

曹丽英, 姚玉霞, 于合龙, 等. 基于模糊本体的玉米病害诊断模型的构建[J]. 华南农业大学学报, 2014, 35(2): 101-104.

基于模糊本体的玉米病害诊断模型的构建

曹丽英¹, 姚玉霞², 于合龙¹, 马 丽¹, 陈桂芬¹

(1 吉林农业大学 信息技术学院, 吉林 长春 130118, 2 长春科技学院 信息工程学院, 吉林 长春 130600)

摘要:【目的】建立一种模糊本体的玉米病害知识表示及推理模型。【方法】将利用数据源得到的数据构建模糊本体, 结合模糊理论, 利用本体技术构建病害诊断模型。【结果和结论】该模型给本体的形式化定义和模糊理论的不确定表示的优势, 将模糊理论引入到本体构建中, 突破了本体只能表示确定知识的局限性。将该方法用于玉米病害诊治中, 对农业生产中病害诊治起到了很好的指导作用。

关键词: 本体; 模糊理论; 诊断模型; 玉米病害

中图分类号: S435.13

文献标志码: A

文章编号: 1001-411X(2014)02-0101-04

Construction of the model in maize disease diagnosis based on fuzzy ontology

CAO Liying¹, YAO Yuxia², YU Helong¹, MA Li¹, CHEN Guifen¹

(1 School of Information Technology, Jilin Agricultural University, Changchun 130118, China;

2 Institute of Information Engineering, Changchun University of Science and Technology, Changchun 130600, China)

Abstract: 【Objective】 To construct a fuzzy ontology of maize disease knowledge representation and reasoning model. 【Method】 Using data from sources to construct fuzzy ontologies, which were combined with fuzzy theory to construct the disease diagnosis model by ontology technology. 【Result and conclusion】 The theory of fuzzy was introduced into the description of ontology form and the shortcoming of an ontology representing the narrow defined knowledge only was overcame by this model. Combining the formal definition of ontology with the indeterminacy presentation advantage of the theory of fuzzy, the robustness of maize expert system for diagnosis and treatment of diseases was increased. Meanwhile, this model improved the accuracy of diagnosis and treatment of maize diseases and presented a new knowledge representation reference for other agricultural production expert systems.

Key words: ontology; fuzzy theory; diagnosis model; maize diseases

网络信息化时代的快速发展, 给传统农业带来了巨大的冲击和挑战, 依靠信息技术, 转变农业生产的管理、经营模式, 快速向信息农业、精准农业、智能农业转变是农业现代化亟待解决的问题。精确农业

在美国等发达国家已取得长足发展, 在我国也取得一些成绩, 但是仍处在探索阶段^[1]。精确农业的发展是解决“三农”问题的重要手段。农业专家系统是精准农业的一部分, 农业专家系统的研制与应用, 是农

收稿日期: 2013-01-01 优先出版时间: 2014-01-03

优先出版网址: <http://www.cnki.net/kcms/detail/44.1110.S.20140103.0828.015.html>

作者简介: 曹丽英(1978—), 女, 副教授, 博士, E-mail: caoliying99@163.com; 通信作者: 陈桂芬(1956—), 女, 教授, 博士, E-mail: guifchen@163.com

基金项目: 863 计划项目(2006AA10A309); 长春市科技特派员项目(2009245); 长春市科技局项目(09KT4); 吉林省世行贷款项目(2011-Z20); “十二五”吉林省教育厅科学技术研究项目(201247, 201248); 吉林省教育科学规划课题(GH11161); 吉林省科技发展计划项目(201101114)

业信息化、现代化的新途径. 农业专家系统也被称为农业智能系统, 是一个具有大量农业专门知识与经验的计算机系统, 它应用人工智能、数据挖掘等技术, 依据一个或多个农业专家提供的特殊领域知识、经验和方法进行推理和判断, 模拟农业专家就某一复杂农业问题进行决策^[2], 能及时解决农业专家不在现场的问题^[3].

本体的应用使计算机人工智能领域获得了许多新的研究热点. 哲学中的本体反映的是特定分类的观察方法^[4]; 而计算机人工智能领域的本体是一个复杂的概念系统, 用一些词汇描述事先规定的概念和规则体系, 是客观世界实体概念、实体关系以及实体属性等信息的形式化描述, 而且还可以加上一些与这些描述相关的领域约束^[5], 构成用于专家系统中知识推理的逻辑基础.

1 模糊本体数据库的构建

模糊逻辑提供了一种处理不确定信息的处理方式, 它把隶属函数与集合中的元素对应起来, 使这个元素与 $[1, 0]$ 区间的某个数建立联系, 称这个数为相对于集合的隶属等级^[6]. $X = \{x\}$ 表示对象空间, x 表示 X 的一般元素, 在 X 中的模糊集 F 可以表示为: $F = \{x, \mu_F(x) | x \in X\}$, 其中, $\mu_F(x)$ 表示 x 在模糊集 F 中的隶属等级^[7]. 当 $\mu_F(x) = 0$ 或 $\mu_F(x) = 1$ 时, F 是非模糊集; 当 $\mu_F(x)$ 在 $0 \sim 1$ 之间取值时, F 是一个模糊集, 此时 x 是否属于 F 不是一个明晰的概念.

模糊领域本体是一个六元组^[8]:

$$O_F = \{C_x, C_t, P_F, O, R, A_o\},$$

其中, C_x 是对实体概念的描述, 是一个模糊的概念, 例如发病时期、发病部位等概念的描述; C_t 是模糊概念 C_x 的集合, 是其所有可能的取值, 以病斑颜色为例, $C_t = \{\text{浅褐色、褐色、黄褐色、灰褐色、深褐色}\dots\}$; P_F 是概念属性集, 表示为四元组 $P_F(C_P, V_P, U, \mu_P)$, C_P 是本体概念集, V_P 是模糊属性值, $U = \{X_1, X_2, \dots, X_n\}$ 是论域, μ_P 为模糊集中的隶属等级; O 是隶属函数集合, 它与模糊概念集 C_t 是同构的. 每个隶属函数具有表示精确模糊概念的语义作用; $R \subseteq C_t \times C_t$ 是概念关系集, 是一个二元关系, 可以表示概念间相交、不相交、包含、补等基本关系, 如没有病斑与病斑颜色和病斑形状是不相交的关系; 也可表示概念之间模糊的关系, 例如浅褐色、褐色、黄褐色、灰褐色、深褐色这些颜色之间的包含关系; A_o 是公理的集合, 包括演绎规则和反应式规则, 用于对知识的推理或演绎, 用来处理知识库.

模糊本体的每个类、属性和个体在创建的时候

都给定了一个表示概念间关系的亲密程度的量——模糊度^[9]. 在添加、删除、修改这些对象的时候, 就会考虑这个概念应该是在什么位置, 添加后个体到底是哪个类的个体; 删除不管是属性、个体还是一个类之后, 都会对本体系统的结构产生影响; 同样修改一个属性的权值, 也许会造成整个本体构建模型大的改变. 模糊本体如果用来处理具体模糊属性值的模糊信息比较容易, 但对于处理具有模糊概念的属性值时, 由于模糊概念的量化问题, 不能很好地表达模糊概念间的语义关系和相关程度, 特别是在不同本体概念之间, 模糊语义的整合有很大的局限性, 而且对同一个属性集, 在不同的概念中可能有不同的取值^[10], 这是本模型的实际应用中遇到难点问题, 本研究针对具体病害特征, 通过咨询领域专业, 得到模糊属性的量化值.

模糊本体数据库的数据来源于咨询领域专家、查阅专业书籍和网络资源等, 在提取属性数据和专业术语的基础上, 结合模糊理论, 利用本体技术进行构建. 模糊本体数据库的生成如图1所示.

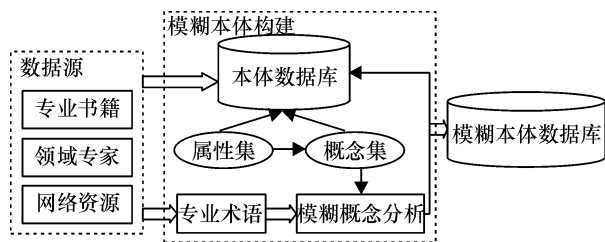


图1 模糊本体数据库的构建过程

Fig. 1 Construction process of fuzzy ontology database

2 玉米病害的模糊知识本体的构建和诊断

2.1 研究数据来源

咨询该领域专家, 如吉林农业大学、吉林省农业科学院植物保护研究所等单位的有关专家得到相关数据资料; 查阅本领域专家出版的专著、教材及科技期刊上发表的文章、学位论文等科技文献资料, 整理、归纳后抽提相关知识. 结合课题组积累的试验研究数据, 根据玉米病害与产量关系的试验, 收集整理玉米苗期、大喇叭口期和成熟期等玉米生长性状指标的数据 29 250 条; 收集整理玉米病害发生程度的数据共 4 320 条, 拍摄照片 189 张; 玉米虫害田间数据 5 760 条, 拍摄照片 306 张; 玉米草害田间数据 4 320 条, 拍摄照片 286 张. 吸收科研成果和数据资料, 收集各示范区的农民和农技人员的知识和实践经验, 补充知识来源. 图片数据来源于吉林农业大学植物保护教研室.

本研究整理了吉林省常见的13种玉米病害,包括玉米矮花叶病、玉米粗缩病,玉米大斑病、玉米顶腐病、玉米褐斑病、玉米灰斑病、玉米茎腐病、玉米青枯病、玉米丝黑穗病、玉米弯孢菌叶斑病、玉米纹枯病、玉米小斑病和玉米圆斑病。

2.2 玉米病害模糊本体知识库的构建

由于玉米病害诊断的原则和基本方法不同,另外真菌、细菌、病毒、线虫等不同病原物在玉米苗期和成熟期引起重要病害的症状特点和发病规律,各种非侵染性病害的发生等原因,导致玉米病害存在大量模糊与不确定现象。例如在玉米病诊断过程中,关注的信息有:发病时期(苗期、生长期、成熟期),发病部位(根、茎、叶等),有无病斑(有、无),病斑颜色(褪绿色、草绿色、浅褐色、褐色、黄褐色、灰褐色、深褐色等),病斑形状(圆形、椭圆形、长形、线形等),诊断过程中的这些特征描述语言值都是模糊的,在不同情景、不同角度,判断会存在一定的差异。这就不可避免地要解决模糊概念的形式化表示问题^[11]。

采用模糊本体模型对玉米病害诊治建立知识本体。玉米病害模糊本体结构如图2所示,包括各种病害的概念集合、病害属性间的关系集合、病害症状的属性集合以及专家提供的属性量化模糊值等。

在此基础上,利用计算机中的结构化查询语言,可以针对提供的病害特征,进行玉米病害的模糊诊断^[12-13]。

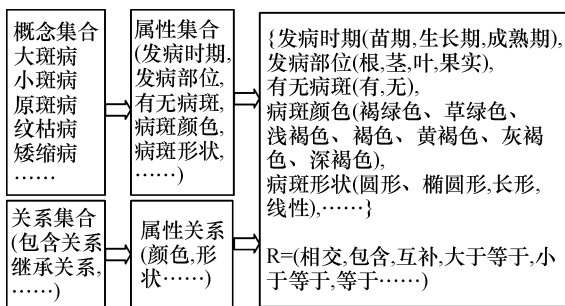


图2 玉米病害模糊本体结构图

Fig. 2 Fuzzy ontology structure graph of maize diseases

2.3 模糊本体推理规则的建立

本系统的规则库使用的模糊推理过程为:病虫害症状信息输入后,将文字信息经过事先专家给定的经验值进行数个转化等数据预处理工作,得出该数据的综合值;然后将元数据库中储存的数据与模糊评价规则库中符合要求的规则进行匹配,从成功匹配的规则中计算隶属度值^[14];最后按照最大隶属度原则,得出最终的诊断结论,提供给用户。

2.4 玉米病害模糊本体的应用

对于农业领域知识通常需要一些模糊语言描

述。在玉米病害诊断本体中,“病害”有属性{发病时期,发病部位,有无病斑,……},其语言值(发病时期{苗期,生长期,成熟期},发病部位{根,茎,叶,果实}……病斑颜色{褪绿色,草绿色,……,浅褐色}等。本文考虑到模糊本体推理机应用于文本信息推理的局限性,采用的推理原则是:最优决策、多标准决策、知识库图像决策、经验决策、逻辑的模糊控制理论决策。选取具有类似特征的实例数据,以较为相近特征的病害为研究对象,以发病时期、病斑颜色和病斑形状为属性,根据领域专家的意见设定准则给出初始值,同时计算出评价准则的权值,发生程度1、3、5、7、9分别表示一般相似、中等相似、很相似、非常相似和极其相似。

例如诊断玉米大斑病,玉米小斑病和玉米灰斑病3种病害,先要对它们的发病时期,病斑颜色和病斑形状进行评价,再根据诊断标准对症状的描述性文字进行数据量化,对发病时期为初期,量化的权值为1。根据发生程度量化的数据,计算出评价准则的权值 $W_{i,j}$, $W_{i,j}$ 是相对于标准的相似程度取值。根据 $\bar{W}_i = (\prod_{j=1}^n W_{i,j})^{1/k}$, $\bar{W} = \sum_{i=1}^k \bar{W}_i$, 其中 $W_i = \bar{W}_i / \bar{W}$ 进行计算诊断标准的权值,其中 i 表示病害名称, j 表示发病时的参数^[15]。结果如表1所示。

表1 诊断标准的初始值

Tab. 1 The value of initial diagnosis standard

项目	玉米大斑病	玉米小斑病	玉米灰斑病	相似度 $W_{i,j}$
发病时期	1	3	1/3	0.258
病斑颜色	1/3	1	1/5	0.105
病斑形状	3	5	1	0.637

从表1可见,从而得出各个属性的诊断权值分别是0.258、0.105和0.637。利用三角函数模糊数对它的程度进行定性判断,得出决策诊断结果,大斑病的发病时期为初期,病斑颜色清晰,病斑形状为连续;小斑病的发病时期为中期,病斑颜色为非常清晰,病斑形状为连续;灰斑病的发病时期为初期,病斑颜色的值为浅,病斑形状为不连续。

在这里,模糊本体模型的框架中,不但包括了传统本体描述方法中的5要素,而且还引入了模糊理论,这种方法的提出,一方面能更好地解决了农业知识的不确定性,另一方面还能提高检索的效率,可以实现模糊概念的语义查询。如模糊查询语句[SELECT Maize disease FROM Data source WHERE (property of concept) < > / > = / < / < = / = = / ! = > / (linguistic value of fuzzy concept)],式中用关系运算

符进行条件比较,也可使用 SQL 语言中的其他运算符表示查询条件,如 in, except 等。

玉米病害属性描述的部分代码:

```
<fuzzyowl2 fuzzytype = "concept" >
<Concepttype = "weightedsum" >
< Concepttype = "weighted" value = "1" base = "
C1"/>
< Concepttype = "weighted" value = "3" base = "
C2"/>
< Concepttype = "weighted" value = "5" base = "
C3"/>
< Concepttype = "weighted" value = "7" base = "
C4"/>
< Concepttype = "weighted" value = "9" base = "
C5"/>
</Concept >
</fuzzyowl2 >
```

通过对玉米病害的概念描述,我们可以利用上文的定义对玉米病害知识提取,得出模糊本体描述的伪代码:

$$P_L(\text{Corn big spot}) = \{ (\text{Growth period, leaf, sick spot/ dark brown/hazel/pewter, } \dots, \text{ Boolean, 1}) \}$$

$$P_U(\text{Corn big spot}) = \{ (\text{Growth period, leaf, yellow/light brown/ dark brown, often/usually/yes/no, 0} \dots 2) \}$$

利用该模型,针对不同病害进行识别,诊断结果如表 2 所示。从表 2 可以看出各玉米病害的识别平均较高,试验识别出来的病害平均准确率为 80.67%。通过将试验识别出来的病害的数据和真实数据病害的数据作对比,该诊断模型的诊断结果与实际发生病害结果基本相符,病害的识别可以应用在玉米农业生产中,说明该基于模糊本体的玉米病害诊断模型在实际的农业领域中应用是可行的、有效的。

表 2 诊断结果表

Tab.2 Results of diagnosis

玉米病害类别	识别发现的病害数目	样本中实际病害数	识别率/%
大斑病	29	34	85
小斑病	35	32	91
丝黑穗病	13	18	72
青枯病	11	16	69
圆斑病	17	24	71
褐斑病	23	22	96

3 结论

农业领域的评议知识的信息量正在不断的增加,对农业知识的探索早已不局限在精确信息上,目前在其他领域不精确信息的研究较多,由于农业领域的知识不但具有不确定性,而且具有动态性,所以在不确定性信息研究中,具有一定实际困难,如何在不确定性推理的基础上,实现效率高、范围广的推理将是农业知识领域的研究重点。本研究对传统的本体模型进行了扩展,引入了模糊理论,并将其应用到玉米病害诊断中,实现了模糊本体的互相表示性、模糊性和演化性,并通过试验验证了该模型的有效性。模型的研究目的是使农业专家系统更好更完全地抽象描述现实世界,模拟农业专家,从而解决农业生产中出现的一些诊断问题。

参考文献:

- [1] 葛中一. 农业专家系统的应用及发展[J]. 山东师范大学学报:自然科学版,2008,23(4):150-153.
- [2] 赵林峰,熊兴耀. 农业专家系统的问题与发展[J]. 农业网络信息,2007(6):43-45.
- [3] 谢琪. 基于本体方法构建中医药概念信息模型的方法学示范研究[D]. 北京:中国中医科学院,2011.
- [4] 刘军锋. 基于本体与对象的地理空间数据模型研究与应用[D]. 桂林:桂林理工大学,2009年.
- [5] 杨青,陈薇,闻彬. 面向语义信息查询的模糊本体模型[J]. 计算机工程,2010,36(8):194-196.
- [6] 刘小寅. 不精确本体模型与构建研究[D]. 大连:大连海事大学,2011.
- [7] 陈薇. 面向语义信息检索的模糊本体自动化构建的研究[D]. 上海:华中师范大学,2010.
- [8] 杨柳. 模糊本体建模方法及语义信息处理策略研究[D]. 长沙:中南大学,2011.
- [9] 翟军,陈燕,沈立新. 基于模糊本体的智能交通系统知识建模[J]. 大连海事大学学报,2008,34(2):91-94.
- [10] 刘昕民,桂卫华,杨柳,等. 基于模糊领域本体的专家遴选服务研究[J]. 北京理工大学学报,2013,33(5):484-489.
- [11] 王霞,李功振. 基于模糊本体的水环境评价系统的研究[J]. 安徽农业科学,2010,38(35):20306-20308.
- [12] 戴才萍,黄义德,钱平,等. 水稻病虫害草害本体的构建研究[J]. 广东农业科学,2011(1):191-194.
- [13] 林潇. 基于本体的水稻病虫害诊断专家系统研究[D]. 合肥:安徽农业大学,2011.
- [14] 王皓. 基于本体论的玉米病虫害诊断系统的设计与实现[D]. 吉林:吉林大学,2011.
- [15] 姜大庆,蔡银杰. 基于本体的蔬菜病虫害知识库构建[J]. 江苏农业科学,2012,40(7):368-370.

【责任编辑 霍欢】