

基于拟人化推理策略的行为模仿

尚 涛^{1,2}, 吴 威^{1,2}, 刘建伟³

(1. 虚拟现实技术与系统国家重点实验室, 北京 100191; 2. 北京航空航天大学计算机学院, 北京 100191;
3. 北京航空航天大学电子信息工程学院, 北京 100191)

摘 要: 本文从模仿机制的再现环节入手, 采用知识使用和行动预见为主的拟人化推理策略, 来实现有效的行为模仿. 具体地, 以模糊集合同的 Hausdorff 距离作为知识使用的尺度, 导入知识半径到距离型模糊推理方法当中实现知识的选择使用; 以预见控制的基本思想为指导, 设计行动预见模型作为知识使用的高层决策, 优化知识使用策略的参数. 并以驾驶行为模仿为例, 验证了拟人化推理策略对于动态知识使用的有效性, 实现了即学即仿的模仿效果.

关键词: 模仿; 模糊推理; 知识使用; 行动预见

中图分类号: TP181 **文献标识码:** A **文章编号:** 0372-2112 (2010) 2A-083-06

Behavior Imitation Based on Humanoid Reasoning Strategy

SHANG Tao^{1,2}, WU Wei^{1,2}, LIU Jian-wei³

(1. State Key Laboratory of Virtual Reality Technology and Systems, Beijing 100191, China;
2. School of Computer Science, BeiHang University, Beijing 100191, China;
3. School of Electronic and Information Engineering, BeiHang University, Beijing 100191, China)

Abstract: This paper focuses on the recurrence component of imitation mechanism and presents humanoid reasoning strategies such as knowledge use and action preview to implement effective behavior imitation. Concretely, by defining the Hausdorff distance between fuzzy sets as the metric of knowledge use, we combined knowledge radius into the distance-type fuzzy reasoning method for the implementation of selective knowledge use; according to the main idea of preview control, we designed the action preview model as high level decision of knowledge use to optimize the parameters of knowledge use strategy. Finally we presented the results of driving behavior imitation, demonstrated the validity of humanoid reasoning strategies in dynamic knowledge use, and realized quick learning-and-imitating performance.

Key words: imitation; fuzzy reasoning; knowledge use; action preview

1 引言

模仿是指有意或无意仿效他人的语言或行为, 使自己的语言或行为与对方相似. 从婴幼儿最初的言语行为的获取到成人间的相互影响, 模仿在问题解决过程中表现出来的有效性为众人所共知. 模仿与高级形式的智力有关, 是一种从根本上与人类特有的智力形式相联系的能力, 特别是语言、文化和理解他人思想的能力. 社会学理论创始人·认知理论之父——阿尔伯特·班杜拉 (Albert Bandura) 对模仿学习进行了深入的研究, 认为婴儿主要是通过对各种社会言语模式的模仿学习而获得言语能力的, 并将社会模仿学习分为下述四个必备的过程: ①注意、②保持、③运动再现、④动机建立^[1]. 鉴于模仿的有效性, 国内外智能科学和机器人研究人员高度关

注模仿的研究和应用^[2~5], 尤其在机器人研究制造领域, 通过模仿人的结构和行为, 使机器人能够自主适应环境的变化. 近年, 随着神经科学和认知科学迅速发展, 相关的研究成果为揭示模仿机制的本质提供了新的参考依据^[6~8]. 应用模仿任务的神经成像研究表明人脑中存在类似猴镜像神经元的皮层, 包括额下回、顶叶和颞上沟, 这些皮层在对行为的理解、模仿学习新的复杂行为以及固有行为的协调方面起着重要作用, 且这些皮层在不同实验的模仿任务中表现出相当一致的激活, 体现了模仿过程中某种动态的知识运用策略. 可见, 建立体现人脑运用的拟人化知识使用策略对于新型模仿机制的构建是一个重要的发展方向, 对于模仿的研究和应用具有重要的参考意义.

2 行为模仿机制

从工学研究角度,可以将模仿细化为观察、学习、记忆、再现、评价的五个环节.具体地说,观察是通过视觉以一定编码方式将环境信息和目标模仿对象传递到人脑,编码方式是对环境信息和目标模仿对象的等价描述,既可能是具体物理信息,也可能是抽象拓扑信息;学习是将输入信息划分为条件信息和行为信息,采用某种机制从获取来的信息中提取更高抽象层次的知识,并以一定编码方式存储到代表不同功能的各个脑部位;记忆是将知识存储到人脑记忆单元,知识的有效性依据在存储单元的差异在脑内存储时间的长短也不同;再现是在相同或相似的环境下依据脑内已经存储的知识,判断行为,并给出行为指令,指令将进一步支配肌肉;评价是对行为与环境作用的效果进行比较,用于反馈给大脑,进行自身调整.总之,模仿机制可以归纳为图1所示的原理图.①-⑤描绘了模仿发生的主要过程,其中①-④内在模型具有输入-输出功能,能实现依赖知识的感知-运动状态,⑤是以内在模型为基础进行更新行为,为达到某一目标所需要运动计划而准备.虚线部分表示用于没有强化条件的新输入,表现更强适应性的模仿.

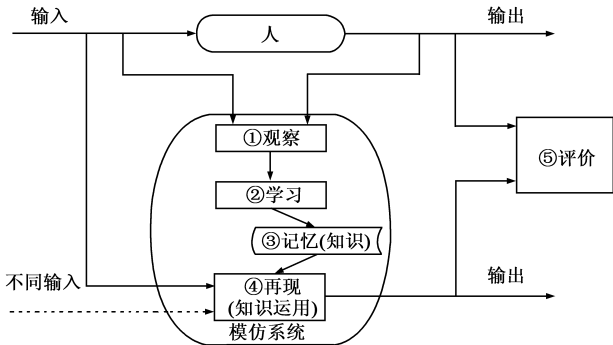


图1 模仿机制原理图

本文以行为模仿为研究对象,构建模仿机制的具体模型.通过观察环节的数据采集,可以完成环境-行为的信息表达,常用相对关系的物理量就可以用来描述人与环境的关系.通常行为模仿过程的数据有限性,可以简化对机器学习方法的要求.因而,集中在再现环节的推理方法建立体现人脑运用的拟人化策略是行为模仿机制的关键环节.尽管推理方法多种多样^[9,10],但推理的过程本质是求解问题,问题求解的质量与效率不仅依赖于所采用的求解方法,而且还依赖于求解问题的策略,即推理策略.由于模仿行为的产生具有模糊特性,基于模糊集合的模糊性和模糊推理的鲁棒性,本研究采用模糊推理方法实现再现环节,通过模拟人脑知识运用的特征,建立多样的拟人化推理策略,来实现有效的行为模仿.

3 拟人化推理策略(一)知识使用

行为模仿的过程中,人通过观察获取目标信息,激活大脑不同部位的知识加以运用来决策自己的行为.根据知识的关联程度,知识使用策略可以在推理方法中灵活运用.

3.1 知识使用尺度

知识使用的多样性取决于如何表示事实与知识之间关联程度.定量化事实与知识之间的关联程度主要包括两种:通常的隶属度函数的计算(经典的 Zadeh、Mamdani 模糊推理方法),和模糊集合之间的距离值(距离型模糊推理方法^[11]).当采用模糊集合之间的 Hausdorff 空间距离作为模糊集合之间关联程度的评价基准时,距离越大,相关性越小;距离为零时,模糊集合完全一致,相关性最大.其有效性已在理论分析和实际应用中得到了验证^[11,12],因而本文采用模糊集合之间的 Hausdorff 空间距离作为一种的知识使用尺度来实现推理过程中知识使用策略.在 Hausdorff 空间内,以输入事实为圆心,以各规则的前提为离散点,计算圆心和各个离散点的距离,通过用某一距离值作为圆半径来圈定使用的规则.随着圆半径的变化,圈定使用的规则的个数也发生变化.为了简化运算,以某一个规则的离散点对应距离值作为半径,定义圈定的规则个数为知识半径.导入离散的整数值范围的知识半径代替连续的实数距离值有三个主要特点:知识半径的范围明确,为不连续区间的大于1的正整数;用较少的计算量可探索最优的知识半径;可实现规则的选择性使用.例如,图2所示规则库存储有3个规则,事实与各规则的前提的距离值依次递减,知识半径的变化范围 $\{2,3\}$.

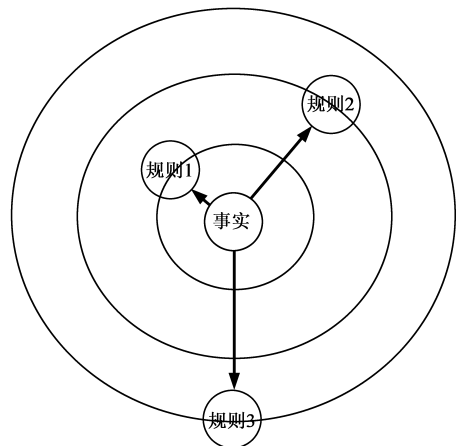


图2 知识半径概念示意图

知识半径用来实现在模仿过程中知识的选择使用,如果知识半径在推理过程中为常数,其描述在推理过程中使用的知识个数不变,称之为静态知识半径;如果知识半径在推理过程中为变量,其描述在模仿过程

中根据场景变化动态改变使用知识的个数,称之为动态知识半径.可见,知识半径是一个非常灵活的知识运用控制策略.

3.2 结合知识半径的模糊推理方法

基于 Hausdorff 空间距离运算的距离型模糊推理方法满足渐近特性和分离原则(Modus Ponens).本文导入知识半径到距离型模糊推理方法当中,在推理过程中利用知识半径的变化对知识加以选择使用.

通常使用以下的推理对象.

R^i : if $x_1 = A^{i1}, x_2 = A^{i2}, \dots, x_m = A^{im}$ then $y = B^i$

$i = 1, 2, \dots, n, j = 1, 2, \dots, m.$

事实: $x_1 = A^1, x_2 = A^2, \dots, x_m = A^m$

输出: B

其中, x_1, \dots, x_m 和 y 分别表示输入变量和输出变量, $A^{i1}, \dots, A^{im}, B^i$ 是相应的模糊集合, n 是规则的个数, m 是前提的个数.

导入知识半径的距离型模糊推理方法由以下四个步骤构成.

STEP1: 根据式(1)计算模糊集合的事实 A^i 和第 i 个规则中第 j 个前提 A^{ij} 的距离值 d_{ij} .

$$d_{ij} = d_{ij}(A^{ij}, A^i) = |a^{ij} - a^i| \quad (1)$$

在式(1)中,在适用于单点型模糊集合的情况下,两个模糊集间的距离计算公式 A^{ij} 和 A^i 简化为上式.事实 A^i 和前提 A^{ij} 相距越远,计算的距離值 d_{ij} 越大.当事实 A^i 和前提 A^{ij} 完全一致的时候,距离值 d_{ij} 为 0.

STEP2: 根据式(2),计算第 i 个规则的所有前提与事实的距离值 d_i .

$$d_i = \sum_{j=1}^m d_{ij}(A^i, A^{ij}) \quad (2)$$

距离值 d_i 定量表示事实与第 i 个规则的差别. d_i 值越大,意味着事实与第 i 个规则条件的关联性越弱.

STEP3: 按着从小到大的顺序排序距离值 d_i , 根据知识半径大小 q 和相应距离阈值 δ , 如式(3)进行距离变换,其中 $\text{sgn}(\cdot)$ 是信号函数.

$$d'_i := d_i \cdot \epsilon^{\text{sgn}(\delta - d_i) - 1}, \epsilon \rightarrow 0 \quad (3)$$

$$\text{sgn}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases} \quad (4)$$

STEP4: 根据式(5),计算输出结果 y_0 .

$$y_0 = \frac{\sum_{i=1}^n [b^i \prod_{j=1, j \neq i}^n d'_j]}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d'_j} = \frac{\sum_{i=1}^n [b^i \prod_{j=1, j \neq i}^n \{d_j \cdot \epsilon^{\text{sgn}(\delta - d_j) - 1}\}]}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n \{d_j \cdot \epsilon^{\text{sgn}(\delta - d_j) - 1}\}} \quad (5)$$

从式(5)可知,前提与事实的距离越近,其规则的结论部分在推理结果上起的作用越大.尤其是,当前提与事实完全一致时,其规则的结论部分作为输出结果

输出.

3.3 选择性知识使用的性质

导入知识半径的距离型模糊推理方法不仅继承了原方法的性质(推导过程详见文献[11]),而且派生了关于选择性知识使用的新特性.

性质: 知识半径的值 q 作用于模糊推理中所有规则的重要性,知识半径限定的规则的权值增大,非限定的规则的权值减小为零.

证明: 原推理方法的输出如式(6)表示,可转化为式(7),其中 w_i 表示规则的权值.

$$y_0 = \frac{\sum_{i=1}^n [b^i \prod_{j=1, j \neq i}^n d_j]}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d_j} = \sum_{i=1}^n \left[\frac{b^i \prod_{j=1, j \neq i}^n d_j}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d_j} \right] \quad (6)$$

$$y_0 = \sum_{i=1}^n [b^i w_i], w_i = \frac{\prod_{j=1, j \neq i}^n d_j}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d_j} \quad (7)$$

当 w_i 为 1 时,规则完全被选择;当 w_i 为 0 时,规则未被选择;当 w_i 介于 0 与 1 之间,规则被不完全选择.因此权值变化意味着规则选择使用的程度.

导入知识半径的距离型模糊推理方法的输出如公式(5)表示,可转化为式(8).

$$y'_0 = \sum_{i=1}^n [b^i w'_i], w'_i = \frac{\prod_{j=1, j \neq i}^n d'_j}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d'_j} \quad (8)$$

因此,新的权值 w'_i 变化可被用来评价知识半径的效用.

(1) 当事实 A^1, \dots, A^m 满足分离原则,即条件满足 $\exists p \in \{1, 2, \dots, n\}$ 且 $\sum_{j=1}^m d(A^j, A^{pj}) = 0, d'_p = d_p \cdot \epsilon^{\text{sgn}(\delta - d_p) - 1} = 0 \cdot \epsilon^{\text{sgn}(\delta - 0) - 1} = 0$, 规则的权值变为 $w'_p = w_p = 1$. 同时,其他规则的权值 $w'_i = w_i = 0, i = 1, \dots, n, i \neq p$. 结果上,唯一被选中的第 p 规则被使用,输出 $y'_0 = y_0$, 推理仍就满足分离原则. 与原方法相比,所有规则的权值并没有发生改变.

(2) 当事实 A^1, \dots, A^m 不满足分离原则,即条件满足 $\forall i \in \{1, 2, \dots, n\}, \sum_{j=1}^m d(A^j, A^{ij}) \neq 0$, 事实与规则 R_i 前提的距离 d_i 都不为 0. 知识半径限定的规则权值变化如式(9)、(10)和(11).

原方法的权值描述为公式(9).

$$\therefore \forall i \in \{1, 2, \dots, n\}, d_i > 0$$

$$\therefore w_i = \frac{\prod_{j=1, j \neq i}^n d_j}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d_j} = \frac{\frac{1}{d_i}}{\sum_{j=1}^n \left(\frac{1}{d_j} \right)} \quad (9)$$

知识半径限定的规则权值如式(10).

$$w'_i = \frac{\prod_{j=1, j \neq i}^n d'_j}{\sum_{i=1}^n \prod_{j=1, j \neq i}^n d'_j} = \frac{1}{d'_i} \quad (10)$$

当 $S = \{i | d_i \leq \delta, i = 1, \dots, n\}$ 条件满足, $d'_i = d_i$; 当 $\{i | d_i \leq \delta, i = 1, \dots, n\}$ 不满足, $d'_i \rightarrow \infty$. 其中 δ 是知识半径 q 对应的距离阈值. 因此, 知识半径限定的规则的权值如下描述.

$$w'_i = \frac{1}{d'_i} = \frac{1}{d_i} \geq w_i \quad (11)$$

另外, 由于 $d'_i \rightarrow \infty$, 知识半径不限定的规则的权值变为 $w'_i \rightarrow 0$.

总之, 知识半径的引入实现了当前规则的不同程度使用, 保证基于知识半径的推理性能必然强于原方法推理性能.

4 拟人化推理策略(二)行动预见

行为模仿的过程中, 人通常需要预测行为产生的结果才会决定自己的行为. 这种策略是在知识使用的前提下, 根据多种知识使用的假想和优化标准, 才明确知识使用的最佳策略, 可认为是知识使用的高层决策.

4.1 预见控制的思想

预见控制理论^[13]是随着自动控制理论和工程应用技术的发展而逐步完善起来的, 不仅注意过去及现在的目标值, 而且注意未来的目标值, 核心思想是通过注意未来的目标值, 使目标值与受控量间的偏差整体地最小, 属于使横贯全控制期间的某一评价函数取最小值的最优控制理论, 如图 3 表示.

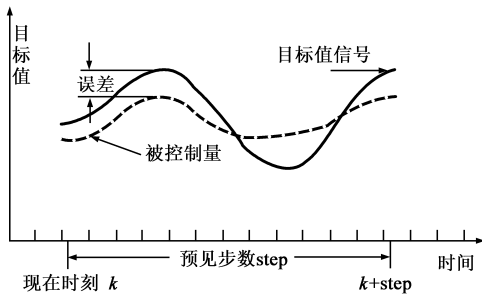


图3 预见控制的概念示意图

本文将预见思想应用到推理策略当中: 利用示范者的行为信息作为系统的当前状态的将来目标值, 通过产生一个虚拟目标值, 让系统连续跟踪这个目标值信号, 从而提高行为模仿的精度.

4.2 行动预见模型

在推理模型的基础上, 利用示范者的行为信息作为目标, 通过预见评价来确定知识使用策略的最优参

数, 包括知识半径和预见步数的值, 如图 4 所示.

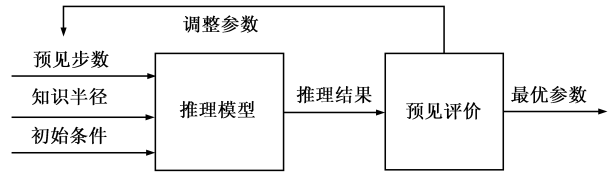


图4 行动预见模型

预见评价函数与预见步数有关. 预见控制的极限性质表明使用未来目标值 1 步信息即可实现最优控制, 时间较远的未来信息并不利于效果提高. 作为比较的基准, 通常使用目标值与被控制量的误差函数作为评价函数.

基于上述模型, 行动预见策略如下记述:

STEP1: 在知识半径范围内, 将知识半径候选值和初始化条件, 共同输入到推理模型;

STEP2: 利用推理模型计算预见步数内的推理结果, 计算基于推理结果的评价值, 确定最优的知识半径;

STEP3: 基于最优的知识半径进行推理.

5 实验与分析

5.1 驾驶行为模仿

在驾驶环境中, 人需要依据环境变化回避障碍物到达目标点, 是一种典型可以应用到移动机器人路径规划的场景, 且驾驶过程中行为连续变化, 模仿起来难度较大, 适于验证方法的有效性. 开发的驾驶模拟环境中包含有目标、障碍物、驾驶代体. 且通过设置障碍物的属性, 如个数、面积、速度等可以改变驾驶环境的复杂程度. 更主要的是, 驾驶代体可以由人或者推理模块智能控制进行模仿演示.

5.2 学习过程

示范者任意完成一次操作避障且成功到达目标. 采用结合发生概率的距离型学习算法^[12]从示范者的行为数据当中获取模糊规则. 该学习算法对知识描述要求简单, 生成用单点模糊集表示的模糊规则. 学习前的参数设定和学习后的结果如表 1 所示.

表 1 学习过程的参数

参数含义	参数值
障碍物坐标	O1(600, 100), O2(300, 200), O3(500, 500)
数据的个数	333 个
学习指定误差	方向 15 度, 力 3 牛顿
结果规则个数	方向 101 个, 力 17 个

* 坐标系左上角为原点, 向右向下为正

5.3 模仿效果分析

以知识半径作为知识使用策略, 以路径面积误差作为行为预见评价函数, 在推理过程中, 动态地确定最优知识半径. 依据不同的预见步数实现了有效的模仿.

模仿效果如图 5 所示,策略使用的参数如表 2 所示.

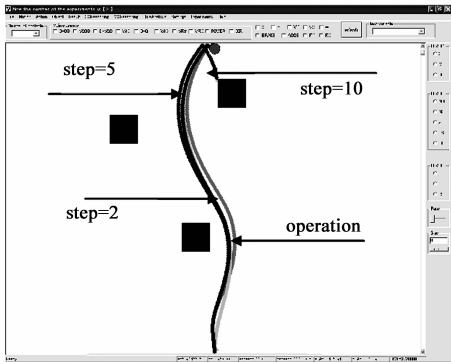


图5 不同的模仿路径

表 2 策略的参数

预见步数	知识半径值的变化		面积误差
	最终值	平均值	
2	3	10.79	2.23
5	3	11.57	1.79
10	2	9.88	99.59

5.4 推理策略有效性分析

如图 6、7、8 所示,推理过程中知识半径的值自适应地变化,不仅使用动态知识半径实现了行为模仿,也肯定了知识使用的重要性,可见基于行动预见策略的动态知识半径更加接近人脑的知识运用策略情况.即使是不同的预见步数,在推理过程中都是经过知识半径的振荡并最终快速安定下来,而且平均值皆接近于 10. 预见步数是 2 时,较大振幅的振荡数是 6;预见步数是 5 时,较大振幅的振荡数是 5;预见步数是 10 时,较大振幅的振荡数是 2. 总之,随着预见步数的增长,知识半径的振荡数有所减少.

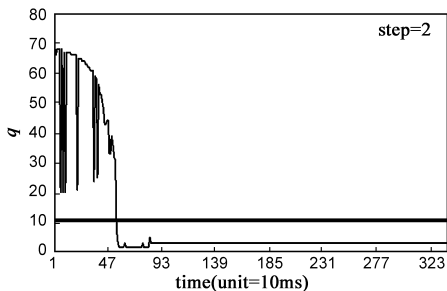


图6 预见步数是2时动态知识半径的变化

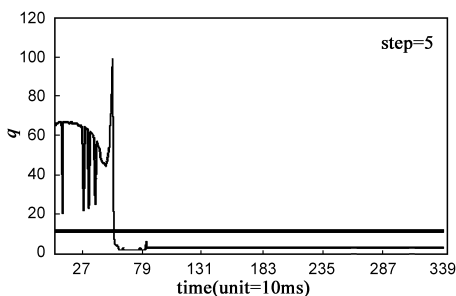


图7 预见步数是5时知识半径的变化

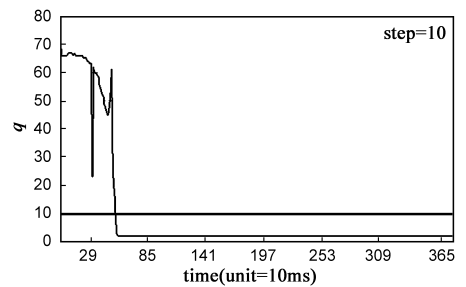


图8 预见步数是10时知识半径的变化

同时,图 9 显示了在不同预见步数作用下,推理过程中评价函数值的变化.随着预见步数的增长,评价函数值也增长,自然计算量增大.综合考虑振荡数和评价函数值,选择合适的预见步数对于实现行动预见策略也是非常重要的.因为在知识半径变化过程中,较少的振荡状态保证平稳的控制;较小的评价函数值可以提高模仿的实时性.

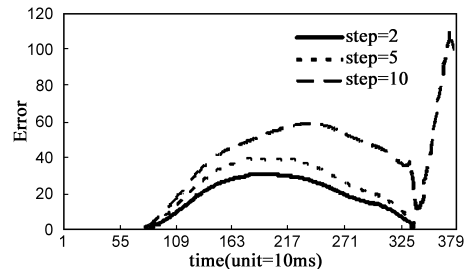


图9 不同预见步数的作用下评价函数值的变化

6 结论

本文从知识使用和行动预见角度深入研究了拟人化的推理策略.以人行为模仿的机理为参考,一方面,以导入知识半径的距离型模糊推理方法作为知识使用策略的基础,并从理论上证明了选择性知识使用的性质;另一方面,将预见控制思想与知识使用策略结合到推理模型中生成行动预见模型,运用基于预见步数的性能评价函数,确定知识使用策略的最优参数,使得知识使用策略更加精确、有效、实用.实验结果表明提出的拟人化策略对于行为模仿是有效的,且需要进一步讨论动态的知识使用问题,如预见步数的选取问题,来优化推理过程中知识使用环节,实现高效的行为模仿.

参考文献:

- [1] Albert Bandura. Social Foundations of Thought and Action: A Social Cognitive Theory[M]. Prentice Hall, 1985.
- [2] Stefan Schall. Is imitation learning the route to humanoid robots [J]? Trends in Cognitive Science, 1999(3):233 - 242.
- [3] Dautenhahn K, Nehaniv C L. Imitation in Animals and Artifacts [M]. Cambridge, Massachusetts: the MIT press, 2002.
- [4] Wang S, Zhao X. Biomimetic pattern recognition theory and its applications[J]. Chinese Journal of Electronics, 2004, 13(3):

373 – 377.

- [5] 张利格, 黄强, 杨洁, 时有, 王志杰, Jafri Ali Raza. 仿人机器人复杂动态动作设计及相似性研究[J]. 自动化学报, 2007, 33(5): 522 – 528.

Zhang Lige, Huang Qiang, Yang Jie, Shi You, Wang Zhijie, and Jafri Ali Razal. Design of Humanoid complicated dynamic motion with similarity considered [J]. Acta Automatica Sinica, 2007, 33(5): 522 – 528. (in Chinese)

- [6] Carr L, Iaconi M, Dubeau M-C, Mazziotta JC, Lenzi GL. Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas [J]. Proceedings of the National Academy of Sciences of USA, 2003 (100): 5497 – 5502.

- [7] Buccino G, Binkofski F, Fink GR, et al. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study [J]. European Journal of Neuroscience, 2001 (3): 400 – 404.

- [8] 陈建敏, 金花. 人类动作模仿的神经机制研究进展[J]. 体育学刊, 2007, 14(8): 47 – 50.

Chen Jian-min, Jin Hua. Progress in researches on the neural mechanism of human motion imitation [J]. Journal of Physical

Education, 2007, 14(8): 47 – 50. (in Chinese)

- [9] 史忠植. 高级人工智能[M]. 北京: 科学出版社, 2006.
- [10] 王万良, 涂序彦. 人工智能及其应用[M]. 北京: 高等教育出版社, 2005.
- [11] S Y Wang, T Tsuchiya, M Mizumoto. A learning algorithm for distance-type reasoning method [J]. Biomedical Soft Computing and Human Science, 2000, 6(1): 61 – 68.
- [12] Tao Shang, Shuoyu Wang. Knowledge acquisition and evolution methods for human driving intelligence [J]. International Journal of Innovative Computing, Information and Control, 2006, 2(1): 221 – 236.
- [13] 土谷武士, 江上正. 数字预见控制[M]. 北京: 北京教育出版社, 1994.

作者简介:

尚 涛 男, 1976 年生于辽宁营口. 北京航空航天大学计算机学院博士后. 研究方向为智能系统、无线网络.

E-mail: shangtao@vrlab.buaa.edu.cn.

吴 威 男, 1961 年生于黑龙江哈尔滨. 北京航空航天大学计算机学院教授. 研究方向为虚拟现实及智能应用.