

动物营养性疾病诊断专家系统产生式模型研究^{*}

杨毅¹, 张曦², 吴兴勇¹, 陶琳丽², 曹志勇¹

(1 云南农业大学 计算机科学系; 2 动物营养重点实验室, 云南 昆明 650201)

[摘 要] 以经典产生式系统的观点研究了动物营养性疾病诊断专家系统模型, 提出了基于模糊诊断表的交互式状态空间搜索方法, 并应用模糊推理及模式识别算法, 实现了专家知识的组织与动物营养性疾病诊断。

[关键词] 专家系统; 营养性疾病; 产生式系统; 模糊诊断

[中图分类号] TP182; S818.9

[文献标识码] A

[文章编号] 1671-9387(2003)04-0080-03

产生式系统(Production System)是人工智能中最典型最普遍的一种结构^[1], 由于它非常符合专家的问诊过程, 所以许多诊断专家系统以此为基础构建。经典的产生式系统将知识以规则的形式存入计算机中, 用符号推理的方式进行诊断。这些基于规则的专家系统存在的最主要问题是: 规则库的组合爆炸与无效匹配对搜索效率的影响^[2,3], 以及非完整和不精确的数据给诊断结果带来的误差^[4]。

以经典产生式系统的观点, 即综合数据库、规则及控制策略, 应用模糊数学方法, 提出了基于模糊诊断表的推理策略, 可有效缩小搜索范围。对于规则知识, 采用模糊目标特征向量集的表达式(在文献[5]中, 提出了一种类特征向量知识组织方法, 但其元素含义与本系统完全不同), 具有案例学习与知识学习的能力, 且推理及评价也采用模糊系统方法, 能够在一定程度上解决不精确、不完整数据及诊断中观察数据噪声^[4]的影响。

由于动物疾病与人体疾病诊断在主体主诉特性上有着巨大差别——动物临床症状全部靠第3者的观察得到, 而且动物诊断系统的用户一般专业水平较低, 甚至可以不具有任何专业知识背景, 所以动物疾病诊断专家系统与人体疾病诊断专家系统相比, 更注重对症状空间的有效提问及搜索, 其数据模糊特性也更加明显。目前, 对动物疾病诊断专家系统的报道尚不多, 本研究是针对动物疾病诊断系统特点的一种新的尝试。

1 结构设计

整理收集了动物营养性疾病 33 种共 281 个症

状。疾病的症状向量空间设为 $(A_1, A_2, \dots, A_n)$, 其中 A_i 是模糊集合。对 1 次诊断, 得到 1 个实例空间 $(x_1, x_2, \dots, x_n)$, $x_i \in [0, 1]$ 为症状属于 A_i 的隶属度, 其值通过该症状的明显程度得到。取值如下: 1 十分明显; 0.8 明显; 0.6 比较明显; 0.4 好象; 0.2 好象没有; 0 无此症状; -1 不知道。在疾病诊断中, 要解决的 3 个问题是: 第一, 对初始症状空间如何确定搜索方向, 以对最有可能疾病的关键症状进行提问; 第二, 对每一种疾病如何产生目标特征; 第三, 问诊结束后如何对症状特征向量与目标特征向量的匹配进行评价, 从而给出合理诊断。

1.1 模糊诊断表

设某类疾病 m 种, 共有 n 个症状, $m \times n$ 阵 (P_{ij}) 表示第 i 种疾病出现第 j 种症状的可能性, 取值如下: 1 肯定; 0.8 几乎可以肯定; 0.6 很可能; 0.5 可能; 0.2 很少可能; 0 不可能; 空 未知。

模糊诊断表由行业专家根据专业知识、具体案例和经验进行填写, 综合书本知识进行完善。应用模糊诊断表的规则如下: (1) 如果某次诊断的症状都和某种疾病有关, 则出现这种疾病的可能性比较大。(2) 如果某种疾病出现某种症状的可能性很高, 则出现某种症状时发生这种疾病的可能性就越大。(3) 如果某种疾病不可能出现某种症状, 则出现该症状时这种疾病被排除的可能性就比较大。根据上述规则, 即可以模糊诊断表确定搜索方向, 并对最有可能的疾病关键症状进行提问。

1.2 目标特征的提取

目标特征的提取主要依靠大量的实际诊断案

* [收稿日期] 2003-03-06

[基金项目] 云南省科技攻关项目“云南省家畜饲料高效利用技术及饲料资源优化利用研究”(云计科技[98]253号)

[作者简介] 杨毅(1966-), 男, 云南昆明人, 副教授, 主要从事数据库应用、GIS 集成开发和专家系统研究。

例, 结合专家经验与书本知识获得。对某一种疾病, 可以有多个实例向量。目标特征表示如下:

$$X_k^i = (x_1^{ki}, x_2^{ki}, \dots, x_n^{ki}) \rightarrow P_k^i,$$

式中, X_k^i 为第 k 种疾病的第 i 个实例向量, P_k^i $[0, 1]$ 是第 i 个实例向量诊断中患第 k 种疾病的可能性。作为目标特征, P_k^i 的值通常接近 1。

通过强制性的学习, 随着系统使用时间的增加, 疾病目标特征知识库将趋于完善, 因而系统能够诊断的情况将越来越多, 诊断准确性也越高。

1.3 诊断结果评价

通过交互式的状态空间搜索, 得到某次疾病诊断的一个实例, $Z = (z_1, z_2, \dots, z_n)$, 对目标特征 X_k^i , 定义两个向量的贴进度^[6]:

$$(Z, X_k^i) = \frac{2 \sum_{m=1}^n (x_m^{ki} \wedge z_m)}{\sum_{m=1}^n (x_m^{ki} + z_m)} \quad (x_m^{ki} \geq 0, z_m \geq 0)$$

又 Z 与 X_k^i 的标准化欧氏距离为

$$d_k^i = \frac{\sqrt{\sum_{m=1}^n (x_m^{ki} - z_m)^2}}{\sqrt{\sum_{m=1}^n (x_m^{ki})^2} + \sqrt{\sum_{m=1}^n (z_m)^2}}$$

选择 j , 使 $(Z, X_k^j) + (1 - d_k^j)$ 最大, 则诊断实例与目标特征 X_k^j 匹配程度最好, 诊断患 K 种疾病的置信度是 $P_k^j \wedge (Z, X_k^j)$ 。

1.4 结构示意图

据此, 可以得到诊断系统结构示意图(图 1)。

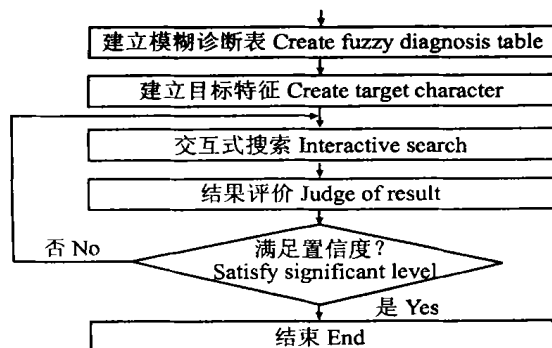


图 1 诊断系统结构示意图

Fig 1 Diagnosis system structure sketch map

2 算法设计

系统综合数据库为 $(x_1, x_2, \dots, x_n)$, 产生式规则在 1.1 中进行了描述。与经典产生式系统不同的是: 对规则的使用与控制策略体现在算法逻辑中, 而不是使规则直接作用于综合数据库。其诊断与控制过

程可描述如下:

根据 (P_{ij}) 确定搜索方向, 采用单层回溯(即每次回溯否定一种疾病)。输入初始症状后, 首先选用规则(1), 若能满足规则(1)的条件, 则得到可信度比较高的结果, 再对结果进行进一步诊断询问, 看是否满足其余症状, 是否和规则(3)相冲突, 此时产生的诊断比较确定。若规则(1)使用失败, 选用规则(2), 对有可能产生初始症状的若干疾病根据可能性大小进行排序搜索, 依照该种疾病可能产生的症状进行提问, 再与目标特征进行匹配, 达不到匹配置信度则进行回溯。

定义如下命题:

R_i 为第 i 种疾病发生; Q_j 为第 j 种症状出现。

算法描述如下:

建立症状状态指标集 J_1, J_2 , 其中 $J_1 = \{j | Q_j > 0\}$, J_1 为阳性和中性症状集, $J_2 = \{j | Q_j = 0\}$, J_2 为阴性症状集;

建立搜索指标集 I ,

$$I = \{i | \bigcup_{j \in J_2} P_{ij} = 0, \bigcup_{j \in J_1} P_{ij} > 0\};$$

设重复诊断标志 $\text{flag} = 0$;

判断是否成立, 对 $\forall j \in J_1, \exists k \in I$, 使 $P_{kj} > 0.6$, 若不是则转 ;

将该症状状态与第 k 种疾病的目标 $X_k^i \rightarrow P_k^i$ 进行匹配, 找到最好的匹配特征向量 X_k^j , 若置信度达到所设, 则给出诊断结果 R_k , 置信度为 $P_k^j \wedge (Z, X_k^j)$, 转结束;

检验重复诊断标志 flag , 若为真, 转 ;

建立询问指标集 $SJ = \{j | P_{kj} \geq 0, j \notin J_1\}$, 对 SJ 中指标进行询问, $\text{flag} = 1$, 构成新的症状状态, 转 ;

显示“你的家畜可能患有第 k 种疾病, 但还不能确诊, 你是否按要求回答了所提问的症状和检查结果”;如果回答是, 则设 $\text{flag} = 0$, 从 I 中删除 k , 若 I 不空, 转 , 否则转结束, 如果回答否, 则询问是否要再重复上一次的诊断, 如果回答是, 则转 , 否则转结束;

$m = 1, 0.8, 0.6, 0.5, 0.2$, 依次在 I 中查询指标 $k \in I$, 使 $j \in J_1$ 且 $P_{kj} = m$ 转。

3 诊断实例

观察到的症状为精神不振(明显), 食欲减退(明显), 行走不稳(十分明显), 喜卧不喜动(十分明显)。则

1) 建立阳性症状指标集 $J_1 = \{j_8: 0.8, j_{12}: 0.8, j_{60}: 1, j_{75}: 1\}$, j_i 表示该症状在症状向量空间的序号。

2) 根据模糊诊断表, 搜索指标集为 $I = (i_1: \text{酮病}, i_5: \text{维生素 B}_2 \text{ 缺乏症}, i_7: \text{维生素 PP 缺乏症}, i_{10}: \text{维生素 B}_{12} \text{ 缺乏症}, i_{13}: \text{白肌病}, i_{17}: \text{铜缺乏症}, i_{22}: \text{佝偻病}, i_{26}: \text{骨营养不良}, i_{27}: \text{衰竭症}, i_{29}: \text{铜中毒})$ 。

3) 其中, 当 $k = 22$ 时, $\forall j \in J_1, P_{kj} > 0.6$, 故患此种疾病(佝偻病)的可能性最大。

4) 在模糊诊断表中按照佝偻病出现症状的可能性(P_{22j})大小进行询问: 四肢僵硬 $\rightarrow$ 关节肿大 $\rightarrow$ 气喘 $\rightarrow$ 异食, 根据回答情况修改症状集 $J_1 = \{j_8: 0.8, j_{25}: 0.6, j_{38}: 0.6, j_{12}: 0.8, j_{59}: 0.8, j_{60}: 1, j_{72}: 1, j_{75}: 1\}$, 佝偻病的目标特征向量之一是 $X_{22}^1 = (j_8: 0.6, j_{25}: 0.8, j_{38}: 1, j_{12}: 0.8, j_{59}: 1, j_{60}: 1, j_{72}: 1, j_{75}: 1) \rightarrow$ 佝偻病(置信度为 0.98)。经过计算, J_1 与 X_{22}^1 贴近度是 0.96, 欧氏距离为 0.11, 在佝偻病的目标特征中

$0.96 + (1 - 0.11) = 1.85$ 为最大, 所以是最好的匹配, 由此可得患佝偻病的可能性是 $(0.96 \wedge 0.98) = 0.96$ 。

4 结束语

动物营养性疾病的产生式诊断模型是一种通用疾病诊断模型, 适合于对某一类疾病的诊断, 可面向完全没有专业知识背景的用户。它以算法逻辑实现了产生式系统的规则选取和搜索控制, 诊断逻辑符合专家思维, 对专家思维进行了较好的模型化表述与实现。是经典产生式系统一种推广意义上的应用。不足之处是系统还缺乏自主学习的功能, 目标特征专家知识库中特征向量间缺少相应的联系, 冲突解决还比较麻烦。对此, 可应用模糊神经网络系统自主学习与自适应功能^[2, 7, 8]加以完善。

[参考文献]

- [1] 林尧瑞, 马少平. 人工智能导论[M]. 北京: 清华大学出版社, 1999.
- [2] 孙佰清, 张长胜, 王雪峰, 等. 人工神经网络在医疗智能诊断系统中的应用[J]. 哈尔滨商业大学学报(自然科学版), 2002, (4): 394-395.
- [3] 姚一波, 王纪亭. 医疗诊断系统专家知识的表达与获取方法[J]. 信息技术, 2002, (2): 31-34.
- [4] 陈真诚, 蒋勇, 胥明玉, 等. 人工智能技术及其在医学诊断中的应用及发展[J]. 生物医学工程杂志, 2002, (3): 505-509.
- [5] 张红梅, 王永诚. 一个仿人疾病诊断专家系统模型[J]. 计算机应用研究, 2002, (1): 41-43.
- [6] 邹开其, 徐扬. 模糊系统与专家系统[M]. 成都: 西南交通大学出版社, 1989.
- [7] 张乃尧, 阎平凡. 神经网络与模糊控制[M]. 北京: 清华大学出版社, 1999.
- [8] 高黎, 卜淮原, 胡曙. 一种医疗智能诊断推理机的设计与实现[J]. 计算机应用与软件, 2002, (6): 44-46.

The research of animal nutritional disease diagnosis expert system based on production system

YANG Yi¹, ZHANG Xi², WU Xing-yong¹, TAO Lin-li², CHAO Zhi-yong¹

(¹ Department of Computer Science, Yunnan Agricultural University, Kunming, Yunnan 650201, China;

² Key Laboratory of Animal Nutrition, Yunnan Agricultural University, Kunming, Yunnan 650201, China)

Abstract: By the point of view of classical production system, this article studies the expert system model of animal nutritional disease diagnosis. Basing on the fuzzy diagnosis form, we put forward with the way of alternation state space search and realize the organization of expert knowledge and disease diagnosis by using fuzzy reasoning and model recognition arithmetic.

Key words: expert system; nutritional disease; production system; fuzzy diagnosis