system; operant conditioning; Skinner pigeon experiment; Wiener worm experiment; cognitive robotics; neural networks

1 引言(Introduction)

探索认知机理、理解认知行为是认知科学、人工智能、机器人学等诸多学科共同关心的课题. 围绕这一主题, 各学科相互交叉、彼此融合, 直接导致了“认知机器人学”(cognitive robotics)的诞生. Brooks教授领导的AP团队构造了一个名为Cog的人形机器人^[1], 开展机器人认知行为的研究, 并提出了认知机器人学的概念^[2].

在认知机器人研究领域, 认知计算模型或认知算

法是研究的重要内容, 涉及到认知形式化即认知建模的工作. 从感觉运动神经系统切入的认知建模是研究思路之一. 有关感觉运动神经计算认知模型的代表性研究包括: Natale等人^[3]为人形机器人设计的感觉运动系统, 能通过认知, 由简单的初始化形态形成感觉运动神经协调机制; Kuniyoshi和Sangawa^[4]建立了肌肉骨骼模型, 并进行了相关感觉运动神经认知发育和模拟实验; Kimura等人^[5]为四足机器人设计了仿生的感觉运动神经系统, 具有所谓的中心模式生成器,

收稿日期: 2014-03-25; 录用日期: 2015-02-10.

†通信作者. E-mail: aiandrobot@163.com.

北京高等学校青年英才计划项目(YETP1610), 国家自然科学基金项目(61075110, 61375086), 北京市自然科学基金项目/北京市教育委员会科技计划重点项目(KZ201210005001), 国家重点基础研究发展计划(973计划)项目(2012CB720000), 高等学校博士学科点专项科研基金课题(20101103110007)资助.

Supported by Beijing Universities Young Talent Plan (YETP1610), National Natural Science Foundation of China (61075110, 61375086), Key Project of S&T Plan of Beijing Municipal Commission of Education (KZ201210005001), National Basic Research Program of China (973 Program) (2012CB720000) and Specialized Research Fund for the Doctoral Program of Higher Education (20101103110007).

以实现自适应的动态行走; Chao和Lee^[6]为自主机器人手眼系统设计了感觉运动神经系统模型及其认知算法, 可通过发育形成手眼协调机制; Mathews等人^[7]构建了一个名为PASAR的人工感觉运动系统, 可模拟期盼、注意、记忆、感觉、响应等认知机制, PASAR量化了相关认知机制, 为生理学和心理学实验提供了具体的假设和预测。

这些关于人工感觉运动系统的研究多强调视觉、触觉等人类感知机能与肌肉骨骼系统之间的联系, 较少涉及以思维为中心的高级认知机能, 例如学习、决策、问题解决等。为了弥补这一缺失, 本文从斯金纳(Skinner)操作条件反射理论中得到启发, 设计并实现了一个具有操作条件反射机能的人工感觉运动系统。

斯金纳操作条件反射理论由美国著名心理学家斯金纳(B.F. Skinner)提出^[8]。他继承并发展了巴甫洛夫的经典条件反射, 并对其中的核心概念“强化”进行了革新。他把“强化”分为正强化(positive reinforcement)和负强化(negative reinforcement)两种, 正强化促使有机体对刺激的反应概率增加, 而负强化则促使有机体消除该刺激的反应增加。在生物体内在的取向性影响下, 生物在感知环境刺激的同时根据以上原则做出反应, 从而促使自身状态向理想状态趋近, 进而学习到环境变化的内在规律, 建立操作条件反射。斯金纳的操作条件反射理论反映了生物的感知与运动之间的联系, 感知产生反应, 反应影响运动的概率, 这正是斯金纳理论的核心。

斯金纳操作条件反射理论的合理性在大量的动物学习实验及人类教学实践中得到了证明。自上世纪90年代中期以来, 国内外有不少学者及研究机构致力于对其进行计算建模, 试图由此再现智能体的自学习行为。Zalama等人^[9]借鉴了Grossberg的条件反射模型, 以此来解决机器人的避障问题; 他们的建模侧重于经典条件反射, 未能凸显操作条件反射理论的核心思想。Itoh等人^[10]基于操作条件反射的Hull理论进行建模, 使机器人能从预先定义的行为列表中自动的选择适合特定情境的行为模式。Salotti等人^[11]提出了一种基于事件预测的模型, 该模型由Rescorla和Wagner的经典条件反射模型发展而来, 结合贝叶斯网络对其进行改进, 最终实现了对操作条件反射和经典条件反射的整合。Cyr等人^[12]采用人工尖峰神经网络(ASNN)构建了一个具有操作条件反射功能的人工脑, 该模型的核心部分仅包含数个神经元, 结构上极其简化, 通过数个复杂度逐渐增加的仿真实验验证了模型的有效性。戴丽珍等人^[13]根据自动机原理, 以操作条件反射学习机制为基础, 运用仿生的自组织学习方法, 提出一种自治操作条件反射自动机; 黄静等人^[14-15]提出了一种操作条件反射学习模型, 将其应用于机器人避障问题, 并证明了模型的收敛性。以上工作虽然从

不同角度对操作条件反射进行了计算建模, 但是并未涉及感觉运动系统, 没有探讨操作条件反射在感觉运动神经系统的认知建模中可起到的作用。

本文借鉴了斯金纳的操作条件反射理论, 将其融于感觉运动系统的设计, 使其在反映从感知到运动的映射关系的同时还具有操作条件反射机能, 能够较好的模拟和再现生物自学习行为。为了证明模型的有效性, 设计了仿真实验对其进行验证。两个仿真实验分别为斯金纳鸽子实验和维纳机器蠕虫实验。经典实验的复现结果证明了本模型的合理性。

2 模型构建(Model description)

具有操作条件反射机能的人工感觉运动系统(artificial sensorimotor system with operant conditioning function, ASSOCF)由以下两大模块组成, 如图1所示。

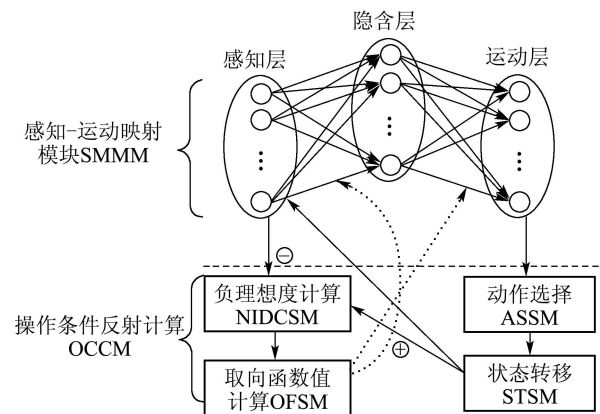


图 1 ASSOCF结构示意图

Fig. 1 Structure of ASSOCF

1) 感知-运动映射模块(sensorimotor mapping module, SMMM). 此模块为一个3层前馈神经网络 $N^3[l, m, n]$, 其中, 输入层与隐含层之间的权值矩阵为 $W1$, 隐含层与输出层之间的权值矩阵为 $W2$. 其输入层包含 l 个神经元, 以编码形式表征所感知到的信息, 构成所谓“感知层”; 隐含层包含 m 个神经元, 对感知层传输的信息进行计算和处理, 隐含层及两个权值矩阵 $W1, W2$, 从功能上模拟了生物感知运动系统中的信息处理中枢; 输出层包含 n 个神经元, 代表动作集合中的 n 个动作, 构成“运动层”. 信息按前馈(feed forward)方式传播, 经感知层、隐含层、运动层前向流动, 实现了从感知到运动的映射关系。

在这里需要说明的是, 生物感知信息包括外部感知(如视觉、听觉、触觉、味觉、嗅觉等)和内部感知(如饥饿感知、干渴感知等)两种, 虽然涉及到的感受器官各异, 但它们所承载的感知信息在神经系统中都是以统一的编码形式传递及处理的。因此, 本模型中的感知层对感知信息的类型不加区分。

2) 操作条件反射计算模块(operant conditioning computational module, OCCM). 如图1所示, 此模块由

负理想度计算、取向函数值计算、动作选择、状态转移等4个子模块组成。

① 负理想度计算子模块(negative ideal degree computational sub module, NIDCSM). 所谓负理想度是对感知信息对应的状态进行评估的一个概念. 当感知信息对应的状态与理想状态之间的差距越大, 负理想度值就会越大, 反之则越小. 负理想度计算子模块分别从感知层及状态转移子模块中接收信息, 针对具体应用问题对感知信息对应的状态进行评价, 计算出这两个不同状态的负理想度值之差. 记感知层输入信息对应的状态为 s_1 , 状态转移子模块输出信息对应的状态为 s_2 , 负理想度函数为 $\varepsilon(\cdot)$, 则本模块的输出为 $\Delta\varepsilon_{12} = \varepsilon(s_2) - \varepsilon(s_1)$.

② 取向函数值计算子模块(orientation function sub module, OFSM). 取向函数 δ 模拟了自然界中生物的取向性, 反映了生物对状态变化的倾向程度. 记 δ_{ij} 表示状态 s_i 执行动作 a_k 转移至状态 s_j 后的取向函数值, 和生物取向性概念一致, $\delta > 0$ 时, 为正取向, 说明状态变化趋向于理想状态, 系统性能趋向变好; $\delta < 0$ 时, 为负取向, 说明系统性能趋向变差; $\delta = 0$ 时, 为零取向, 说明系统性能没有变化. 该模块从NIDCSM中接收负理想度差值作为输入, 按式(1)规定的计算方法计算取向函数值, 然后输出到SMMM中参与权值 $W1$ 及 $W2$ 的调整.

$$\delta_{ij} = \delta(\Delta\varepsilon_{ij}) \begin{cases} > 0, & \Delta\varepsilon_{ij} < 0, \\ = 0, & \Delta\varepsilon_{ij} = 0, \\ < 0, & \Delta\varepsilon_{ij} > 0, \end{cases} \quad (1)$$

式中 $\Delta\varepsilon_{ij} = \varepsilon(s_j) - \varepsilon(s_i)$, 取向函数 δ 在定义区间上连续, 为对 $\Delta\varepsilon_{ij}$ 单调递减函数, 其绝对值随 $\Delta\varepsilon_{ij}$ 绝对值单调递增. 当 $\Delta\varepsilon_{ij} > 0$ 时, 负理想度增大, 系统性能趋向变差, 因而取向函数 $\delta < 0$, 且 $\Delta\varepsilon_{ij}$ 越大, 取向函数 δ 越小; 反之, 当 $\Delta\varepsilon_{ij} < 0$ 时, 负理想度变小, 系统性能趋向变好, 因而取向函数 $\delta > 0$, 且 $\Delta\varepsilon_{ij}$ 越大, 取向函数 δ 越小; 当 $\Delta\varepsilon_{ij} = 0$ 时, 负理想度不变, 系统性能趋向也不变化, 因而取向函数 $\delta = 0$.

③ 行动选择子模块(action selection sub module, ASSM). 该模块接收SMMM的输出, 由此获得动作集合的概率分布, 按照“胜者全拿”(winner-take-all)的原则选择概率最大的动作执行.

④ 状态转移子模块(state transfer sub module, STSM). 该模块承接ASSM之后, 计算执行动作之后引发的状态转移, 记录新状态, 并将其输出到SMMM中的感知层; 同时输出到NIDCSM, 以计算新旧状态的负理想度差值.

3 系统运行流程(Operate process of ASSOCF)

感知层第 i 个神经元的输入为 s_i , 第 i 个神经元的输出为 s_i^* , $i = 1, 2, \dots, l$; 隐含层第 j 个神经元的输入

为 h_j , 第 j 个神经元的输出为 h_j^* ; 运动层第 o 个神经元的输入为 m_o , 第 o 个神经元的输出为 m_o^* ; w_{ij} 表示感知层第 i 个神经元与隐层第 j 个神经元之间的连接权值, w_{jo} 表示隐层第 j 个神经元与运动层第 o 个神经元之间的连接权值, 则系统将遵循如下步骤运行:

步骤 1 初始化. 确定感知-运动映射模块SMMM结构(即感知层、隐层及运动层等各层神经元数量), 令权值矩阵 $W1, W2$ 在 $[0, 1]$ 上随机取值, 确定初始感知状态, 设定学习速率;

步骤 2 将当前状态的感知信息输入感知层;

步骤 3 按照前馈神经网络的工作算法计算从感知到运动的映射, 即:

步骤 3.1 依据式(2)计算感知层输出

$$s_i^* = s_i. \quad (2)$$

步骤 3.2 依据式(3)-(4)计算隐层输出

$$h_j = \sum_i w_{ij} s_i^*, \quad (3)$$

$$h_j^* = \frac{1}{1 + e^{-h_j}}. \quad (4)$$

步骤 3.3 根据式(5)-(6)计算运动层输出

$$m_o = \sum_j w_{jo} h_j^*, \quad (5)$$

$$m_o^* = m_o; \quad (6)$$

步骤 4 根据运动层输出, 按照式(7)计算得到动作集合概率分布

$$p_o = \frac{|m_o^*|}{\sum_o |m_o^*|}, \quad (7)$$

式中: p_o 表示第 o 个动作对应的概率, $|m_o^*|$ 表示求 m_o^* 的绝对值.

步骤 5 按照“胜者全拿”(winner-take-all)的原则选择概率最大的动作执行;

步骤 6 计算动作执行后的状态转移, 记录新状态;

步骤 7 计算状态转移前后的负理想度值之差. 设动作执行前状态为 s_{old} , 执行后状态为 s_{new} , 负理想度函数为 ε , 则状态转移前后的负理想度值之差为

$$\Delta\varepsilon = \varepsilon(s_{new}) - \varepsilon(s_{old}); \quad (8)$$

步骤 8 按照式(1)的说明计算取向函数值, 记为 δ ;

步骤 9 参照负梯度下降误差反传算法, 依据式(9)-(12)调整权值:

步骤 9.1 计算动作层与隐层之间的权值改变:

$$\Delta w_{jo} = \delta \eta h_j^*, \quad (9)$$

$$w_{jo}(k+1) = w_{jo}(k) + \Delta w_{jo}. \quad (10)$$

步骤 9.2 计算隐层与感知层之间的权值改变:

$$\Delta w_{ij} = \delta \eta w_{jo}(k) h_j^*(1 - h_j^*) s_i, \quad (11)$$

$$w_{ij}(k + 1) = w_{ij}(k) + \Delta w_{ij}, \quad (12)$$

其中 η 为学习速率. 显然, 当 $\delta > 0$ 时, 表示系统性能趋于变好, $\Delta w_{jo} = \delta \eta h_j^* > 0$, $\Delta w_{ij} = \delta \eta w_{jo}(k) h_j^*(1 - h_j^*) s_i > 0$, 则 w_{jo}, w_{ij} 增大, 对应动作输出增大, 被选择的概率也相应增加; 当 $\delta < 0$ 时, Δw_{jo} 及 Δw_{ij} 均为负, 则 w_{jo}, w_{ij} 减小, 对应动作输出减小, 被选择的概率也相应减小; 当 $\delta = 0$ 时, Δw_{jo} 及 Δw_{ij} 均为零, 则 w_{jo}, w_{ij} 为零, 对应动作输出不变, 被选择的概率也保持不变. 由此实现了操作条件反射功能.

步骤 10 令新状态为当前状态;

步骤 11 判断是否满足结束条件, 如果满足, 则结束程序, 否则, 转第2步.

系统运行流程的简化描述如图2所示.

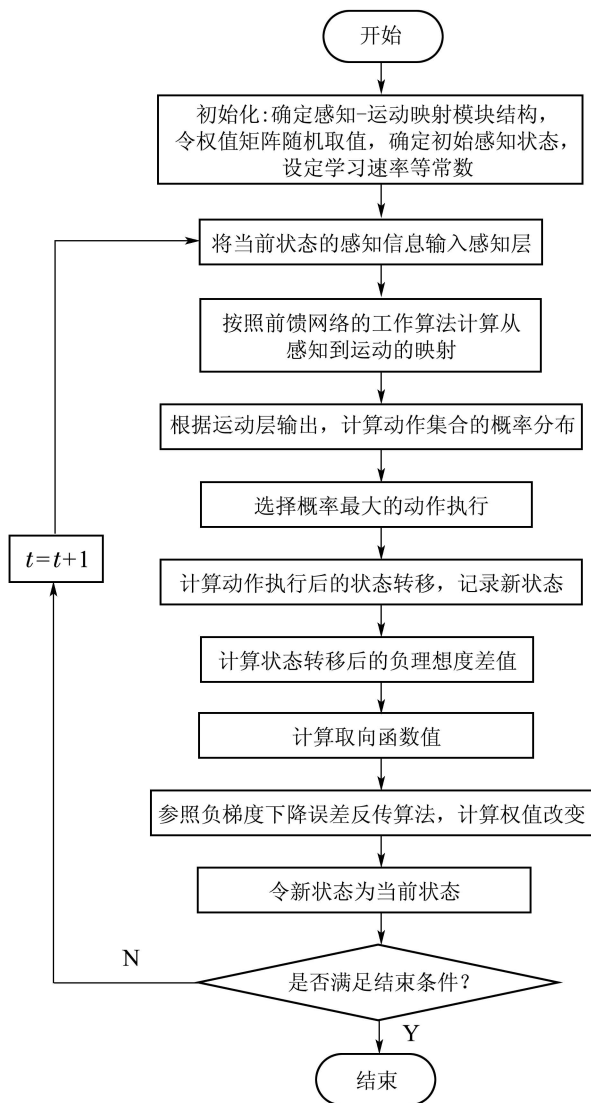


图 2 系统运行流程图

Fig. 2 Flow diagram of ASSOCF

4 仿真实验(Simulation experiment)

为了验证本系统模型的正确有效性, 本节将通过两个仿真实验来达到这一目的.

自然界中生物的感知-运动机制常常是在和生存环境交互的过程中自治地 and 自组织的建立起来的. 出于仿生的考虑, 本文分别将本模型作为人工脑应用在机器鸽及机器蠕虫上, 复现了行为心理学及控制论中两个经典的实验. 同时, 为进一步说明本模型性能, 设计了对比实验, 将其与黄静等人^[14-15]提出的一种操作条件反射学习模型(operant conditioning learning model, OCLM)相比较.

4.1 斯金纳鸽子实验(Skinner pigeon experiment)

如前所述, 斯金纳为了证明操作条件反射理论, 进行了大量的动物行为实验, 著名的斯金纳鸽子实验就是其中之一. 在这个实验中, 鸽子被放在一个箱中(即著名的斯金纳箱), 面对红、黄、蓝3色按钮. 啄红色按钮, 鸽子将会得到食物; 啄黄色按钮, 无任何刺激; 啄蓝色按钮, 则会给予电击. 斯金纳发现, 开始时鸽子啄3色按钮的次数基本相等, 但实验进行一段时间后, 鸽子啄红色按钮的次数明显增多, 反之啄蓝色、黄色按钮的次数则明显下降, 从而证明了理论的正确性.

仿真实验中按图2所示流程对该实验进行了复现.

首先, 令感知层神经元个数 $l = 1$, 代表“饥饿”、“半饱”、“饱足”3种感知信息之一, 其中, 感知层输入为1时代表感知状态为“饥饿”, 输入为2时代表“半饱”, 输入为3时代表“饱足”; 运动层神经元个数 $n = 3$, 代表“啄红色按钮”、“啄黄色按钮”及“啄蓝色按钮”3种动作; 根据多次实验结果, 将隐层神经元个数 m 确定为100, 令权值矩阵 $W1, W2$ 在 $[0, 0.05]$ 随机取值. 令初始状态为“饥饿”. 然后, 将当前状态的感知信息输入感知层, 依前馈神经网络的工作算法如式(2)-(6)计算出运动层的输出. 随后, 根据运动层输出, 按照式(7)计算得到动作集合分布, 从中选择概率最大的动作执行. 之后, 根据表1所示的状态转移矩阵计算该动作执行后的状态转移, 并将其记录为新状态.

表 1 机器鸽状态转移矩阵

Table 1 State transition matrix for machine pigeon

	啄红色按钮	啄黄色按钮	啄蓝色按钮
饥饿	半饱	饥饿	饥饿
半饱	饱足	饥饿	饥饿
饱足	饱足	半饱	饥饿

各状态的负理想度值依表2进行计算, 并可由式(8)得到状态转移前后的负理想度值之差.

表2 各状态负理想度值

Table 2 Negative ideal degree value for each state

状态	饱足	半饱	饥饿
负理想度值	-10	0	10

根据式(1)要求,取式(13)为取向函数值的计算方法

$$\delta_{ij} = \delta(\Delta\varepsilon_{ij}) = \begin{cases} \exp(1/\Delta\varepsilon_{ij}), & \Delta\varepsilon_{ij} < 0, \\ 0, & \Delta\varepsilon_{ij} = 0, \\ -\exp(-1/\Delta\varepsilon_{ij}), & \Delta\varepsilon_{ij} > 0, \end{cases} \quad (13)$$

其中: $\Delta\varepsilon_{ij} = \varepsilon_j - \varepsilon_i$ 表示机器人从状态 s_i 转移到状态 s_j 之后负理想度值的变化, δ_{ij} 表示状态转移的取向函数值。

之后,参照梯度下降误差反传算法,依据式(9)–(12)调整权值,设定学习速率 $\eta = 0.1$,学习100次之后程序终止。

实验结果如图3所示。当学习开始时,机器鸽随机选择啄取按钮,因此3种动作的概率相等;随着学习的进行,机器鸽啄红色按钮的概率逐渐上升,啄黄色、蓝色按钮的概率逐渐下降,且啄蓝色按钮概率下降更快;经过100次学习后,机器鸽啄红色按钮的概率已经趋近于1,而啄黄色、蓝色按钮的概率则趋近于0,这说明机器鸽已经学得最优动作,操作条件反射已经建立。

另一方面,随着机器鸽逐渐学会最优动作,其感知到的状态也逐渐趋于理想,如图4所示。开始时,机器鸽处于饥饿状态,随着学习的进行,机器鸽逐渐学会啄红色按钮,从而状态渐渐向饱足转变,当学习到近40次时,机器鸽已稳定在饱足状态。机器鸽的学习展现了生物的“感知–行动”机制的形成过程:机器鸽通过内部感知获得自身当前状态(饥饿、半饱或饱足),当动作执行改变感知状态时,机器鸽的动作概率也因此改变;由此,行动可以改变感知,而感知也能反馈影响行动,实现了感知与行动的闭环系统,最终体现了生物自学习、自适应的智能特征。

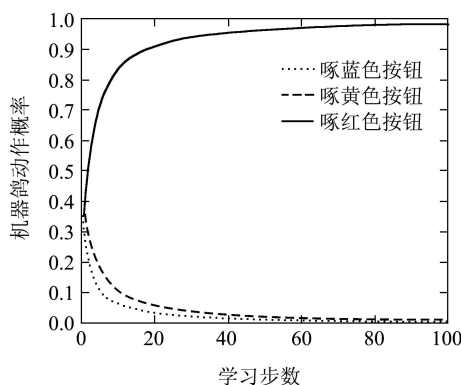


图3 机器鸽动作概率变化情况

Fig. 3 Change of the action probabilities for the pigeon

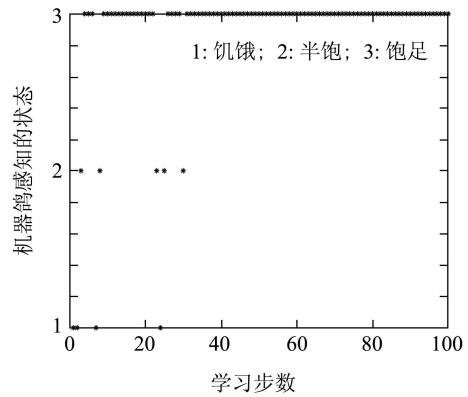


图4 机器鸽状态变化情况

Fig. 4 Change of the pigeon's states

4.2 维纳机器蠕虫实验(Wiener worm experiment)

控制论奠基者Wiener在《控制论》一书中说^[16-17]:“一定形式的视觉–肌肉反馈系统即便在蠕虫这样低等的动物世界中也是极为重要的”。为了让机器表现出蠕虫的习性与行为,并由此论述动物和机器中相似的控制机制,Wiener在《控制论》中虚拟了一个机器蠕虫。他所构想的这一机器蠕虫虽然与真实蠕虫的形态和构成不同,但都具有相似的感觉–运动系统^[17]。

自然界中的某些蠕虫,如蚯蚓,具有负趋光的特性。本环节将通过负趋光实验再现Wiener机器蠕虫,而本文所提出的系统模型则作为机器蠕虫的神经系统发挥作用。

正如维纳在《控制论》中所表达的那样,机器蠕虫的形态并不重要,重要的是机器蠕虫具有类似自然生命的感知–运动系统,具有生物蠕虫的习性和行为。因此本仿真实验环节并不考虑形态上的仿生,将机器蠕虫设定为一个带有光强传感器的轮式圆形机器人。如图5所示,该机器人半径为20 cm,行走机构采用双轮差动式运动底盘,在机器人左右两侧安装有轮 w_L 和 w_R ,由直流伺服电机驱动,尾部有一个起支撑作用的方向轮 w_F 。

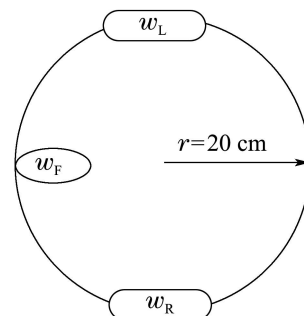


图5 机器人结构简图

Fig. 5 Structure of the wheeled robot

实验环境为一个 $4\text{ m} \times 4\text{ m}$ 的正方形空间,在几何中心(2,2)处放置一点光源,如图6所示。实验环境中

各处光强依据与光源距离呈放射状均匀分布, 设光源的坐标为 $(x_{\text{light}}, y_{\text{light}})$, 则环境中坐标 (x, y) 处的光照强度 ϕ 依据式(14)计算:

$$\phi = K \frac{\Phi_{\max}}{(x - x_{\text{light}})^2 + (y - y_{\text{light}})^2}, \quad (14)$$

式中: Φ_{\max} 为最大光强, 即光源处光强, 本实验中设定为10单位; K 为光强系数, 本实验中设定为1. 显然, 某处的光强与 Φ_{\max} 成正比, 与该处和光源距离的平方成反比.

初始化阶段, 令感知层神经元个数 $l = 1$, 其输入代表感知到的光强; 运动层神经元个数 $n = 6$, 代表机器人可以选择6种角度前进, 即 $(0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ)$, 这些角度均以当前前进方向为基准, 如 60° 表示选择从当前方向逆时针旋转 60° 所确定的方向, 其余以此类推; 隐层神经元个数 m 设定为10; 令权值矩阵 $W1, W2$ 在 $[0, 0.1]$ 随机取值, 令机器蠕虫从光源处出发, 即坐标 $(2, 2)$ 处, 前进速度设定为10 cm/s, 学习速率 η 设定为0.5.

由于本实验模拟的是蠕虫的负趋光特性, 因此将式(14)设定为负理想度计算公式, 即: 某点处机器蠕虫状态的负理想度等于该点光照强度, 光照强度越大, 则负理想度越大. 同时, 以式(13)作为取向函数的计算公式, 并且按照式(15)计算机器蠕虫的位置转移:

$$\begin{cases} x_{\text{new}} = x_{\text{old}} + v \times t_s \times \cos \theta_k, \\ y_{\text{new}} = y_{\text{old}} + v \times t_s \times \sin \theta_k, \end{cases} \quad k = 1, 2, \dots, 6, \quad (15)$$

式中: $x_{\text{new}}, y_{\text{new}}$ 代表动作选择后机器人新的横、纵坐标, 类似的, $x_{\text{old}}, y_{\text{old}}$ 代表选择前机器人的横、纵坐标; v 为机器人移动速度, t_s 为机器人传感器采样时间, θ_k 表示所选方向在以机器人圆心为极点、前进方向为极轴建立的极坐标系中的弧度值.

各模块结构及相关常数设定好以后, 模型根据如图2所示的流程运行. 实验结果如图6所示.

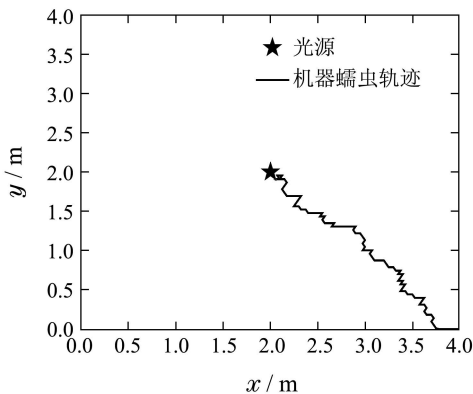


图 6 机器蠕虫实验结果

Fig. 6 Result of the Wiener worm experiment

从图6可以看出, 机器蠕虫从光源点出发, 沿着远离光源的路径移动, 直至抵达环境边界, 成功地模拟(表现)了生物蠕虫的负趋光行为. 在这一过程中, 机器蠕虫表现出具有和生物蠕虫类似的“感觉-运动”系统: 都能对感知到的光强信息作出反应, 进而根据生物取向性(负趋光)选择最优动作, 而动作的选择又将改变智能体所处状态, 使得感知信息也随之变化.

这样的“感知-行动闭环”建立在操作条件反射的基础上, 从图6中可以看出, 机器蠕虫开始的轨迹并非严格按照负趋光要求前行, 有折返的现象, 但随着学习次数增加, “选择光强弱的地方前行”这一行动由于符合生物内在取向, 其概率不断增加, 反映在实验结果上就是机器蠕虫行动的轨迹逐渐向光强弱的地方延伸, 折返现象消失. 这一实验结果表明, 在“刺激-反应-强化”的反复作用下, 机器蠕虫已经建立起操作条件反射, 表现出负趋光的行为.

4.3 对比实验(Contrast experiment)

黄静等人^[14-15]提出了一种基于操作条件反射原理的学习模型OCLM. 该学习模型为九元组, 由状态空间、操作行为空间、概率分布函数、仿生学习机制、系统熵等组成, 其核心是学习机制的设计与定义. 该学习机制以“正强化操作概率增加, 负强化操作概率减少”为原则来调整各状态操作行为概率分布, 由此实现操作条件反射. 虽然该模型对操作条件反射学习进行了数学建模, 但是没有明确阐述生物的感知-运动机制与模型的联系, 缺乏对感觉运动神经系统的仿生.

为了进一步探讨本文所提出模型ASSOCF在模拟生物感知-运动机制、实现操作条件反射学习方面的能力, 进行了两组对比实验. 对比实验采用OCLM作为人工脑, 在关键参数(学习速率、仿真环境设置、机器蠕虫行走速率等)保持一致的条件下分别复现了斯金纳鸽子实验及维纳蠕虫实验, 实验结果如图7-8所示.

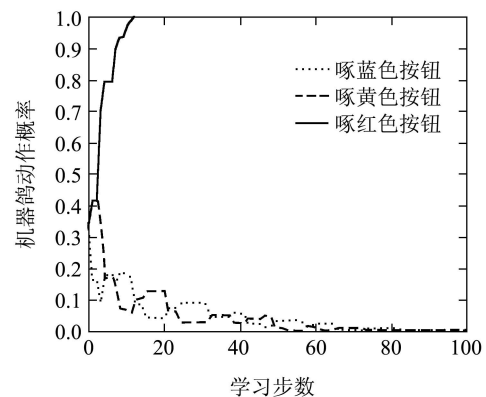


图 7 对比试验中机器鸽动作概率变化情况

Fig. 7 Change of the action probabilities for the pigeon in the contrast experiment

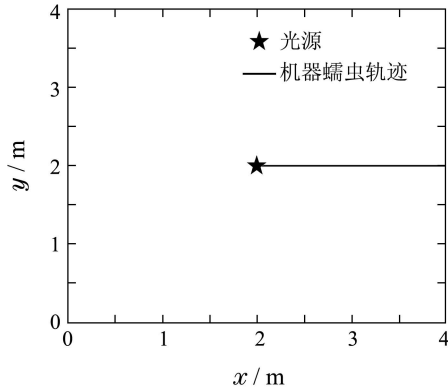


图8 维纳机器蠕虫对比实验结果

Fig. 8 Result of the Wiener worm experiment in the contrast experiment

从图7-8所展示的实验效果可以看出,基于OCLM的对比实验基本达到了预期的效果,实现了对两个经典实验的复现.例如,图7中的机器鸽在经过学习之后,啄红色按钮的概率逐渐上升到1,而啄黄、蓝两色按钮的概率则逐渐趋近于0,说明已习得了最优动作(啄红色按钮).图8中的机器蠕虫也能从光源出发,以负趋光为行为逻辑,远离光源运动.

但是,对比实验结果也反映出OCLM相对于ASSOCF的不足.从图7中可以看到,啄红色按钮概率的上升不如ASSOCF平滑(见图3),提示该动作概率在学习过程中并非严格递增,有一定时间内保持不变的现象.同时,图7中啄黄色按钮的概率与啄蓝色按钮的概率并非完全单调递减,有增加的情况存在.另外,啄黄色按钮的概率有时甚至会低于啄蓝色按钮的概率,这些现象与斯金纳鸽子实验并不吻合.图8中机器蠕虫虽然能表现负趋光行为,但行为选择单一,多次实验(20次)所得行走路线始终如图8所示.而ASSOCF则可表现出多样化的行走轨迹,图9所示即为机器蠕虫选择的另一条负趋光路线.显然,ASSOCF表现出的多样性更符合实际情况,说明该模型能更好的模拟和复现蠕虫负趋光行为.

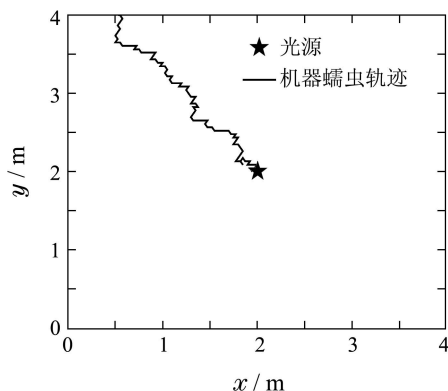


图9 另一次机器蠕虫实验结果(基于ASSOCF模型)

Fig. 9 Another result of the Wiener worm experiment based on ASSOCF

为了分析两种模型在收敛速度上的差异,在完成对比实验时对程序的运行时间进行了记录.同时,为了排除单次实验中可能出现的偶然因素,以程序运行10次之后的平均时间作为考量依据,结果如表3所示.

表3 程序运行时间对比

Table 3 Program running time comparison

	斯金纳鸽子实验	维纳机器蠕虫实验
OCLM	0.4037 s	129.7701 s
ASSOCF	1.9398 s	0.4179 s

从实验数据可以看到,当状态空间较小时(斯金纳鸽子实验),两种模型的收敛速度差距不大,ASSOCF并没有明显优势,甚至略慢;但对于状态空间较大问题(维纳机器蠕虫实验),ASSOCF收敛速度明显快于OCLM,说明相对于OCLM,ASSOCF对状态空间大小不敏感,更适于解决复杂问题,在收敛速度方面表现出更好的性能.

5 结论(Conclusions)

认知科学、人工智能及机器人学之间的相互渗透融合直接导致了认知机器人学这一新兴研究领域的诞生.这一交叉学科诞生背后所蕴含的思想与维纳的《控制论》一脉相承,是《控制论》宏观思想在具体问题的延伸和发挥.基于这样的背景,本文试图通过仿生的思路以模拟生物感觉运动系统的方式构造一种人工脑;同时,本文借鉴了美国心理学家斯金纳的操作条件反射理论,使其具有操作条件反射机能,进而提高机器人的认知及智能水平.

本文所构造的具有操作条件反射机能的人工感觉系统ASSOCF由感知-运动映射模块SMMM及操作条件反射计算模块OCCM组成,前者基于前馈神经网络,模拟(表现)了生物感觉运动系统中从感知到运动的映射关系;后者包括负理想度计算NIDCSM、取向函数值计算OFSM、动作选择ASSM、状态转移STSM等4个子模块,体现了操作条件反射机能中感知与运动相互影响的闭环联系.

验证模型的正确性,本文复现了两个经典实验,分别是斯金纳鸽子实验和维纳机器蠕虫实验,同时还进行了对比实验,与黄静等人提出的操作条件反射模型OCLM在仿生效果、收敛速度等方面进行了对比.实验结果表明,本模型能够模拟(表现)生物的感觉运动神经系统,具有较强的自学习能力,以建立操作条件反射的形式习得最优动作,表现出类似生物适应环境的自适应性.较之OCLM,本系统能更好的表现生物的自学习行为,在收敛速度上具有更好的性能,更适宜于解决状态空间大的复杂问题.

本模型的价值可以从科学及工程两个角度来进行分析.从科学的角度,本模型可以作为认知模型应用

于机器人, 开展人工智能及认知心理学的具身化(embodied)研究: 从工程的角度, 本模型可以作为机器人的人工脑, 实现对机器人的智能控制, 应用于包括机器人的避障导航、路径规划以及运动控制等问题。

另一方面, 本模型对生物感觉运动神经系统的模拟还仅仅停留在功能层面, 缺乏更细致的神经生理学方面的讨论。这些将作为未来的研究方向, 有待进一步的深入研究。

参考文献(References):

- [1] BROOKS R A. From earwigs to humans [J]. *Robotics and Autonomous Systems*, 1997, 20(2): 291 – 304.
- [2] BROOKS R A, BREAZEAL C, MARJANOVIC M, et al. *The Cog Project: Building a Humanoid Robot* [M]. Berlin: Springer-Verlag, 1999: 52 – 87.
- [3] NATALE L, ORABONA F, BERTON F, et al. From sensorimotor development to object perception [C] // *Humanoid Robots, 2005 5th IEEE-RAS International Conference on*. Piscataway, NJ: IEEE Press, 2005: 226 – 231.
- [4] KUNIYOSHI Y, SANGAWA S. Early motor development from partially ordered neural-body dynamics: experiments with a cortico-spinal-musculo-skeletal model [J]. *Biological Cybernetics*, 2006, 95(6): 589 – 605.
- [5] KIMURA H, FUKUOKA Y, COHEN A H. Adaptive dynamic walking of a quadruped robot on natural ground based on biological concepts [J]. *The International Journal of Robotics Research*, 2007, 26(5): 475 – 490.
- [6] CHAO F, LEE M H. An autonomous developmental learning approach for robotic eye-hand coordination [C] // *Proceedings of the IASTED International Conference*. Calgary: ACTA Press, 2009: 13 – 18.
- [7] MATHEWS Z, VERSCHURE P F M J. PASAR: an integrated model of prediction, anticipation, sensation, attention and response for artificial sensorimotor systems [J]. *Information Sciences*, 2012, 186(1): 1 – 19.
- [8] SKINNER B F. *The Behavior of Organisms: an Experimental Analysis* [M]. New York: D. Appleton-Century Company, 1938: 110 – 150.
- [9] ZALAMA E, GAUDIANO P, CORONADO J L. *Obstacle Avoidance by Neans of an Operant Conditioning Model* [M]. Berlin: Springer, 1995: 471 – 477.
- [10] ITOH K, MIWA H, MATSUMOTO M, et al. Behavior model of humanoid robots based on operant conditioning [C] // *2005 the 5th IEEE-RAS International Conference on Humanoid Robots*. Tsukuba: IEEE, 2005: 220 – 225.
- [11] SALOTTI J M, LEPRETRE F. Classical and operant conditioning as roots of interaction for robots [C] // *Proceedings of the Workshop From Motor to Interaction Learning in Robots, Conference on Intelligent Robotics Systems*. Nice: Springer, 2008: 253 – 262.
- [12] CYR A, BOUKADOUM M, THERIAULT F. Operant conditioning: a minimal components requirement in artificial spiking neurons designed for bio-inspired robot's controller [J]. *Frontiers in Neurobotics*, 2014, doi: 10.3389/fnbot.2014.00021.
- [13] 阮晓钢, 戴丽珍, 于乃功, 等. 一种自治操作条件反射自动机 [J]. *控制理论与应用*, 2012, 29(11): 1452 – 1457. (RUAN Xiaogang, DAI Lizhen, YU Naigong, et al. An autonomous operant conditioning automaton [J]. *Control Theory & Applications*, 2012, 29(11): 1452 – 1457.)
- [14] HUANG J, RUAN X G, FAN Q W, et al. Operant conditioning learning model in the bionic experiment [J]. *Applied Mechanics and Materials*, 2013, 373: 255 – 264.
- [15] 阮晓钢, 黄静, 范青武, 等. 一种基于操作条件反射原理的学习模型 [J]. *控制与决策*, 2014, 29(6): 1016 – 1020. (RUAN Xiaogang, HUANG Jing, FAN Qingwu, et al. A learning model based on operant conditioning principles [J]. *Control and Decision*, 2014, 29(6): 1016 – 1020.)
- [16] WIENER N. *Cybernetics* [M]. New York: J. Wiley, 1948: 167 – 178.
- [17] 阮晓钢. *神经计算科学——在细胞的水平上模拟脑功能* [M]. 北京: 国防工业出版社, 2006: 548 – 549. (RUAN Xiaogang. *Computational Neuroscience: Simulating the Brain Mechanism on the Level of Neuron* [M]. Beijing: National Defence Industry Press, 2006: 548 – 549.)

作者简介:

黄静 (1979–), 女, 博士研究生, 讲师, 研究领域为人工智能、智能控制及认知科学, E-mail: aiandrobot@163.com;

阮晓钢 (1958–), 男, 博士, 教授, 研究领域为机器人、人工智能及认知科学, E-mail: adrxg@bjut.edu.cn;

于乃功 (1966–), 男, 博士, 教授, 研究领域为机器人、人工智能及认知科学, E-mail: mymailhj@sohu.com;

张晓平 (1991–), 女, 博士研究生, 研究领域为机器人、人工智能及认知科学, E-mail: icandai@163.com;

魏若岩 (1985–), 男, 博士研究生, 研究领域为图像处理、人工智能、制导与控制等, E-mail: databasehomework@163.com;

范青武 (1977–), 男, 博士, 副教授, 研究领域为遗传算法、人工智能, E-mail: databaseforu@163.com.