

成绩影响因素分析模糊专家系统的设计与实现^①

Design and Implementation of Fuzzy Expert System in Analysis of Grade-Affected Factors

潘道华 董春游 (黑龙江科技学院 计算机与信息工程学院 黑龙江 哈尔滨 150027)

摘要: 针对大学生学习成绩不稳定及影响因素多而杂的现实情况,提出了一种用于确定众多影响因素中主要因素的方法。利用专家系统与网络数据库结合的结构模型,采用正向与逆向相结合的推理策略进行推理,建立模糊推理机制,阐述推理机的实现过程,并开发了相应模糊专家系统。通过实测,说明了该模糊专家系统在分析影响大学生学习成绩的因素中的必要性和价值。

关键词: 学习成绩 模糊专家系统 模糊规则 模糊推理 不确定性推理

1 引言

从学生跨入小学门口的一刹那开始,学习成绩就是每个学生、家庭及老师所关心和牵挂的热门话题。在素质教育的今天,高校里大学生学习成绩依然是衡量学生的一个重要标准。由于大学生一般都是外地求学,所以大学生的学习成绩受父母的教养方式的影响逐步减弱,但是随着大学生自身身心的变化、学习环境的改变以及时代的发展而呈现出更多的复杂性和不可预测性,学生的学习成绩具有极大的不稳定性,容易受到学生自身心理因素和外部环境因素的影响。教师们根据多年的教学经验以及长期与学生共处的体会,将影响大学生学习成绩的各方面因素按其属性范围划分逻辑层,分析成如下图所示的树状图。

在此树形结构中,最高层的外部环境因素和内部个人因素称为根节点,中间一层上的八个节点称之为中间节点,最下面一层上各个直接原因节点称之为叶子节点。通常,学生的学习成绩升降是由多方面原因引起的,各原因之间又有可能互相关联,而且各原因对成绩的影响作用也不相同,但对学生个体而言,总有一个原因是导致成绩浮动的主要原因,所以找出主要原因,处理掉主要矛盾,某些次要原因也会消除,这对提高大学生学习成绩至关重要。本文设计的模糊专家系统则可以有效地从众多影响学生成绩的因素中找到主要因素。

2 模糊专家系统简介

专家系统是一个具有大量的专门知识与经验的程序系统,它应用人工智能技术和计算机技术,根据某领域一个或多个专家提供的知识和经验,进行推理和判断,模拟人类专家的决策过程,以便解决那些需要人类专家处理的复杂问题^[1]。经典专家系统是一个二值逻辑系统,它只能处理确定性(即只能在0或1中取值)的数据和命题。而实际上,人类专家的很多知识都是模糊的、不确定的,这使得经典专家系统在处理这些问题时效果不理想,难以实现预定的功能,在此情况下模糊专家系统就得以产生和发展起来。模糊专家系统是建立在模糊集合等理论的基础上,采用模

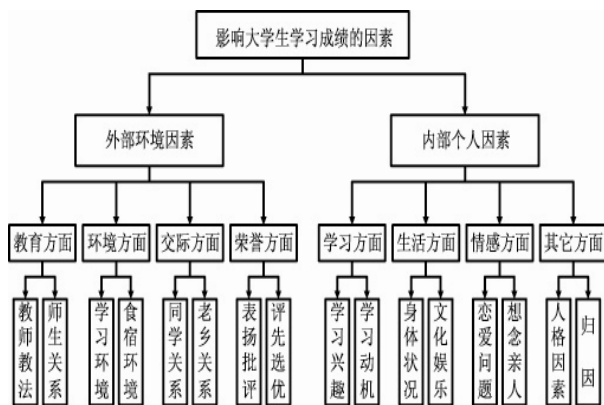


图1 影响大学生学习成绩的因素

① 收稿时间:2008-08-02

糊技术来处理不确定性的一类专家系统。模糊专家系统一般由人机接口模块、模糊数据库、模糊知识库、模糊推理机、知识获取模块和解释模块等部分构成，其结构如图 2 所示^[2]：

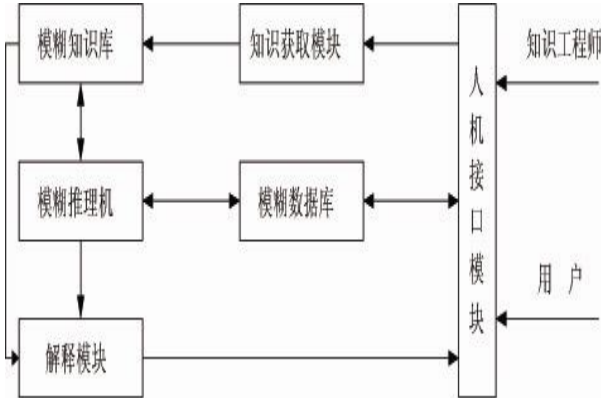


图 2 模糊专家系统的一般结构

人机接口模块：实现用户和专家系统之间传递信息的交互接口。用于输入系统的初始信息，这些信息允许是模糊的、随机的、或不完备的；也用于输出系统的最终结论，这些结论也允许是带不确定性的；显示系统推理过程的解释及系统运行过程中所有的人—机对话；打印必要的输出信息等。

模糊数据库：用来存储问题的原始特征信息和系统推理过程中产生的中间信息及推理最终结论等。它与一般的专家系统中的综合数据库相类似，只不过存储的信息都可能是不确定的。

模糊知识库：存放由领域专家总结出来的与特定问题求解领域相关的事实和规则，包括事实性知识和启发性知识两种。与一般的知识库所不同的是这些事实与规则可以是模糊的或不完全可靠的。因此在模糊专家系统中通常要求为各事实附上可信度标志并为各规则附上强度标志。

知识获取模块：接受领域专家以自然语言形式描述的领域知识，将之转换成标准的规则和事实表达形式，存入模糊知识库。

解释模块：用于回答用户提出的问题，解释推理的结论。

模糊推理机：是模糊专家系统的核心。它可根据系统输入的初始不确定性信息，利用模糊知识库中的

不确定性知识，按一定的模糊推理策略，较理想地处理待解决的问题，给出恰当的结论，从而实现基于知识的推理求解过程。

3 专家系统与网络数据库结合的结构模型

专家系统可以处理非结构化决策问题，可进行知识表达，运用启发式搜索和推理处理某些专家水平的困难问题，同时具有知识获取和解释功能。但专家系统缺乏有效地搜索和利用大型知识库的能力，知识库管理能力比较差，工作效率较低，这些问题阻碍了专家系统的发展和应用。而目前应用最广泛的网络数据库管理系统用于解决结构化决策问题，在数据共享、并发控制、数据存取与查询方面功能很强，且具有低冗余的特点，但缺乏数据抽象能力，不能处理基于规则的知识，没有推理演绎能力缺乏自然语言理解能力。为摆脱单方面进行研究的局限性，专家系统和网络数据库技术相结合，彼此取长补短，相互促进，共同发展。专家系统与数据库结合后，专家系统的结构模型如图 3 所示。为提高专家系统的性能，将正、逆向推理规则分开组织、分开调用，并用数据库来构造广义的知识库^[3]。

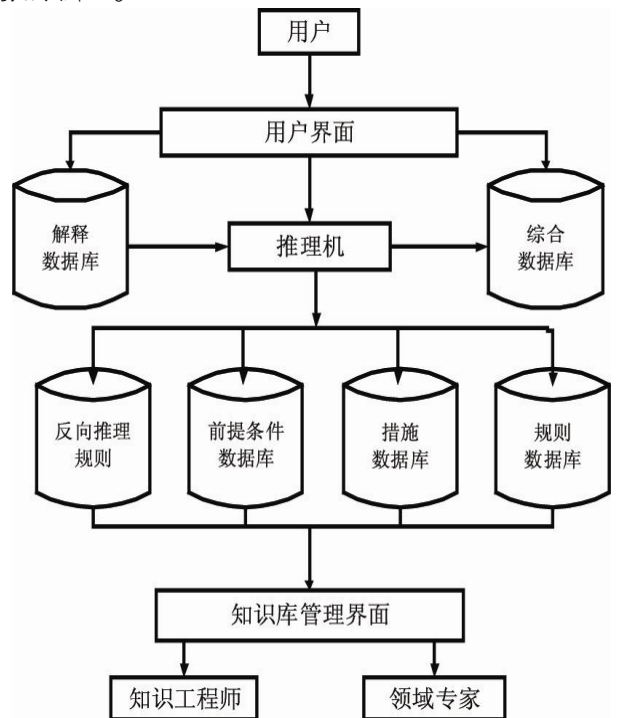


图 3 专家系统的结构模型

4 正向推理的规则表示

正向推理是从已知的事实出发, 向结论方向进行推导的推理方式。它根据知识库中已有的知识, 正向使用规则进行推理, 将采集到的特征信息与知识库中的模糊规则前件进行匹配来选用可用规则, 若匹配成功则将该规则结论作为中间结果继续匹配, 直到问题解决。若推理过程中有多条模糊规则匹配可用, 则按照冲突消解策略从中选出一条最优规则执行^[4]。

正向推理用多输入、单输出的模糊产生式规则表示。模糊产生式规则具有可以有效表达启发式知识等优点, 并可以根据数据可靠性给出可信度因子, 从而实现模糊推理^[5]。

模糊产生式规则的形式如下:

IF $E_1(w_1, cf_1)$ and $E_2(w_2, cf_2)$ and...and $E_n(w_n, cf_n)$ THEN $H(cf_h)$ (CF, λ) (1)

式中, E_1, E_2, \dots, E_n 为规则的模糊前提条件(前件); w_1, \dots, w_n 为用户提供的每个事实(规则的前提条件)的证据重要程度, 即权重; cf_1, \dots, cf_n 为规则的各个前提条件的可信度; H 为规则的结论(规则后件), cf_h 为结论的可信度; CF 为该规则的置信度或事件发生的可能性程度($0 < CF \leq 1$), 若多个前提条件推导出一个结论, 且每个前提条件都对结论起到一定的影响, 即学生的学习成绩浮动一般是受多方面因素叠加影响的, 那么该条规则条件的置信度为多个前提条件权重的最大值并根据条件数多少按公式

$$CF = \max(w(i))(m-1)/m \quad (i=1, 2, \dots, n) \quad (2)$$

取值; λ ($0 < \lambda < 1$) 是规则的阈值, 即规则可被使用的限值, 只有规则的置信度大于阈值时, 该规则才被激活, 系统相信的假设成立, 规则方可被使用, 结论才可以提交给用户, 否则, 系统相信的假设不成立, 结论不提交给用户。

推理方法: $cf_h = cf_1 * cf_2 * \dots * cf_n * CF$ (3)

利用这个公式可以把推理过程的每一步引入的不确定性全部组合起来, 产生最后结论的总可信度。

5 逆向推理的算法模型

因素集合 $A=(a_1, a_2, \dots, a_n)$, $a_i(i=1, 2, \dots, n)$ 为学生学习成绩浮动的各影响因素, 对应树状图中树的各逻辑层; 模糊向量集 $B=(b_1, b_2, \dots, b_m)$, $b_i(i=1, 2, \dots, m)$ 为各影响因素的权值, 由于在某个学习阶段或某个环境下, 每个因素对某位学生的影响程度各不相同, 可以对每个影响因素赋以不同的权值。

对单个影响因素 a_i 的评判, 得到 B 上的模糊集 $(r_{i1}, r_{i2}, \dots, r_{im})$, 所以它是从 A 到 B 的一个模糊映射。

$$f: A \rightarrow \varphi(B) \quad (4)$$

模糊映射可以确定一个模糊关系 $R \in a_{n \times m}$,

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{bmatrix} \quad (5)$$

它是由所有对单因素评判的模糊集组成的, 称为评判矩阵。

设已给定的影响大学生学习成绩的因素可信度模糊集 $C = \{c_1, c_2, \dots, c_n\}$, 则定义模糊合成运算为,

$$C \odot R = E = \{e_1, e_2, \dots, e_n\} \quad (6)$$

其中, $e_j = \vee_i (c_i \wedge r_{ij})$ ($j=1, 2, \dots, n$) 或者 $e_i = (c_1 * r_{1j})(c_2 * r_{2j}) \dots (c_n * r_{nj})$ ($i=1, 2, \dots, n, j=1, 2, \dots, m$)。这是完成该系统反向推理算法的依据。

6 模糊推理机的实现过程

模糊专家系统的推理实际是不确定性推理, 它根据接收到的模糊输入信息(加权值)和规则库中现有的模糊规则(规则可信度), 产生相应的模糊输出变量(结论可信度)。即推理过程中信息不确定性(有数据置信度、推理可信度、结论可信度等)的传播。本系统采用正向与逆向相结合的双向推理策略。

①正向推理: 根据大学生表现出的学习问题和消息以及搜索证据的置信度, 运用正向推理模糊产生式规则得出满足规则条件的可能影响该学生学习的因素, 作为初步结论。

②逆向推理: 首先通过学生的个人汇报、座谈会、同学反映以及老师多年的教学经验, 对教师分析出的影响学生学习成绩的 16 项因素, 给出某位大学生在当前阶段影响其学习的各因素的程度值(权值); 将正

向推理得出的 16 个前提因素可信度与对应的因素的权值做乘积,对隶属于同一中间节点因素的乘积取两者较大的为中间节点因素可信度值。并把对应前提因素保存到推理历史库表中;然后将刚得到的 8 个中间节点因素可信度值与对应中间节点因素的权值做乘积,对隶属于同一根节点因素的 4 者中取最大的为根节点因素的可信度值。同时将对应的中间节点因素保存到推理历史库表中;对两个根节点因素的可信度值进行比较,选取可信度较大的根节点因素作为影响该学生学习的主要因素。同时将对应的根节点因素保存到推理历史库表中^[3]。

运用推理历史库表中的记录,由根节点因素经过中间节点因素向回寻找,直到找到作为叶子节点的前提因素,确定影响该大学生学习浮动的主要原因。再结合措施库表,找出合适的解决方法与措施。

7 结束语

本文概述了模糊专家系统的基本构成,详细叙述了推理机的正向模糊产生式知识表示、逆向推理算法模型以及模糊推理机的实现过程,模糊产生式较好的解决了模糊系统中的不确定性知识表示的问题,本系统利用专家系统与网络数据库结合的办法,采用正向

与逆向相结合的推理策略,达到推理全面可靠的目的。该学生学习成绩影响因素分析模糊专家系统采用 JSP^[6]语言开发完成,以计算机与信息工程学院 06 级 5 班的学生学习成绩情况为例进行实际测试,结果表明系统运行稳定、可靠,该系统易于管理、易于修改,可以满足用户要求,具有极大的应用及推广价值。

参考文献

- 1 蔡自兴,徐光佑.人工智能及其应用(第三版).北京:清华大学出版社,2004:200.
- 2 程伟良.广义专家系统.北京:北京理工大学出版社,2005.
- 3 岳一领,王希术,李东生.新兵经常性思想工作模糊专家系统推理机的实现.太原理工大学学报,2003,36(5).
- 4 Xu KL, Zhou M, Ren JW. An object-oriented power system fault diagnosis expert system. Proceedings of International Conference on Electrical Engineering. Sunnyvale: Omni Press, 1999:166-117.
- 5 Chen SM. Fuzzy group decision making for evaluating the rate of aggregative risk software development. Fuzzy Set and Systems, 2001,118(1):75-88.
- 6 Chopra V, Eavas J, Jones R. JSP 程序设计.北京:人民邮电出版社,2006.