

# 基于 DFL 的自主学习子空间的公理体系研究<sup>\*</sup>

王 静 李凡长

(苏州大学计算机科学与技术学院 苏州 215006)

**摘 要** 动态模糊逻辑理论是一种新的有效处理动态、不精确、不确定、含糊信息的理论,自主学习历来是教育和心理学家共同关注的一个重要问题,也是当前机器学习领域研究中的一个热点问题。自主学习公理体系研究是自主学习理论及应用研究的基础。利用动态模糊关系,建立自主学习子空间的公理体系。该公理体系由几个相互独立的非常简洁的表达式构成,进一步深化了动态模糊和自主学习理论。

**关键词** 动态模糊逻辑,自主学习,子空间,动态模糊关系,公理体系

## Axiomatic System of Autonomic Learning Subspace Based on DFL

WANG Jing LI Fan-zhang

(College of Computer Science and Technology, Soochow University, Suzhou 215006, China)

**Abstract** Dynamic fuzzy logic theory is a new and effective theory when handling dynamic, imprecise, uncertain, or vague information. Autonomic learning has always been an important issue which is concerned by both educationalists and psychologists, and it is also a hot issue of machine learning research. Axiomatic system of autonomic learning is the foundation of the theory and application of autonomic learning. By using dynamic fuzzy relations, the axiomatic system of autonomic learning subspace was established. The axiomatic system was constituted by three independent expressions, further deepening the theory of dynamic fuzzy logic and autonomic learning.

**Keywords** Dynamic fuzzy logic, Autonomic learning, Subspace, Dynamic fuzzy relationship, Axiomatic system

### 1 引言

学习是人类智能的主要标志和获得智慧的基本手段。机器学习的研究就是希望计算机能像人类那样具有从现实世界获取知识的能力,同时进一步发现人类学习的机理和揭示人脑的奥秘。在 20 世纪 50 年代末和 60 年代初,人们就开展了基于神经网络的联接学习机制和学习算法的研究,但由于线性感知器的局限性,同时,传统的 Von Neuman 型数字计算机此时正处在发展的全盛时期,而且基于符号的人工智能得到迅速的发展,使得早期的基于联接机制的非符号学习被符号学习的研究所替代。20 世纪 80 年代以来,基于符号机制的机器学习的研究发展甚为迅速,其间研究了各种学习算法,如讲授学习、解释学习、类比学习、示例学习、概念聚类。

机器学习中的方法或范式有很多种分类体系,例如,按学习策略分类、按知识表示方法分类、按应用领域分类和按对人类学习的模拟方式分类等。如果按照学习策略分类,即按照学习中所使用的推理方法,机器学习可分为:记忆学习、传授学习、演绎学习、类比学习、示例学习、发现学习等;如果按获取知识的表示方法分类,机器学习可分为:代数表达式参数、决策树、形式文法、产生式规则、形式逻辑表达式、图和网络、框架和模式、计算机程序和它的过程编码、神经网络、多种表示形式的组合;如果按应用领域分类,机器学习可分为:专家系统学习、机器人学习、自然语言理解学习等;如果按对人类学习的模拟方式分类,机器学习可分为:从功能上模拟的符号学习、从结构上模拟的连接学习。综合分类:经验性归纳学

习、分析学习、类比学习、遗传算法、联接学习、强化学习等等。

纵观机器学习的种种分类体系不难发现,无论是基于联接机制的机器学习研究,还是基于符号机制的机器学习研究,其方法和研究手段都带有一定的局限性和片面性,不利于揭示人脑的思维机制和学习机理,它们都只从单一的层次来描述学习过程。对于人脑如何接受外界信息,通过其内部联接机制的变化,从而反映在系统行为的变化这样的问题,现有的理论不能给出一个大家都能接受的、合理的解释<sup>[1]</sup>。

事实上,学习是受一定的意志支配的(即有特定的学习目的),其内部表现为一定的结构(即基于联接机制),其外部表现为一定的行为变化(即基于符号学习)的复杂过程。它涉及到联接理论、认知科学、行为科学、神经科学等多门科学。因此,对于机器学习的研究,必须采用计算机科学、控制论、人工智能、认知科学、神经科学、心理学等多学科交叉的方法,才有望取得机器学习研究的更大进展<sup>[2]</sup>。

我们根据认知科学的理论,同时,利用计算机科学、数学及工程科学的方法,提出了一种自主式机器学习方法(自主学习, Autonomic Learning, 简记为 AL)<sup>[7]</sup>。该方法尝试从微观层次、中间层次、宏观层次,根据人脑的学习机理模拟人脑的学习过程和学习方法。

在不确定性条件下的数据自主式机器学习方法,或者称为自主式学习方法,是人工智能知识获取研究中的一个难点问题。实现自主式机器学习,摆脱知识获取过程中对领域专家先验知识依赖的约束(当然,不排斥适当利用领域专家拥有的先验知识),是实现智能知识获取的一个具有挑战性的研究

<sup>\*</sup>基金项目:国家自然科学基金项目(60775045),江苏省自然基金项目(BK2005027),苏州大学 211 项目(2118005)。王 静 硕士研究生,主要研究方向为自主学习、动态模糊逻辑等;李凡长 教授,硕士生导师,主要研究方向为机器学习、动态模糊逻辑等。

课题<sup>[3]</sup>。

自主学习(Autonomic Learning, AL)是指由学习者态度、能力和学习策略综合而成的一种自导学习的内在机制。例如学习个体指导和控制自己的学习能力,制定学习目标的能力,针对不同学习任务选择不同的学习方法和学习活动的的能力,对学习过程进行监控的能力,对学习结果进行评价的能力,当然也有根据情况向他人求助的能力。自主学习通常指主动、自觉、独立的学习,它与被动、机械、接受式的学习相比,是个体终身学习和毕生发展的基础。自主学习历来是教育和心理学家共同关注的一个重要问题,也是当前机器学习领域研究中的一个热点问题<sup>[4]</sup>。

现实世界中的各种问题都可以用学习空间来描述,即将学习问题映射到学习空间中,一个学习问题能映射到多个学习空间中,而一个学习空间也可用于描述多个学习问题<sup>[5]</sup>。我们按照学习个体的自主程度,将学习空间进一步分为自主学习空间和非自主学习空间。由于在实际的学习情境中,完全自主的和完全不自主的学习都较少,多数学习介于这两极之间(Zimmerman,1994)。因此,与其把学习空间截然地划分为自主的或不自主的,不如说学习的自主程度的大小(强弱),并进而分清自主学习个体在学习的哪些方面是自主的,在哪些方面是不自主的,这样更有利于对自主学习个体的学习有针对性地施加影响,从而更好地完成学习任务<sup>[4]</sup>。

总之,自主学习系统是具有动态模糊性的系统。目前,人们用于研究动态性和模糊性的理论工具主要有:模态逻辑、时序逻辑、进程代数、action 理论、对策论、模糊逻辑、Z-语言、情景演算、Reinforcement Learning 及 Rough Set 等。在这些理论工具中,如时序逻辑、进程代数、action 理论、对策论、情景演算及 Reinforcement Learning 等可以描述动态问题,但对模糊问题的处理是不足的;而模糊逻辑、Rough Set 等这些理论可以解决模糊问题,但对解决动态问题却又是它们的弱点。而自主学习既具有动态性又具有模糊性。上述这些理论工具在解决动态模糊性问题方面是不能令人满意的。

因此,对于自主学习来说,选择一个可以解决动态模糊性问题的理论工具是必要的。自主学习公理体系研究又是自主学习理论及应用研究的基础。基于这些事实,本文选择动态模糊逻辑(Dynamic Fuzzy Logic, DFL)<sup>[6]</sup>来研究自主学习,在文献<sup>[7]</sup>的基础上研究自主学习的公理体系,从而提出了“一种基于 DFL 的自主学习子空间学习的公理体系”。

本文的内容组织如下:首先给出自主学习子空间学习的公理体系,以便形式化地描述自主学习系统,并介绍公式变形规则和动作选择策略,然后对公理进行合理性解释。

## 2 自主学习子空间学习的公理体系

公理体系由两部分组成:(1)一组初始公式,即公理;(2)若干条变形规则,即推演规则。

首先,我们在 DFL 的基础上,给出自主学习的相关定义如下:

**定义 1** (自主学习空间(Autonomic Learning Space) 由一切自主学习要素构成的用于描述学习过程的空间称为自主学习空间。它由{自主学习样例,自主学习算法,输入数据,输出数据,表示理论}5 个元素组成,可表示为 $(\vec{S}, \vec{S}) = \{(\vec{E}x, \vec{E}x), ER, (\vec{X}, \vec{X}), (\vec{Y}, \vec{Y}), ET\}$ 。

**定义 2** 自主学习子空间(Autonomic Learning Sub-

space):我们将自主学习空间 $(\vec{S}, \vec{S})$ 划分为  $n$  个子空间, $(\vec{S}, \vec{S}) = \{(\vec{S}_1, \vec{S}_1), (\vec{S}_2, \vec{S}_2), \dots, (\vec{S}_n, \vec{S}_n)\}$ 。

该划分是完备的,可得公理如下:

**公理 1**  $(\vec{S}, \vec{S}) = \bigcup_{i=1, \dots, n} (\vec{S}_i, \vec{S}_i)$ 。

由文献<sup>[6]</sup>可知,自主学习子空间之间存在动态模糊关系 $(\vec{R}, \vec{R})$ 。

**公理 2**  $(\vec{R}, \vec{R}) \in DF((\vec{S}_i, \vec{S}_i) \times (\vec{S}_j, \vec{S}_j)), i, j = 1, \dots, n$

所述二元动态模糊关系可以推广为  $n$  元动态模糊关系。

下面给出自主学习中动态模糊关系的相关性质:

**公理 3** 设 $(\vec{R}, \vec{R}), (\vec{S}, \vec{S}), (\vec{P}, \vec{P}), (\vec{R}, \vec{R})_{(\tau, \tau)} \in DF((\vec{S}_i, \vec{S}_i) \times (\vec{S}_j, \vec{S}_j)), (\vec{t}, \vec{t}) \in (\vec{T}, \vec{T})$ , 则有以下性质:

$$(1) (\vec{R}, \vec{R}) \vee (\vec{R}, \vec{R}) = (\vec{R}, \vec{R}), (\vec{R}, \vec{R}) \wedge (\vec{R}, \vec{R}) = (\vec{R}, \vec{R})$$

$$(2) (\vec{R}, \vec{R}) \vee (\vec{S}, \vec{S}) = (\vec{S}, \vec{S}) \vee (\vec{R}, \vec{R}), (\vec{R}, \vec{R}) \wedge (\vec{S}, \vec{S}) = (\vec{S}, \vec{S}) \wedge (\vec{R}, \vec{R})$$

$$(3) ((\vec{R}, \vec{R}) \vee (\vec{S}, \vec{S})) \vee (\vec{P}, \vec{P}) = (\vec{R}, \vec{R}) \vee ((\vec{S}, \vec{S}) \vee (\vec{P}, \vec{P}))$$

$$((\vec{R}, \vec{R}) \wedge (\vec{S}, \vec{S})) \wedge (\vec{P}, \vec{P}) = (\vec{R}, \vec{R}) \wedge ((\vec{S}, \vec{S}) \wedge (\vec{P}, \vec{P}))$$

$$(4) ((\vec{R}, \vec{R}) \vee (\vec{S}, \vec{S})) \wedge (\vec{R}, \vec{R}) = (\vec{R}, \vec{R})$$

$$((\vec{R}, \vec{R}) \wedge (\vec{S}, \vec{S})) \vee (\vec{R}, \vec{R}) = (\vec{R}, \vec{R})$$

$$(5) (\vec{R}, \vec{R}) \wedge \left( \bigvee_{(\tau, \tau) \in (\vec{T}, \vec{T})} (\vec{R}, \vec{R})_{(\tau, \tau)} \right) = \bigvee_{(\tau, \tau) \in (\vec{T}, \vec{T})} (\vec{R}, \vec{R})_{(\tau, \tau)}$$

$$((\vec{R}, \vec{R}) \wedge (\vec{R}, \vec{R})_{(\tau, \tau)}) \vee \left( \bigwedge_{(\tau, \tau) \in (\vec{T}, \vec{T})} (\vec{R}, \vec{R})_{(\tau, \tau)} \right) = \bigwedge_{(\tau, \tau) \in (\vec{T}, \vec{T})} ((\vec{R}, \vec{R}) \vee (\vec{R}, \vec{R})_{(\tau, \tau)})$$

$$(6) (\vec{R}, \vec{R}) \vee (\vec{I}, \vec{I}) = (\vec{I}, \vec{I}), (\vec{R}, \vec{R}) \wedge (\vec{I}, \vec{I}) = (\vec{R}, \vec{R})$$

$$(\vec{R}, \vec{R}) \vee (\vec{O}, \vec{O}) = (\vec{R}, \vec{R}), (\vec{R}, \vec{R}) \wedge (\vec{O}, \vec{O}) = (\vec{O}, \vec{O})$$

其中 $(\vec{I}, \vec{I}), (\vec{O}, \vec{O}) \in DF((\vec{S}_i, \vec{S}_i) \times (\vec{S}_j, \vec{S}_j)), (\vec{I}, \vec{I})((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) = (\vec{I}, \vec{I}), (\vec{O}, \vec{O})((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) = (\vec{O}, \vec{O})$   
 $(\forall ((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) \in ((\vec{S}_i, \vec{S}_i), (\vec{S}_j, \vec{S}_j)))$  分别称为 $(\vec{S}_i, \vec{S}_i)$ 到 $(\vec{S}_j, \vec{S}_j)$ 的动态模糊全称关系和动态模糊零关系。

**定义 3** 若 $(\vec{R}_1, \vec{R}_1), (\vec{R}_2, \vec{R}_2) \in DF((\vec{S}_i, \vec{S}_i), (\vec{S}_j, \vec{S}_j)), \forall ((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) \in ((\vec{S}_i, \vec{S}_i), (\vec{S}_j, \vec{S}_j))$ , 记:

$$(1) (\vec{R}_1, \vec{R}_1) \subset (\vec{R}_2, \vec{R}_2) \Leftrightarrow (\vec{R}_1, \vec{R}_1)((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) \leq (\vec{R}_2, \vec{R}_2)((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j))$$

$$(2) ((\vec{R}_1, \vec{R}_1) \cup (\vec{R}_2, \vec{R}_2))((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) = (\vec{R}_1, \vec{R}_1)((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) \vee (\vec{R}_2, \vec{R}_2)((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j))$$

$$((\vec{R}_1, \vec{R}_1) \cap (\vec{R}_2, \vec{R}_2))((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) = (\vec{R}_1, \vec{R}_1)((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) \wedge (\vec{R}_2, \vec{R}_2)((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j))$$

$$(3) (\vec{R}, \vec{R})^{-1}((\vec{s}_i, \vec{s}_i), (\vec{s}_j, \vec{s}_j)) = (\vec{R}, \vec{R})((\vec{s}_j, \vec{s}_j), (\vec{s}_i, \vec{s}_i))$$

**公理 4**  $(\vec{R}, \vec{R})^{-1}$ 有以下性质:

$$(1) (\vec{R}_1, \vec{R}_1) \subset (\vec{R}_2, \vec{R}_2) \Rightarrow (\vec{R}_1, \vec{R}_1)^{-1} \subset (\vec{R}_2, \vec{R}_2)^{-1}$$

$$(2) ((\vec{R}, \vec{R})^{-1})^{-1} = (\vec{R}, \vec{R})$$

$$(3) ((\vec{R}_1, \vec{R}_1) \cup (\vec{R}_2, \vec{R}_2))^{-1} = (\vec{R}_1, \vec{R}_1)^{-1} \cup (\vec{R}_2, \vec{R}_2)^{-1}$$

$$(4) ((\vec{R}_1, \vec{R}_1) \cap (\vec{R}_2, \vec{R}_2))^{-1} = (\vec{R}_1, \vec{R}_1)^{-1} \cap (\vec{R}_2, \vec{R}_2)^{-1}$$

现在对自主学习系统采用离散化处理, $(\vec{S}, \vec{S}), (\vec{L}, \vec{L})$ 均用状态子空间表示,于是有:

**定义 4** 自主学习模型 ALM 描述为:

$$(\vec{x}, \vec{x})(k+1) = G1((\vec{x}, \vec{x})(k), (\vec{u}, \vec{u})(k), \xi_1(k))$$

(2.2.1)

(下转第 170 页)

其中的一部分。图中所有的矩形对象表示粒网络建立过程中产生的粒。

### 3.3.5 分类规则的表达

一个分类规则可以用(,)、=、\* (合取)、-(非)、+(析取)等符号表示。每一个文本类都有一个或多个分类规则。例如:

Class1=f3 \* (-f8)

Class2=f2 \* (-f6)

Class2=f8

Class2=(f2 \* (-f6)) + f8

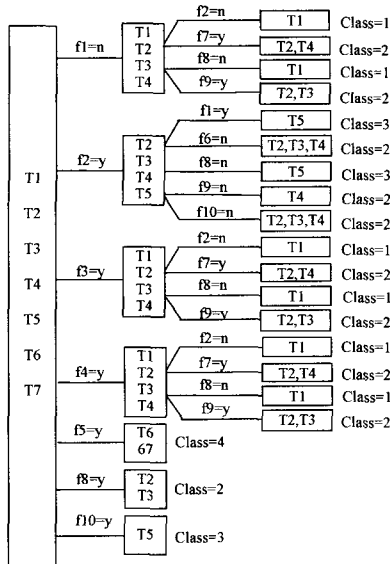


图1 部分粒网络

**结束语** 通过实验,可以得到如下结论:

(1)当类之间的重叠现象比较严重时,如本实验中的class1与class2,如果按基于统计的方法,通过比较训练文本中不同的模板决定文本的类别,则class2中的文本很大比例就会分到class1中,造成分类精度的下降。而本算法表现出较高的精确率。

(2)目前,对论域粒化产生规则的方法,如id3,c4.5决策树,都只是划分的方法,一个对象只能得到一个规则。而本算法是用覆盖的方法产生规则,一个对象可以得到多于一个的规则,适用于语料文档不充裕时的文本分类。而且,由算法的特性决定,覆盖的方法会产生较短的规则。

(3)文本分类中使用此算法,属性值简化为两项,所以运算步骤被大大缩减,便于实践。

### 参考文献

- [1] Sebastiani F. Machine learning in automated text categorization. ACM Computing Surveys, 2002, 34(1): 1-47
- [2] Yao Jingtao, YAO Yiyu. A granular computing approach to machine learning // Proc. of the 1st International Conference on Fuzzy Systems and Knowledge Discovery. Singapore, 2002; 732-736
- [3] Yao Yiyu. On modeling data mining with granular computing // Proc. of the 25th Annual International Computer Software and Applications Conference. Chicago, 2001; 638-643
- [4] Yao Yiyu. Granular computing: Basic issues and possible solutions // Proc. of the 5th Joint Conference on Information Sciences. Atlantic, 2000; 186-189
- [5] Zhao Yan, Yao Yiyu. Interactive classification using a granule network // Proc. of the Fourth IEEE International Conference on Cognitive Informatics. Washington, 2005; 250-259

(上接第 147 页)

$$(\hat{y}, \hat{y})(k) = G2((\hat{x}, \hat{x})(k), (\hat{u}, \hat{u})(k), \xi_2(k)) \quad (2.2.2)$$

$$(\hat{p}, \hat{p})(k) = \sum_{i=1}^k P((\hat{y}, \hat{y})(i)) \quad (2.2.3)$$

其中 $(\hat{x}, \hat{x})(k)$ 是 $(\hat{S}, \hat{S})$ 在时刻 $k$ 的状态变量, $(\hat{u}, \hat{u})(k)$ 是 $(\hat{S}, \hat{S})$ 的动态输出, $\xi_2(k)$ 是状态方程中的随机干扰, $(\hat{y}, \hat{y})(k)$ 是 $(\hat{L}, \hat{L})$ 的动态反馈输出, $\xi_1(k)$ 是观察随机误差。 $k$ 表示时刻,只取整数值。假定其中的向量全部为有限维状态变量。 $(\hat{p}, \hat{p})$ 表示系统自主学习性能指标, $P(i)$ 是一个标量函数,表示时刻 $i$ 时的系统自主学习性能。

根据定义可得如下 3 个命题:

**命题 1** ALS 是随机系统(Random System)。

**命题 2** ALS 是开放系统(Open System)。

**命题 3** ALS 是非线性系统(Nonlinear System)。

证明略。

**命题 4** 现实世界中的各种问题都可以用学习子空间来描述,并且学习子空间包含所有可能的学习路径。若问题可解,则我们必然能从中找到一条学习路径,且为最优(近似最优)学习路径。

**证明:**将学习问题映射到对应的学习子空间中,从问题初始状态出发,在学习过程中,做每一步路径划分时,描述出当前状态的所有可能的后裔状态,并根据具体的衡量标准从中找出最佳状态,然后做进一步的划分,直至找到目标状态,这就保证了我们能够找到路径,并且为所需的目标路径。需要注意的是在具体的学习算法中由于考虑到计算时间和存储空

间的问题,在大规模的学习问题中一般不会扩展出所有的状态,但是无论是全局最优还是局部最优,根据学习问题的提示信息一定能找到最优(近似最优)的学习路径。

**结束语** 自主学习公理体系研究是自主学习理论及应用研究的基础。本文利用动态模糊关系,建立了自主学习子空间的公理体系及相关内容。该公理体系由几条相互独立的非常简洁的表达式构成,进一步深化了动态模糊逻辑和自主学习理论。

### 参考文献

- [1] 闫友彪,陈元琰. 机器学习的主要策略综述. 计算机应用研究, 2004; 4-13
- [2] 王继成. 基于认知模拟的自适应机器学习算法研究. 软件学报, 2001, 12(8): 1205-1211
- [3] 王国胤. 基于 Rough 集理论的自主式机器学习. 计算机科学, 2004, 31(10): 18-20
- [4] 庞维国. 自主学习. 上海: 华东师范大学出版社, 2003
- [5] 陈凤, 李凡长. 李群机器学习(LML)的学习子空间轨道生成理论及算法初探. 苏州大学学报: 自然科学版, 2007, 23(1): 61-66
- [6] Li Fan-zhang. Dynamic Fuzzy Logic and Its Applications. Nova Science Publishers, Sep. 2006
- [7] Wang Jing, Li Fan-zhang. Autonomic Learning Model and Algorithm Based on DFL // 2007 IEEE International Conference on Granular Computing. Silicon Valley, U. S. A
- [8] 刘贵龙. 模糊近似空间上的粗糙模糊集的公理系统. 计算机学报, 2004, 27(9): 1187-1191
- [9] 刘贵龙. Pawlak 粗糙集的公理系统. 系统工程与电子技术, 2006, 28(11): 1752-1755