

基于可信度推理的猪病诊断专家系统

杜少杰 (滨州职业学院计算机系, 山东滨州 256603)

摘要 针对猪病诊断过程中广泛存在的证据不确定性, 建立了能够处理这种不确定性的推理模型。基于证据可信度的推理模型, 给出了证据可信度的表示方法、传递算法。该模型很好地体现了猪病诊断中存在的确定性, 减小了主观因素对诊断结论准确性的影响。

关键词 可信度; 疾病诊断; 专家系统

中图分类号 S858.23 文献标识码 A 文章编号 0517-6611(2008)10-004342-01

An Expert System on Pig Disease Diagnosis Based on Reasoning with Credibility

DU Shaojie (Department of Computer, Binzhou Vocational College, Binzhou, Shandong 256603)

Abstract In order to deal with the uncertainty of evidences mostly in diagnosis process of pig disease, a reasoning model based on evidence credibility was established, and the expressing method and transporting arithmetic of evidence credibility were given. The model could fully reflect the uncertainty in diagnosis of pig disease and reduce the influence of subjective factors on the diagnosing accuracy.

Key words Credibility; Disease diagnosis; Expert system

专家系统是能够模仿某领域的专家应用知识进行推理并解决该领域复杂问题的计算机系统^[1]。在动物医学领域, 疾病诊断是专家系统的主要应用方面。对于基于结论不确定性的推理, 其在实际应用中存在以下问题: 应用的前提是要求条件集合中的各子条件彼此独立; 若一般假设条件的确定程度为1或0, 当用户在疾病表现为非典型或用户主观理解不确定的时候, 人为误差增大, 增大了主观因素对推理结论的影响。为了解决上述问题, 笔者建立了基于证据可信度的推理模型, 并应用该模型开发了猪病诊断专家系统。

1 证据可信度推理模型的建立

1.1 证据可信度推理模型中知识不确定性表示的一般形式
在这种不确定性推理方法中, 知识的表示形式为:

$$\text{IF } E_1(w_1) \text{ AND } E_2(w_2) \dots \text{AND } E_n(w_n) \text{ THEN } H \text{ CF}(H, E), \quad (1)$$

其中, E 为知识的前提条件, 它既可是是一个简单的条件, 也可以是多个简单条件用 AND(和) / OR(或) 连接的复合条件。 H 是结论, 它可以是单一的结论, 也可以是多个结论。 $w_i (i = 1, 2, \dots, n)$ 是加权因子, 其初始值由用户给出。 $\text{CF}(H, E)$ 是该条知识的可信度, 称为可信度因子或规则强度, 而 $\text{CF}(H, E)$ 和 $\text{CF}(E)$ 的具体值是由相关领域专家给出。

1.2 可信度的传递算法^[2] 对于前提条件

$$E = E_1(w_1) \text{ AND } E_2(w_2) \dots \text{AND } E_n(w_n) \quad (2)$$

所对应的组合证据, 其组合证据可信度的计算方法(假设

设 $\sum_{i=1}^n w_i = 1$):

$$\text{CF}(E) = \sum_{i=1}^n (w_i = 1 \times \text{CF}(E_i)), i = 1, 2, 3, \dots, n \quad (3)$$

结论 H 的可信度可通过下式算出:

$$\text{CF}(H) = \text{CF}(H, E) \times \text{CF}(E) \quad (4)$$

当一条规则的结论可信度值 $\text{CF}(H)$ 达到阈值 $\text{CF}(E)$ 时, 即:

$$\text{CF}(H) \geq \text{CF}(E) \quad (5)$$

该规则就可以被应用, 从而推出相应的结论可信度。

1.3 推理策略 推理机的总体推理策略为混合推理。在推理过程中, 系统先进行正向推理, 即由初始症状推理出假设结论; 再进行反向推理, 即通过对假设结论的验证来推理出

最终诊断结论。其具体推理过程见图1。

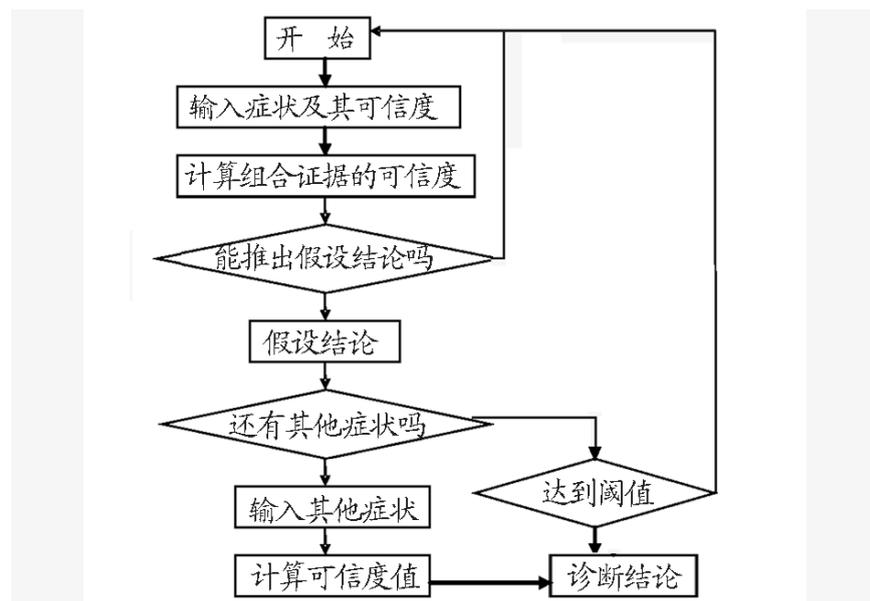


图1 猪病诊断专家系统的推理模式

Fig.1 The sketch map of expert system for pig disease diagnosis

1.4 冲突消解 在应用该系统进行疾病诊断的过程中, 系统要不断用当前已知的事实与知识库中的知识进行匹配, 匹配的结果可能有3种情况: 已知事实不能与知识库中的任何规则成功匹配; 已知事实恰好只与知识库中的一条规则成功匹配; 已知事实可与知识库中的多条规则成功匹配, 或者有多个已知事实可与知识库中的多个知识成功匹配^[3]。

在猪病诊断系统中, 对于以上3种情况, 第1种情况对应的是输入的症状没有找到与之相匹配的疾病, 其原因可能是因为知识库中缺少某些必要的规则, 也可能是因为欲求解的问题超出了系统的能力范围, 此时可根据当时的实际情况进行相应的处理; 对于第2种情况, 由于匹配成功的规则有且只有1个, 所以可以直接应用这条规则进行推理, 不用冲突消解; 对于第3种情况, 由于有多个规则匹配成功, 这时就发生了冲突, 此时需要进行冲突消解。对于猪病诊断系统来说, 就是输入的症状激活了多条规则, 可以推出多种疾病, 这时采用冲突消解的总体原则是选择组合证据可信度最大的规则进行推理^[4]。例如, 假设有如下2条知识:

$$R_1 \text{ IF } E_1(w_1) \text{ AND } E_2(w_2) \text{ THEN } H_1(\text{CF}_1, 1) \quad (6)$$

$$R_2 \text{ IF } E_3(w_3) \text{ AND } E_4(w_4) \text{ AND } E_5(w_5) \text{ THEN } H_2(\text{CF}_2, 2) \quad (7)$$

如果根据当前的已知事实证明这2条规则都可以被激

(下转第4350页)

作者简介 杜少杰(1973-), 女, 河北唐山人, 硕士, 讲师, 高级程序员, 从事XML与Web应用的研究。

收稿日期 2008-02-19

域7)。

将所提取的点、线、区域等复制、粘贴到新建的图层内,从而提取出所需要的图形,另存,就可以建立园林管理系统所需要的特定图层(比如甬路层、灌木层、水域层等)。将所建立的图层进行进一步处理。将需要求算面积的图层通过捕捉工具或拓扑程序,将线段相互连接并形成面域,以求算面积。将图层结构进行维护,建立需要的属性字段(如市、区、地址、面积、长度等),并赋予合适的字符长度,从而建立成层层独立又层层相关的图形文件。由于 MapInfo 的图形和数据同时存在,因而基本的属性数据结构同时建立。

3.2 数据信息编辑的方法 MapInfo 图纸有自身的长度和面积测算工具,对于需要测算的长度和面积可以直接进行计算更新。利用 Mapbasic 语言,可以统计、测算和管理。但是为了避免烦琐的语言学习,通常可以利用常用的 Excel 软件进一步处理。数据库表格在 Excel 里可以简单、直观地进行修改和编辑。MapInfo 可以打开一般的 dbase 数据库文件,所以一般的 Excel 表格另存转化成的 dbase 文件可以被 MapInfo 直接打开,只要 dbase 文件和 MapInfo 图纸的属性字段有相同的检索字段。利用 MapInfo 的更新属性数据列的功能,可以将 dbase 文件数据和 MapInfo 图纸进行自动检索更新,使 MapInfo 数据库的更新相对简单。MapInfo 也可以另存为 dbase 数据库文件,用 Excel 打开进行编辑和统计。

3.3 数据信息的查询和统计 SQL (Structure Query Language) 是 Mapinfo 自带的一套强大的查询系统,语言简单,指令简洁,功能强大,能够完成十分复杂的查询、选择。同时,可以配合 Mapbasic 语言,在处理属性数据时,功能更加完善,而且可以直接进行空间分析。利用 SQL 选择工具不仅可以选择点、线、面,而且可以查出某一字段的特殊数据。MapInfo 又可以把 MapInfo 数据库文件转化为 dbase 文件,输出到 Excel 中进行统计,使数据库的查询和统计变得简单易行。

3.4 打印输出功能 MapInfo 自带的创建专题图纸、打印功能,按照提示的步骤就可以打印所需求的专题图纸。另外,

(上接第4342页)

活,即:

$$CF(E_1(w_1) \text{ AND } E_2(w_2)) > 1 \quad (8)$$

$$CF(E_3(w_3) \text{ AND } E_4(w_4) \text{ AND } E_5(w_5)) > 2 \quad (9)$$

这时就存在了冲突,究竟首先启用哪条规则进行推理就需要进行冲突消解。这里确定的冲突消解策略是首先启用组合证据可信度值最大的规则进行推理。例如:

如果 $CF(E_3(w_3) \text{ AND } E_4(w_4) \text{ AND } E_5(w_5)) > CF(E_1(w_1) \text{ AND } E_2(w_2))$, 则应当首先启用规则 R_2 进行推理。

2 讨论

(1) 基于证据可信度的推理充分考虑了现实客观世界的复杂性以及人们对客观事物认识的不精确性和不完全性,同时减小了主观意识因素对判断和推理过程的影响。当某些不重要的症状没有被提供的情况下,通过引入加权因子实现了在证据不确定性和证据不完全的情况下进行推理。

(2) 在确定加权因子时,一般应考虑以下因素: 相应症

将地图文件输出存为 dbase 文件,在 Excel 进行操作,也可以制作各种专题柱状、折线图和各种报表。

4 软件结合运用的优点

结合运用 AutoCAD 和 MapInfo 2 种软件所形成的文件(图1),具有很多优点: 该文件可直观地看到和打印类似 AutoCAD 的图纸; 利用 MapInfo 中的信息工具和选择工具可方便地查到所要查找的现实地物信息,MapInfo 的标注功能可以为读懂图纸提供帮助; 具有可编辑性,可以随时编辑、修改属性数据和图形,方便对园林绿化动态的随时更新; MapInfo 最主要的特点是结合属性数据库管理软件,依据地理信息系统原理和 GPS 技术,建立一体化区域化的地理信息平台。总之,随着卫星地图影像和 GPS 技术的兴起,对各工程图纸进行进一步的地理坐标化,对图纸配准、拼接,建立所管理区域的园林绿化信息管理平台 and 园林绿化信息数据库才是园林绿化管理的未来——城市园林绿化信息管理系统。



图1 2种软件结合运用效果

Fig.1 The drawing of combination

参考文献

- [1] 罗云启, 罗毅. 数字化地理信息系统 MapInfo 应用大全 [M]. 北京: 希望电子出版社, 2001.
- [2] 王彬华. 中文 AutoCAD 2005 [M]. 上海: 上海交通大学出版社, 2005.
- [3] 范明华, 徐志忠, 李明巨, 等. 基于 3S 技术的城市园林绿化管理信息系统的解决方案 [C] // 2005 数字江苏论坛电子政务与地理信息技术论文集, 南京, 2005.

状对结论成立的重要性。如果 1 个症状对结论成立的重要性较大, 则应使它具有较高的权值。如在猪结核病中, 症状“结核菌素反应阳性”对结论的成立具有很大的重要性, 因此应给予较高的权值; 相应症状的独立性。如果一个症状具有较强的独立性, 而其他症状对它具有依赖关系, 则应使其具有较大的权值; 权值的取值范围一般规定为 [0, 1], 且应该使其满足归一条件, 即每个疾病对应的所有症状的加权因子之和为 1。由于猪病的发病原因和临床表现的复杂性, 权值的确定要反复验证, 才能尽可能地接近临床实际情况。

参考文献

- [1] 武波, 马玉祥. 专家系统 [M]. 北京: 北京理工大学出版社, 2001.
- [2] 李道亮. 智能化水产养殖信息系统的设计与初步实现 [J]. 农业工程学报, 2005, 16(4): 135 - 138.
- [3] 田东, 傅泽田. 网络化淡水虾养殖专家系统的设计 [J]. 计算机应用研究, 2004(6): 24 - 25, 28.
- [4] 王靖飞. 动物疾病诊断专家系统的研究与应用 [D]. 哈尔滨: 东北农业大学, 2002.