

农业智能系统原理与应用

陈桂芬 等编著



科学出版社

www.sciencep.com

农业智能系统原理与应用

陈桂芬 等编著

科学出版社

北 京

内 容 简 介

本书介绍了农业智能系统研制的基本原理、主要技术和应用实例。前六章重点阐述了农业智能系统研制的基本概念、基本原理和主要研制技术;第七章至第十二章通过玉米、水稻、蔬菜、人参、大豆和畜牧业生产智能系统的研制实例,详细说明了农业智能系统开发的主要技巧和操作规范。为了使各相关农业应用和研发人员便于了解和掌握农业智能系统原理,并依此提高解决具体领域问题的能力,书中还给出了许多例题和研制实例。

本书可供从事农业智能系统研究和开发的科技人员参考,也可作为高等院校相关专业的教材或参考书。

图书在版编目(CIP)数据

农业智能系统原理与应用/陈桂芬等编著. —北京:
科学出版社,2006

ISBN 7-03-018124-7

I. 农... II. 陈... III. 人工智能—应用—农业
IV. S-39

中国版本图书馆 CIP 数据核字(2006)第 119558 号

责任编辑:杨富亿/封面设计:肖海福

科学出版社出版

北京东黄城根北街16号

邮政编码 100717

<http://www.sciencep.com>

吉林农业大学印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

2006年10月第一版

开本:787×1092 1/16

2006年10月第一次印刷

印张:19 1/8

印数:1~500

字数:440 000

定价:50.00 元

(如有印装质量问题,我社负责调换)

前 言

近年来,随着计算机技术的飞速发展和全球知识经济一体化的加快,信息技术已成为 21 世纪的两大科技支柱之一;用信息化带动工业化、带动农业现代化这一观点得到越来越多人的共识;从工业化社会走向信息化社会,从农业现代化走向农业信息化,是人类社会进步和现代农业发展的里程碑。未来的农业发展是科技实力和人才竞争的时代。农业信息化是现代农业发展的必然趋势。农民对农业知识的日益增长的需要与农业专家稀缺的矛盾是当前我国农村的一个主要矛盾,这个矛盾不解决就会影响农民的增产和增收,最终影响三农问题的解决。那么,怎样解决这个矛盾呢?唯一的渠道就是要实现农业的智能化。“九五”以来,国家科技部高技术司和农社司先后支持了国家“863”计划农业智能系统和农业数字化技术等农业信息技术研究与应用课题,已在全国建立了 21 个智能化农业信息技术应用示范区和多个农业数字化技术应用示范区;研制开发出一批实用农牧业生产智能系统和精准农业设备,并在应用中取得了很好的经济效益与社会效益。

智能农业和数字农业的实质,是将智能系统和数字化技术引入农业,让其发挥人类专家的作用。农业智能系统和农业数字化技术作为一种现代化的载体,能够方便快捷而生动形象地向各级农业生产人员和管理干部传播各类实用的农业知识、管理知识和各种经济信息,提高广大农业技术人员和农民的科学种田素质;对提高科技成果转化率、提高农业领导者的决策水平、农业科技人员的管理水平和广大农民的科学种田素质所起的作用更是不可低估。有关专家在评述我国智能化农业信息技术实施十多年所取得的丰硕成果时指出,最重要的成果不仅仅是增加了多少产量,取得了多少效益,而农业智能系统更大的贡献还有两条:一是让信息技术走进农家;二是初步探索了农业信息化建设的体系框架。

本书作者有幸参加了国家“863”计划“农业智能系统研究与应用”和“大规模现代农业数字化技术应用研究与开发”项目研究。在项目的实施过程中,深切感受到农业领域中专家、管理人员和种养殖专业户对农业信息技术的迫切需求,由此萌生编著此书的想法。本书以近年来学习和讲授的人工智能与专家系统教材及搜集整理的农业信息技术资料为基础,以上述诸课题的科研成果为主要内容,系统地阐述了农业智能系统研制的基本原理、技术与应用实例。

全书共分十二章。本书总体设计由陈桂芬教授负责,并由陈桂芬教授统稿和最后定稿。

本书中内容提要、前言、第一章、第二章和第五章(部分)由陈桂芬执笔;第三章和第六章由曹丽英执笔;第四章、第五章(部分)和第十一章由于合龙执笔;第七章由王越执笔;第八章和第九章由姚玉霞执笔;第十章和第十二章由司秀丽执笔。另外,王国伟、曹丽英、侯元村、李东明、刘鹤、伞晓辉、张莹和毕春光、邓蕾蕾等同志分别为第七章、第八章、第九章、第十章和第十二章提供了部分基础文稿,在此一并表示感谢。

在编著此书过程中,深深体会到,面对农业信息技术领域研究和实践飞速发展的现状,我们撰写的书稿仅能起到抛砖引玉的作用,圆了我们为“三农”服务的梦。由于时间

和精力的关系,特别是受学识水平限制,书中难免有诸多错误和不足,敬请广大读者给予更多的批评指正。

在此,向将我带入农业智能系统研制领域的刘大有教授、庄铁成研究员、李明树研究员、吴泉源教授、石纯一教授和国家“863 - 306”智能化农业专家系统总体组的全体成员表示衷心感谢;向亲自领导农业智能系统推广应用工作的原吉林省科技厅李广臣副厅长、张伯军副厅长和省科技厅农村处陈忠亮处长等领导表示诚挚谢意。

编 者
2006 年 2 月

编辑委员会

主 任： 陈桂芬

副主任： 姚玉霞 司秀丽 于合龙

编 委(按姓氏笔画为序)

于合龙 王 越

司秀丽 陈桂芬

姚玉霞 曹丽英



陈桂芬 教授、硕士生导师、在读博士。1956年2月生于吉林省长春市，1982年毕业于吉林农业大学，1999年在吉林大学获计算机专业理学硕士，2005年在吉林大学攻读计算机应用技术专业博士学位。现任吉林农业大学信息技术学院院长，计算机应用学科带头人；国家“863”计划智能化农业信息示范工程技术总体组专家、中国农学会计算机应用分会常务理事、吉林省有突出贡献专家。主要从事计算机农业应用、专家系统和数字农业等方面的研究。1996年以来，先后主持、承担“863”项目8项、省部级项目6项；获吉林省科技进步奖一等奖2项、吉林省科技进步奖二等奖1项；获国家软件著作权登记11项、发明专利1项；在国际会议及核心刊物上发表论文28篇，出版专著3部。

目 录

前言

第一章 农业智能系统综述	(1)
第一节 农业智能系统的内涵与外延	(1)
第二节 农业智能系统的特征	(3)
第三节 农业智能系统的结构与功能	(4)
第四节 农业智能系统的研制基础	(6)
第五节 农业专家系统开发平台	(29)
第六节 农业智能系统的应用	(36)
第二章 基于规则表示与智能系统	(38)
第一节 产生式系统的产生前景	(38)
第二节 产生式系统的基本原理	(39)
第三节 匹配、匹配冲突和冲突消解	(52)
第四节 应用实例	(58)
第三章 基于案例推理与智能系统	(63)
第一节 CBR 的技术概述	(63)
第二节 CBR 的基本内容和 CBR 的模型	(66)
第三节 CBR 案例的特征抽取和描述	(73)
第四节 案例检索策略	(74)
第五节 结论改写与机器学习	(80)
第四章 不确定性知识与智能系统	(82)
第一节 概述	(82)
第二节 基于规则的不确定性推理方法	(86)
第三节 贝叶斯网概述	(91)
第四节 贝叶斯网的构建	(97)
第五节 贝叶斯网的推理	(103)
第五章 与地理信息系统和数学模型结合的智能系统	(112)
第一节 概述	(112)
第二节 GIS 与智能系统的结合	(114)
第三节 数学模型与智能系统的结合	(127)
第四节 前景展望	(133)
第六章 数据库与信息管理系统	(144)
第一节 数据库管理系统的概述	(145)
第二节 数据模型及数据库语言 SQL	(147)
第三节 数据库管理系统的设计	(150)

第四节 数据库信息管理系统的研制	(152)
第七章 多媒体玉米生产智能系统	(156)
第一节 系统总体结构	(156)
第二节 主要功能	(157)
第三节 实现方法	(162)
第四节 知识表示与推理机制	(164)
第五节 数据库	(167)
第六节 模型与模型库	(169)
第七节 系统运行	(171)
第八章 多媒体水稻生产智能系统	(196)
第一节 系统总体结构	(196)
第二节 系统的主要功能	(196)
第三节 知识表示与推理	(198)
第四节 系统的实现方法	(204)
第九章 多媒体蔬菜生产智能系统	(207)
第一节 系统总体结构	(207)
第二节 系统(单机版)的实现方法	(212)
第三节 系统(网络版)的实现方法	(216)
第四节 知识表示与推理	(220)
第五节 多媒体教学课件	(231)
第十章 多媒体人参生产智能系统	(236)
第一节 系统总体结构、流程图	(236)
第二节 主要功能	(238)
第三节 实现方法	(240)
第四节 操作方法	(258)
第十一章 多媒体大豆生产智能系统	(265)
第一节 系统的总体结构与功能	(265)
第二节 基于规则的大豆生产智能决策模块	(268)
第三节 基于贝叶斯网的大豆病虫害诊断模块	(270)
第十二章 畜牧业生产智能系统	(279)
第一节 多媒体鹅生产技术专家系统	(279)
第二节 畜禽饲料配方优化系统	(287)
第三节 畜禽疾病诊断专家系统	(291)
第四节 奶牛饲养专家系统	(293)
参考文献	(296)
附页	(298)

第一章 农业智能系统综述

农业智能系统也称专家系统,是智能化农业信息技术研究的一个重要领域。国际上关于农业专家系统的研究兴起于20世纪70年代末期。1978年美国Illinois大学开发的针对Illinois地区的大豆病虫害诊断专家系统PLANT/ds,被认为是第一个农业专家系统。之后Illinois大学于1982年又开发了玉米螟虫害预测专家系统PLANT/cd。早期的农业专家系统还有Virginia理工学院的苹果疾病诊断专家系统POMME;日本千叶大学的西红柿病虫害诊断专家系统MICCS;美国农业部研制的棉花管理专家系统GOSSYM/COMAX;Texas A&M大学开发的得克萨斯州棉花病虫害防治专家系统COTFLEX;加利福尼亚大学的生产管理专家系统CALEX;荷兰Utrecht大学开发的奶牛饲养和肉猪生产专家系统ANIMAL FARM;荷兰Wageningen大学开发的农业企业经济分析专家系统CHESS;美国普渡大学的谷物销售策略指导专家系统Grain Market Advisor等。多数专家系统是由地方政府部门借助高等院校的科研力量开发的,一般具有明显的区域适用特征。近年来,出现一些专门进行农业专家系统研究开发的科研机构,比如Virginia理工学院信息系统和昆虫研究实验室、埃及农业专家系统中心实验室。在将Bayesian网应用于农业专家系统方面,国际上已经有了很多研究工作,但引人注目的产品比较少见。丹麦农业科学信息网络(DINA)是一个由丹麦政府部分资助的农业信息化科研网络,其重心是开展理论数学、统计学和计算机科学在农业领域中的应用研究,其中也包括Bayesian网在农业中的应用。

我国农业智能专家系统研究起步于20世纪80年代初。1985年,中科院合肥智能机械研究所与安徽省农科院土壤肥料研究所合作率先研制“砂姜黑土小麦施肥专家系统”。其后,各地高校、研究所和农科院相继开发了许多农业专家系统。90年代,在国家“863”计划、国家自然科学基金委和科技部、农业部等国家部门的支持下,我国农业智能系统的研究开发在广度和深度上均有了很大进展。由国家“863”计划信息领域计算机软硬件技术主题(前身智能计算机系统主题)持续支持的“智能化农业信息技术应用示范工程”,是我国目前得到政府持续支持时间最长、参与人员最多、实施区域最广的一个项目,在22个省市建立了23个智能化农业信息技术应用示范区,已有5个农业专家系统开发平台。本书也正是在这个课题背景下完成的。

第一节 农业智能系统的内涵与外延

一、农业智能系统的基本定义

智能系统也称专家系统(expert system, ES),是一个智能程序。它能对那些需要专家知识才能解决的应用难题,提供相关领域权威专家水平的解答(Michie, 1979; Feigenbaum, 1977; Hayes-Roth, 1983)。专家系统,可以说是一个运用知识进行推理的计算

机程序。推理就是使用某种符号逻辑,从一些事实得到结论的过程。从功能上,可以把专家系统定义为“一个智能程序,它能对那些需要专家知识才能解决的应用难题提供专家水平的解答”;从结构上讲,可把专家系统定义为“由一个专门领域的知识库,以及一个能获得和运用知识的机构构成的一个问题求解系统”。

总之,专家系统是一个智能程序系统;有大量的、高水平领域专家的知识;有领域专家解决问题的思维方法。专家系统所处理的问题是依据已积累的知识来求得问题解答,一般没有准确的数学公式来表达,这就是专家系统与“一般问题求解”方法的不同之处。数据+算法→传统程序,知识+推理→专家系统。

从而可以引申出农业智能系统的定义:农业智能系统是一个拥有大量权威农业专家的经验、资料、数据与成果构成的知识库,并能利用其知识,模拟农业专家解决问题的思维方法进行判断、推理,以求得解决农业生产问题结论的智能程序系统。

二、农业智能系统的引申含义

从广义上来讲,农业智能系统可涵盖农业生产的智能化。我们知道,农民对农业知识的日益增长的需要与农业专家短缺的矛盾是当前我国农村科技服务体系的一个主要矛盾,这个矛盾不解决就会影响农民的增产和增收,最终影响三农问题的解决。那么,怎样解决这个矛盾呢?主要渠道之一就是要实现农业生产的智能化、数字化和网络化。

其中,智能化是以收集、整理权威农业领域专家的知识,集成单项农业生产新成果,基于专家系统、GPS、GIS、数据库和多媒体等技术,利用面向对象可视化软件开发工具研制的农业生产智能系统;数字化是指对农业生产所涉及的对象和全过程进行数字化表达、设计、控制和管理,它可以使农业生产由粗放生产到精细操作、由自然生长到人为控制,这需要一定的硬件体系来支撑;而网络化是指将智能化和数字化的成果传播出去。这“三化”当然不是截然分开的,而是有紧密的联系。但从一定意义来讲,没有农业生产的智能化,就不可能实现真正意义上的农业过程数字化;没有农业的智能化、数字化,就没有农业信息技术的全面应用,也就不可能实现农业生产管理的网络化。

可以预期,以智能系统为主的农业智能化、数字化、网络化,可以使农业生产由经验性到量化、规范、集成和智能化;由分散封闭到有效获取和利用信息;由经验判断到宏观测报和科学管理,是传统农业的一次深刻革命。这“三化”的大力发展,必将使农业实现更高的效率,农产品达到更高的质量,使农业更好地满足人们生活不断增长的需求。同时,又使农业环境得到更有效的保护,实现农业符合现代化要求的可持续发展。

当前农业智能系统的发展趋势是,与数据库、多媒体、管理信息系统、决策支持系统、专家系统(ES)、地理信息系统(GIS)、全球定位系统(GPS)、遥感技术(RS)、计算机网络、优化模拟、多智能体、数据采集与处理技术、农作物生产新技术和农业工程技术等诸多技术结合,利用构件化技术进行集成,从而提供更为强大的功能、更为简单便捷的操作,使农业生产管理同市场经济、宏观决策有机结合,成为促进农业产业化的重要工具。随着我国农业产业和专家系统理论研究的逐步发展,优化建模、储备完善各种资源数据库、采用更为合理的核心技术、为农业产业的预测、决策、诊断提供精确指导,在农业生产领域实现信息技术的全面应用。这将成为我国农业生产智能系统的重要发展方向。同时,对于从事

农业智能系统研究和开发的科技人员来说,还应注意到在农业智能系统的研究发展与我国当前的农业生产力发展水平尤其是产业化水平之间,存在着强烈的相互制约、相互促进的关系。

第二节 农业智能系统的特征

农业智能系统的最基本目标在于为农业科技人员及农民提供农业方面的各种建议。它可以帮助人们进行设计(如灌溉系统和防洪系统的设计)、选择(如适宜作物品种或市场销路的选择)、诊断(识别)(如作物、畜禽疾病和作物的缺素诊断)、预测(如对洪涝灾害和重大疾病等偶然事件的预测)、决策(如饲料配方、施肥配方等农业生产不同方面的智能决策)。因而,农业智能系统既具有一般专家系统的基本特性,又有其本身的特殊性质。

一、农业智能系统的基本特性

(一)领域知识的权威性与丰富性

由于农业生产智能系统能集权威专家的解决问题的经验、方法、技术于一体,所以一个好的农业生产智能系统应存储大量的实验数据与资料,建立起丰富的数据库、知识库与模型库,集多学科权威专家知识于一体,能准确表达多学科农业权威专家解决实际问题的思路。

(二)具有智能性和可控制性

农业生产智能系统是在农业领域表现出智能行为的系统。其智能水平取决于知识的多寡和质量;其推理机的优劣取决于相关事实、原理的可获得性和推理的完备性及效率。

农业生产智能系统在设施农业领域是表现出可控制性的监控系统。其监控系统能方便用户及时了解设施农业中温室内的相关数据,根据生产需要进行实时调控。其调控精度取决于数据采集的准确性、实时性和专家解决方案的可靠性和权威性。

农业生产智能系统对一般问题的求解能力,还应有较强的自适应性。

(三)推理的准确性和透明性

农业生产智能系统中将包含的知识整理成具有可表示性,即可用符号表示和符号处理(符号表示:用符号体系即形式化体系表示知识;符号处理:用符号推理处理用符号体系表达的知识的过),因而能够运用专家的知识 and 经验进行正确推理、判断和决策;

农业生产智能系统能对系统自身的行为即推理过程给出解释并回答用户提出的问题,给出系统所得出的结论的依据,以便让用户能够了解推理过程,提高对农业智能系统的信赖感;

这些特征使得智能系统的推理变得更为准确和透明,这也是农业智能系统与传统程序的一个显著差别。

(四)界面友好性和实用性

由于农业智能系统的应用对象主要是农业用户,因而大多数农业智能系统采用了多媒体技术和可视化技术,界面友好,操作简单,易于广大农业科技人员和生产用户使用。

(五) 系统的灵活性和快速性

农业专家系统能不断地增长知识,一旦某一领域有了新技术,马上可以对智能系统进行升级,其知识更新的速度要比其他信息媒体快得多,而且成本也低得多,由于这一特点,使得农业智能系统能不断进行补充和完善。

二、农业智能系统的特殊性质

农业智能系统不仅具有智能性和可复制性,可以部分的替代人类专家,而且还具有人类专家及书刊等信息媒体所不具有的优点。

第一,能够适应变化:智能系统可以根据不同的输入通过推理得到相应的输出,这就克服了文献上知识的静态性的缺点;

第二,具有交互性:用户和计算机中的智能系统可以通过友好的界面以问答的形式获取知识;

第三,能够实现知识的自动获取:通过机器学习技术可以从实验数据及案例中自动获取知识,并能够处理不确定的信息;

第四,可以集成多种媒体:智能系统可以集成文本、图像、声音等多媒体信息;

第五,可以集成多领域的知识:某一问题的解决,经常需要结合多领域的知识。智能系统能够有效地利用不同领域的知识使问题得到更好的解决;

第六,传播速度快:把智能系统放在服务器中,使用户可以通过网络来访问智能系统,这加快了知识的传播速度。

总之,农业智能系统作为一种实用工具,为农业知识的获取、发现、保存、传播、使用和评价提供了有效手段和方法。它不仅能汇集问题领域多个专家的知识与经验于一身,每天 24 小时不知疲倦地工作,而且它本身可以通过网络技术和分布式计算方法实现无穷无尽地迅速复制与克隆。

正是由于上述特点,使得农业智能系统获得迅速发展,应用领域从种植业到养殖业、水产业、加工业及农产品流通领域,涉及到农业生产的产前、产中和产后全过程,解决实际问题的能力也越来越强,对国民经济起着重要的作用。

第三节 农业智能系统的结构与功能

同一般的专家系统相同,一个基于规则的农业智能系统一般是由知识库、推理机、数据库(工作存储区,或称黑板结构;或称综合数据库)、人机界面、知识获取、解释机构构成(图 1-1)。其各部分的功能如下。

一、知识库

知识库主要用来存放农业领域专家的专门知识。在这里,知识以一定的形式表示。常用的知识表示法是产生式规则,在这种情况下,知识库包含“事实”和“规则”。知识表示要解决的问题是如何用计算机能够理解的形式表达和存储知识。知识库是推理机工作的重要对象,其中知识表示的好坏直接影响着整个系统的工作效率。

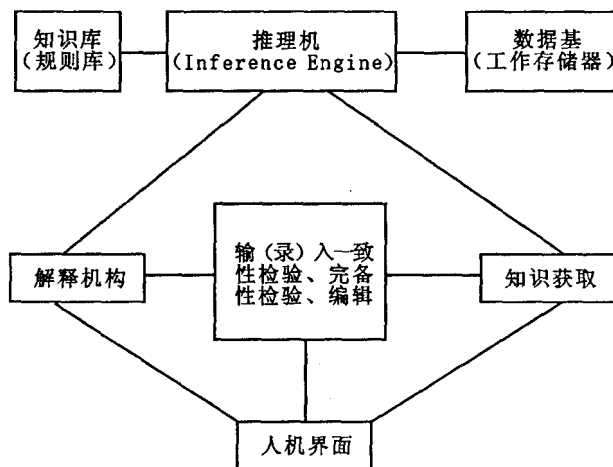


图 1-1 农业智能系统的构成

二、数据基(综合数据库)

综合数据库又称全局数据库或总数据库,它用于存储领域或问题的初始数据和推理过程中得到的中间数据(信息),即被处理对象的一些当前事实。

三、推理机

推理机用于记忆所采用的规则和控制策略的程序,使整个农业专家系统能够以逻辑方式协调地工作。推理机能够根据知识进行推理和导出结论,而不是简单地搜索现成答案。它包括推理方法和控制策略两部分。

四、知识获取

知识获取过程可以看作是农业专家的专业知识从知识源到知识库的转移过程。知识源主要是农业权威专家,也可以是书本或者其他试验资料。知识获取过程包括在知识库创建时识别出必要的知识并将其形式化。建成的知识库经常会发现有错误或不完全,所以知识获取过程还包括对知识库的修改和扩充。

知识获取的过程实质上是一个机器学习的问题。通过知识获取,农业专家系统不断地从外部(人类专家或书本等资料)和内部获取知识以修改和扩充知识库。主要的机器学习方式有:机械学习;基于解释经验的学习;基于事例的学习;基于概念的学习;基于类比的学习和基于神经网络的学习等。

五、解释机构

它能够向用户解释农业智能系统的行为,包括解释推理结论的正确性和系统输出其他候选解的原因。

六、人机界面

它能够使系统与用户进行对话,使用户能够输入必要的数据、提出问题和了解推理过程及推理结果等。系统则通过人机接口,要求用户回答提问,并回答用户提出的问题,进行必要的解释。

第四节 农业智能系统的研制基础

一个基本的农业智能系统的研制需要进行知识获取,确定知识表达方式和推理方式,建立知识库、数据库和模型库,编写推理等程序,最后进行程序的调式与修改。知识获取与知识表示、知识运用是建造一个智能系统的三个关键技术。本节重点介绍一下知识获取和知识表示。

一、知识获取

(一) 知识获取的意义

知识获取和知识表示是专家系统的“瓶颈”和“核心”问题。知识是专家系统推理的基础,知识表示是否恰当,直接影响专家系统运行效率和应用效果。成功的知识获取能够保证专家系统的推理精度,反之,智能系统可能起不到专家的作用。农业领域中的知识是复杂多变的,所以智能系统往往根据不同情况采用不同的知识表示形式。而且,不同种类的知识需要不同的表示形式和数据。在智能系统中,知识库的建造通常是知识工程师与领域专家密切配合的结果。领域专家自身或知识工程师与领域专家共同整理总结领域的知识、模型和专家的实践经验、研究成果等,按所建智能系统规定的知识表示形式,整理成一个个知识单元,放入知识库,这种过程称之为知识获取。

知识获取贯穿于一个专家系统生命周期的始终,是一个长期的、与专家系统共生死的过程,犹如新陈代谢对于一个人一样,粗略地说,可分为早期、中期和后期三个阶段。早期知识获取直接从知识源获取原始知识,这里所说的知识源包括人类专家的知识 and 记录在载体上的知识。中期知识获取是对已经得到的原始知识进行再加工,把那些隐含于原始知识中不能为我们直接利用的知识提炼出来。后期知识获取是用实践来检验已有的知识,达到去伪存真、去粗取精的目的。

(二) 知识获取的任务

知识获取的基本任务是为智能系统获取知识,建立起健全、完善、有效的知识库,以满足求解领域问题的需要。

1. 抽取知识

抽取知识是指把蕴含于知识源(领域专家、书本、相关论文及系统的运行实践等)中的知识经过识别、理解、筛选、归纳等抽取出来,以用于建立知识库。

抽取方法大体上分为四种类型。

第一种是人工分析方法,它主要提供一种分析方法学,由知识工程师或领域专家或两者合作分析文本或数据资料,从中提取知识。

第二种是统计分析方法,它要对试验数据做某种统计分析,以此获取知识。

第三种是自然语言理解方法,它不但用于文字、数据,也用来对有限范围内的技术资料,如对农药、化肥等使用说明书进行理解,获取知识。

第四种是知识编译方法,该方法不但使用某种语言规范(语法、语义、词汇表)来理解语言,而且用一个内在的知识模型来获取语言中所包含的知识,把它转换成内部形式,直至最后组织成知识库。

2. 知识转换

知识转换是指把知识由一种表示形式转换为另一种表示形式。

人类专家或科技文献中的知识通常是用自然语言、图形、表格等形式表示的,而知识库中的知识是用计算机能够识别、运用的形式表示的,两者之间较大的差别。为了把从专家及有关文献中抽取出来的知识送入知识库供求解问题使用,需要进行知识表示形式的转换。知识转换一般分两步进行:第一步是把从专家及文献资料抽取的知识转换为某种知识表示模式,如产生式规则、框架等;第二步是把该模式表示的知识转换为系统可直接利用的内部形式。前一步工作通常由知识工程师完成,后一步工作一般通过输入及编译实现。

3. 知识输入

知识输入是把用适当的知识表示模式表示的知识经过编辑、编译送入知识库的过程。目前,知识的输入是通过两种途径实现:一种是利用计算机系统提供的编辑软件;另一种是用专门编制的知识编辑系统,称之为知识编辑器。前一种的优点是简单,可直接拿来使用,减少了编制专门的知识编辑器的工作。后一种的优点是专门的知识编辑器,可根据实际需要实现相应的功能,使其具有更强的针对性和适用性,更加符合知识输入的需要。

4. 知识整理

知识整理包括以下几个方面的内容。

(1)知识的归类集中。书本知识结构与知识库结构大不一样。有些在知识库中成组、块存放的知识,有可能散布于一本书的各处。

(2)知识的联网。为了组成合理的框架体系,需要检查。

(3)知识的补充配齐。

5. 知识检测

知识库的建立是通过对知识进行抽取、转换、输入等环节实现的,任何环节上的失误都会造成知识错误,直接影响到专家系统的性能。因此,必须对知识库中的知识进行检测,以便尽早发现并纠正错误。另外,经过抽取转换后的知识,可能存在知识的不一致和不完整性等问题,也需要通过知识检测来发现是否有不一致和不完整,并采取相应的修正措施,使专家系统的知识具有一致性和完整性。

(三)知识获取的方式

按知识获取的自动化程度划分,知识获取可分为非自动知识获取和自动知识获取两种方式。

1. 非自动知识获取

在非自动知识获取方式中,知识获取分两步进行。首先由知识工程师从领域专家和

有关技术文献获取知识;然后由知识工程师用某种知识编辑软件输入到知识库中。其主要工作有:

(1)与领域专家进行交流,阅读有关文献、获取专家系统所需要的原始知识。这是一件很费力费时的的工作,知识工程师往往需要从头学习一门新的专业知识。

(2)对获得的原始知识进行分析、整理、归纳,形成用自然语言表述的知识条款,然后交领域专家审查。知识工程师与领域专家可能需要进行多次交流,直至有关的知识条款能完全确定下来。

(3)把最后确定的知识条款用知识表示语言表示出来,通过知识编辑器进行编辑输入。知识编辑器是一种用于知识输入的软件,通常是在建造专家系统时根据实际需要编制的。知识编辑器应具有以下主要功能:把用某种知识表示模式或语言所表示的知识转换成计算机可表示的内部形式,并输入到知识库中;检测输入知识中的语法错误,并报告错误性质与位置,以便进行修正;检测知识的一致性等,报告非一致性的原因,以便知识工程师征询领域专家意见并进行修正。

农业专家系统中的“多媒体玉米生产智能系统 MISMAP”就是利用非自动知识获取方法完成的。

2. 自动知识获取

自动知识获取是指系统自身具有获取知识的能力,它不仅可以直接与领域专家对话,从专家提供的原始信息中“学习”到专家系统所需的知识,而且还能从系统自身的运行实践中总结、归纳出新的知识,发现知识中可能存在的错误,不断自我完善,建立起性能优良、知识完善的知识库。为达到这一目的,自动知识获取至少应具备以下能力:

(1)具备识别语音、文字、图像的能力;

(2)具有理解、分析、归纳的能力;

(3)具有从运行实践中学习的能力。

总之,在自动知识获取系统中,原来需要知识工程师做的工作都由系统来完成,并且还应做更多的工作。自动知识获取是一种理想的知识获取方式,它的实现涉及到人工智能的多个研究领域,例如模糊识别、自然语言理解、机器学习等,对硬件也有更高的要求。

(四)基于模型的知识获取

知识获取不能从空白开始。任何知识获取过程都是以原有的知识为基础来获取新的知识,在这原有的知识中有一部分是核心的,根本的,称为知识模型,有时也叫背景知识。一个到实际单位求职的大学生,首先要填各种各样的表格,表格上有姓名、年龄、性别、学历等各种栏目。这些表格体现的也是一个模型,被人事部门用来获取和管理有关干部的信息。模型的使用是贯穿于所有知识获取过程之中的,只是使用方式不同,有时显示,有时隐含。常用的模型包括:

(1)说明模型。借助软件工程中软件生命周期早期阶段(如需求分析、规格说明)所使用的方法,加以变通,运用到对领域知识的分析中来,形成抽象的知识模型。

(2)领域模型。以某个领域的知识为基础,构成解决该领域中某类问题的公共理论框架。根据这个框架去获取知识,即可生成一个具体的知识库,这里一般指较大的领域,如诊断型、规划型、数据分析型等等。