中国科学 : 信息科学　2010 年　第 40 卷　增刊 : 205 ~ 215
www.scichina.com　info.scichina.com
《中国科学》杂志社
SCIENCE CHINA PRESS

论　文

农业知识工程建设中复杂系统理论与混合计算方法研究

李淼*, 张浩东, 张建, 胡泽林, 高会议

中国科学院合肥物质研究院, 合肥 230031
* 通信作者. E-mail: mli@iim.ac.cn

收稿日期: 2010–05–10; 接受日期: 2010–08–03
国家自然科学基金 (批准号: 30871451) 资助项目

摘要　农业生产是一个复杂的生命系统, 也是一个知识密集型的产业. 本文针对传统农业生产控制和智能农业控制在农业生产这一复杂系统中存在的农业知识工程的“获取”、“表示”、“应用”和“管理”的系统性、效能性、精确性、时效性所表现出的局限性、脆弱性、孤立性、单调性, 以及不能有效地解决农业信息系统所面临的关键性问题 (及时掌控态势、有效控制风险、提高管理效率) 进行分析, 提出了以复杂系统科学理论为指导, 以混合智能计算为基础, 注重开展互动性的语言动力学人机交互与信息抽取、系统性的复杂知识结构设计与模型构造、机理性的领域知识关键因素识别与分析、开放型的农业 (大) 生产模型扩展与改造 4 个理论技术和病虫预报、肥水运筹、模式栽培、风险提示、价格分析、销售匹配 6 个科学实践研究.

关键词
农业知识工程
复杂系统
混合智能计算

1　引言

1984 年, 著名科学家钱学森院士在总结党的十一届三中全会召开以来农业生产蓬勃发展的形式时预计, 到 21 世纪, 将要出现知识密集型的农业, 从而导致整个国家生产体系和生产组织的变革, 并由此形成一个重大课题 [1]. 20 多年过去了, 中国农业的形势确如钱先生预计的那样, 得到了前所未有的大发展, 快速转变 (进入) 成为了以生物技术和信息技术为先导的知识和技术高度密集的产业. 就农业生产体系的要素 (环节) 而言, 从单纯的利用光能、风能、生物质能源 (植物、微生物、动物)、降水等自然资源, 加入了工业 (石油) 产品要素 (如: 化肥、农膜、农药、锄草剂、生根剂、保水剂、温室设施、物理育 (变) 种等) 以及农业生产信息要素. 石元春院士在描绘 80 年代的农业状况时说, 农业生产要素构成已从“土地 + 生产资料 (含技术)”转向“土地 + 技术 + 信息要素”[2].

80 年代初期, 在众多科学家以及科研决策部门的推动下, 开展了以知识工程和软件工程两大理论体系为支撑的专家系统工程研究 [3], 通过专家系统开发平台、基层农业技术人员二次开发两大技术手段, 将各类农业专家系统 (水稻、小麦、烟草、畜牧) 从科研单位的实验台, 推广到了千家万户农民的田间地头. 2003 年在瑞士日内瓦举行的世界信息峰会上, 获得世界信息峰会大奖 (world summit award,

引用格式: 李淼, 张浩东, 张建, 等. 农业知识工程建设中复杂系统理论与混合计算方法研究. 中国科学: 信息科学, 2010, 40: 205–215

WSA). 标志着我国利用信息技术改造传统农业, 为促进农村社会经济发展所做出的巨大努力, 也取得了很大成绩 [4].

进入 90 年代以来, 各项信息技术得到了日新月异的快速发展, 在农业生产科学研究实践过程中, 由于各类信息采集与获取手段的提高, 农业生产过程中各种表象因素的观察结果越来清晰 (分辨率达到了百万级, 尺度从分子级进入了微纳米的基因级别), 越来越多样性 (数据、文本、日志、音频、视频、图像等), 而用常规的微分或差分数理统计方法, 各类以文字 (符号) 为标记的规则、框架方法, 以及以多种智能计算 (人工神经元网络、遗传算法、案例推理) 方法, 来归纳构造各类的农业生产模型, 并以模型“还原”的方式, 来“复述”农业生产的过程方式, 已远远不能有效地解决这些复杂系统所面临的许多关键性问题 (及时掌控态势、有效控制风险、提高管理效率). 同时农业生产经济方式的多样化 (个体劳动、合作组织、经纪人、乡土化小型农业企业), 也推动了农产品的快速更新、快速入市、快速改变方式 (温室、滴灌、地膜、反季、工厂化育秧等), 农产品的价值也快速从“以食为天”, 进入了工业化领域 (工业原料、生物柴油等). 另外, 网络化的信息传输手段, 每天都可以数以万条的信息随时随地涌现在 (送到) 科学家和农民的面前. 知识总量的迅速增长, 使得人们面对的农业生产决策不断进步的过程中, 也变得越来越困惑. 到底种什么最好、用什么方法最好、把握什么时机最好, 甚至吃什么最好也是经常担心的问题. 上述复杂多样的农业生产、社会、经济系统的变化, 也极大地加剧了各类农业系统分析与模型构造的复杂程度 [5].

农业知识工程, 是把人脑的思维和记忆, 转化为计算机可“认识”的、可“计算”的符号. 从计算的角度来看知识, 它也是一个系统, 它揭示的是概念与概念之间, 以及概念的属性与属性之间的关系, 知识体系的广度与深度取决于上述关系的多少. 以领域知识库为核心的农业专家系统及以人工神经元网络、遗传算法、案例推理等一系列智能计算方法为核心的农业生产知识“归纳”、“还原”方法, 属于一种平面化的过程“控制”[5]. 在当前大量农业信息形成“涌现”的情况下, 其农业知识的处理能力 (如模型构造、信息抽取、因子识别、普适计算) 在自适应性、效能性、精确性、时效性等方面都表现出了局限性、脆弱性、孤立性和单调性. 各类信息之间关联关系的提取 (模型构造) 对以数学计算为基础的农业生产模型构造方法进行了挑战. 农业知识工程建设面临着很大的困难, 同时也带来了一系列问题: 离散数据的结构化、农业生产机理的揭示、表象化问题的分析, 甚至许多科学观察的数据因计算方法的限制而无法“拟合”[6], 进而造成了浪费.

同时还看到, 农业生产的复杂性不仅表现在“巨”、“变”、“杂”等方面, 更主要的是其中许多可用来被“其他人”或者“机器”所理解、所使用的领域知识难以以一种清晰的数字化、结构化的形式来描述, 或者说目前还不可能将一些在本质上、机理上原本就无法说清的问题清晰化.

就复杂农业生产自身而言, 光能和生物质能是驱动其生长的系统动力, 这些动力的变化以各种信息反映在植物的表面 (颜色、高度、斑点、状态等); 而对这样一个复杂而又多样的农业生产信息管理体系特别是智能 (能够思考的系统) 而言, 知识是始终贯穿如一的核心驱动力.

因此, 需要在现有可行的解析量化方式理论技术条件下探索一种“两全之策”: 将传统的以数学单元为基础的生长模型所描述的“精度”, 从“确定性”推进到复杂系统的“可能性”[7], 进入到以知识单元为组织结构的知识模型阶段. 这种看似“妥协”的做法, 实际上是希望从更宽的角度、更理性化的角度去面对所处的复杂环境和问题 [8].

因此需要根据农业生产的复杂性, 考虑完善农业知识工程建设的理论方法与技术指导体系. 在“不断探索和改善”的指导原则下, 针对智能农业知识工程建设中存在的问题, 智能型的人机互促交互模式, 在人与机器的共同努力下, 利用混合计算方法, 实现农业知识机理的揭示、知识模型的构造、快速

的普适性利用. 最终建立一个开放的、有序的、网络化的、可供多种需求人员使用的“智能农业研讨厅”[9], 为我国的农业生产提供人机共存互促的农业生产技术服务保障系统, 其体系结构见图 1. 文中的许多想法都是讨论性 (探讨性) 的, 希望能起到“抛砖引玉”的作用.

2 复杂系统理论对农业知识工程建设的理论指导意义

复杂性科学是一门研究复杂系统及其复杂性现象的新兴科学. 90 年代初, 钱学森教授提出了综合集成研讨厅, 将机器体系、专家体系和知识体系有机结合起来而构成的一个高度智能化的人机结合系统 [7]. 系统科学和已有的其他科学不同, 它是从事物的整体与部分、全局与局部以及层次关系的角度来研究客观世界的. 客观世界包括自然、社会和人自身. 能反映事物这个特征最基本的重要概念就是系统, 所以系统也就成为系统科学研究和应用的主要对象. 其行为过程更具有拟人性 (图 2), 其特点是自组织性, 在反复地交互中, 不断地校正 [7], 不断地变得聪敏. 对农业知识工程的建设而言, 希望在与人的交互中, 更好地利用机器的快捷和大容量等优势进行信息抽取、特征识别、计算归纳.

20 世纪 60 年代初, 美国“农业系统工程委员会”曾向政府提出建议: 采用系统工程方法, 把各种专业的人员集中起来进行综合研究, 探讨在军事上和工业上成功的技术方法在农业上应用的可能性, 以促使美国农业获得高产和高效流通. 后来的实践表明, 今天美国居世界首位的发达农业体系的形成, 与系统工程在农业领域的成功应用是分不开的 [10].

我国农业系统工程的发展起步较晚. 虽然一些运筹学方法在农业领域的应用, 可以追溯到 20 世纪 60 年代初期, 如用线性规划方法解决农田作业系统中的人财物分派、农机具合理配备等问题. 但真正作为一个崭新的领域并形成学科力量而大规模开展活动则开始于 20 世纪 80 年代. 1985 年召开了中国系统工程学会农业系统工程委员会成立大会暨第一次学术研讨会, 1987 年创办了学术刊物《农业系统科学与综合研究》, 钱学森先生亲自为《农业系统科学与综合研究》撰写了创刊文章 [1]. 2009 年四川农业科学院召开了第 13 届会议.

由于农业系统工程在计算方法方面, 基本以应用数学 — 运筹学 (OR: operational research) 为基础的, 对系统研制以及使用者都有一定的数学条件要求, 因此在大规模推广应用方面受到了一定的影响. 自 90 年代以来, 随后的人工智能的逐步发展, 农业系统工程的研究在突出“系统”的原则下, 逐步采用了许多人工智能领域中的智能计算方法 (以模糊 (或精确) 为代价, 换取相对解). 农业系统工程与智能农业工程也在相互促进中得到了发展.

人类最主要的两种智能形式是抽象思维和形象思维. 而对于农业这样的复杂巨系统而言, 形象思维与抽象思维是互为交织的. 人们根据田间观察作物的长势, 来判断作物生产所处的阶段以及存在的问题, 属于形象思维; 而我们根据作物的表象, 去做出作物机理性的判断, 提出控制的方法 (如灌水、施肥、打杈等) 属于抽象思维. 虽然在传统农业生产控制 (机理型、自然态) 和智能农业控制 (预测型、主动态) 的理论方法研究中, 各国研究人员都已经获得巨大的进展. 然而各种传统的 (或新颖的) 农业智能模型迄今还不能完全而圆满地对农业生产过程中的思维活动进行解释和模拟. 而复杂系统所研究的开放性、多层次性、涌现性、巨大性能够帮助以“系统”的概念去分析描述农业生产模型、发现寻找控制方法、建立人机共存互促的农业生产技术服务保障系统. 同时我们也要针对农业生产特点, 去关注那些特别需要注意的问题.

1) 科学地描述农业生产过程: 通过开放的观察方法, 引导人们关注农业生产系统与外部环境以及子系统, 如利用土、肥、光等之间存在的能量、信息或物质的交换过程来科学地描述农业生产模型. 同

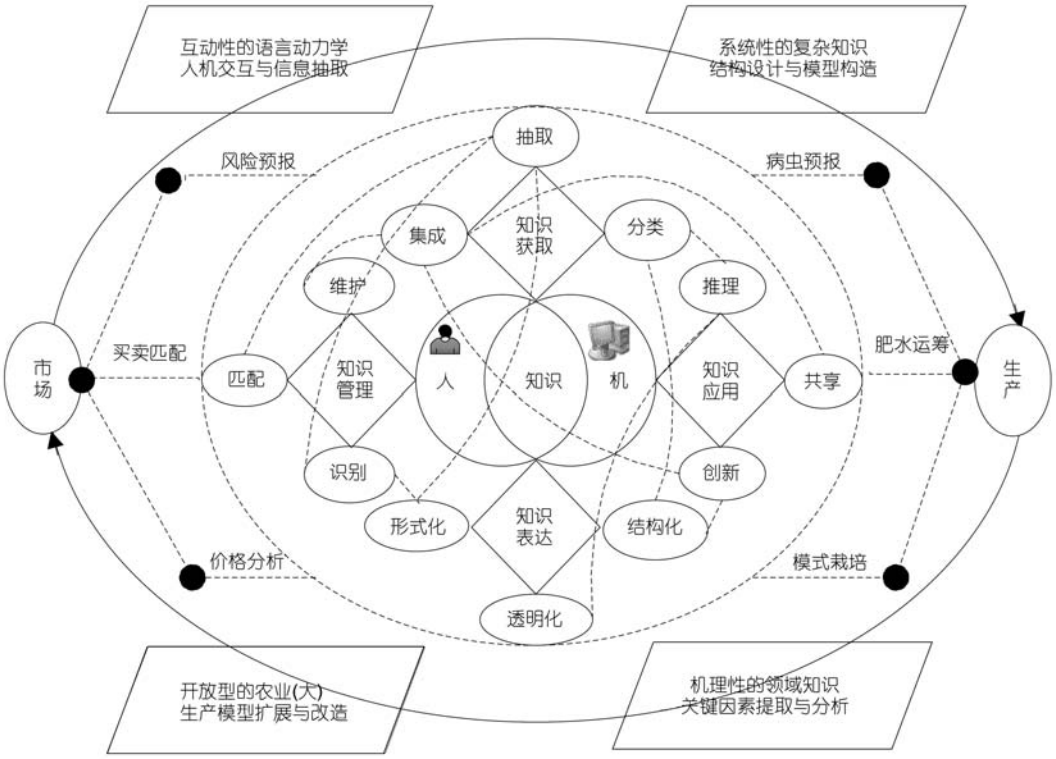


图 1　人机共存互促的农业知识工程体系

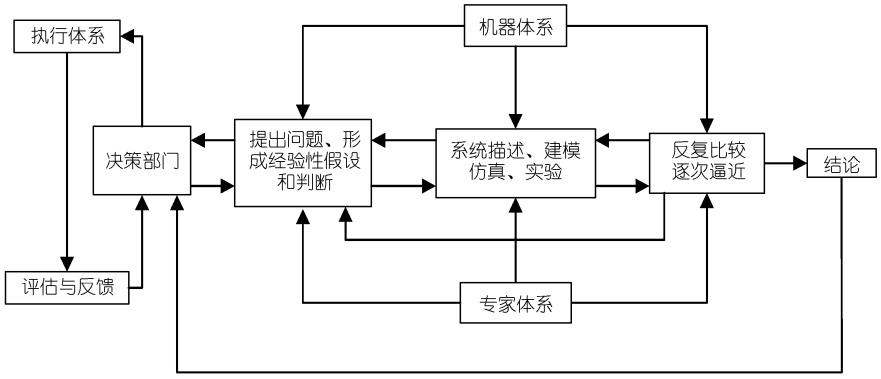


图 2　知识系统工程工作流程

时也要注意到, 农业系统是由环境、作物等不同层次和不同尺度的组分组成的, 描述上述各组分的概念分别由环境 (光、温、水、土、气等) 和植物 (分子、细胞、器官、组织等) 等来确定, 且这些概念一般不能相互套用. 因此, 对于一个复杂智能农业系统来说, 正确区分和界定不同层次的组分并建立相应的概念来刻画它们是至关重要的.

2) 科学地分析农业生产过程: 注意区分与界定复杂系统的组分具有多层次性或多尺度性 [8]. 不同层次或不同尺度的组分不仅有各不相同的结构, 而且有各不相同的特性. 通过多层次性的分析方法,

从已经认识的比较清楚或者容易观察的子系统, 如利用叶面积、叶片颜色、株高等来推断、分析、观测整个系统之间层次、组分以及组成模式或者变异等.

3) 科学地寻找合理地控制方法: 通过涌现性来预测作物生长过程中肩负不同角色的组件间的交互模式、交互规则、系统中子系统或基本单元之间的局部交互以及状态, 以寻找合理地控制方法, 如利用光合作用对作物器官 (分蘖) 的建成、温度对作物的糖份转换等进行预测以及控制.

4) 科学地建立人机共存互促的环境: 在复杂系统中, 基本单元或子系统的数目往往是极其巨大, 从成千上万甚至到数以亿计. 在复杂系统体系中人充当了系统的高级智能组件, 机器充当了记忆、计算、分析与驱动的功能组件, 它们同时互为促进成为系统演化发展的关键因素. 如目前全国有 29700 多个农业资源网站, 其中保存着数以千万兆级的农业实验、观察数据, 同时我国也有数以百万级的农业科研人员和 8 亿农民, 在信息化高度发展的今天, 建立共存互促人 — 机环境, 可以快速推进我国农业现代化的进程.

5) 科学地建立黑箱型的知识模型: 复杂系统的控制并不强调一定要预先知道自组织和涌现现象的行为过程, 而更强调自组织和涌现现象的结果是否满足控制者的要求. 因此知识工程中的各类组分 (模型), 对使用者而言是可以不需要去关心的. 这样可以减低使用者的学习与使用成本. 这方面的一个典型的例子就是作物生长系统 (具有生化反应的复杂系统), 农民主要关心的是作物收获果实质量与产量, 而并不关心肥、水、光、热、土对作物生长过程中的自组织行为和自涌现行为 (二氧化碳光合作用、植物的养分交换、同化物积累与分配、水肥平衡、土壤养分平衡等).

因此, 对这样一个复杂系统的描述、控制就需要借助“系统”. 所以寄希望从技术理论方法和体系上, 进行实用性的探讨. 以系统工程的理论, 去指导农业知识工程的建设, 形成可相互驱动、调配组合的立体化知识工程结构, 见图 3.

3 混合计算对农业知识工程建设的技术指导意义

第 2 节介绍的是农业知识工程建设的指导思想, 那么通过一些什么技术来保障农业知识工程的实现呢?

智能技术自 70 年代广泛进入社会的研究应用阶段以来, 在近 30 年的发展中, 在机器“智能”方面没有取得重要进展. 其重要原因之一就是电脑的工作方式一直停留在它的初始“串行”阶段. 这是它的先天缺陷. 尽管当前的计算机速度最快已经达到 280.6 万亿次/秒, 贮存容量进入了 500 甚至 1000 GB 的实验阶段. 但是电脑与人脑相比结构上有着根本地差别, 人脑不仅约有 $10^{10} \sim 10^{11}$ 个脑细胞 (被称为神经元), 更在于每个神经元约有 $10^3 \sim 10^4$ 个突触, 即一个神经元可通过突触与 8000 个其他神经元交换信息, 当生物电流通过某一突触时, 神经元就将信息传送到下一神经元. 因此人脑有极丰富的联想能力, 可以超越时空, 任意跳跃. 所以计算机的联想、判别与推理能力远不如人脑, 特别是当前采用的人工智能计算方法都是以单项为主体在进行工作的, 它们自身也存在不同的弱点 [11]. 而实际生产中, 特别是农业生产领域, 因其领域的特殊性 (复杂时空变异性、生产条件不可控性、生产过程不可逆性等), 这种单一的处理方法更显得“力不从心”.

针对人工智能研究与具体领域应用中遇到的这些问题, 人们试图开展将多种智能技术以及非智能技术进行综合集成应用的研究. 自 1991 年人工智能领域的著名专家 Minsky(明斯基: 哈佛大学教授, 1969 年“图灵奖”获得者, 获奖原因: 人工智能理论及软件) 提出了“研究不同智能技术组成的人工智能系统的必要性” [11]. 此后, 不同领域的学者从各自的研究目的出发, 开始研究将单个的智能技术同其他的智能技术或非智能术组合研究. 混合智能系统的研究最初从神经网络和专家系统的集成开始,

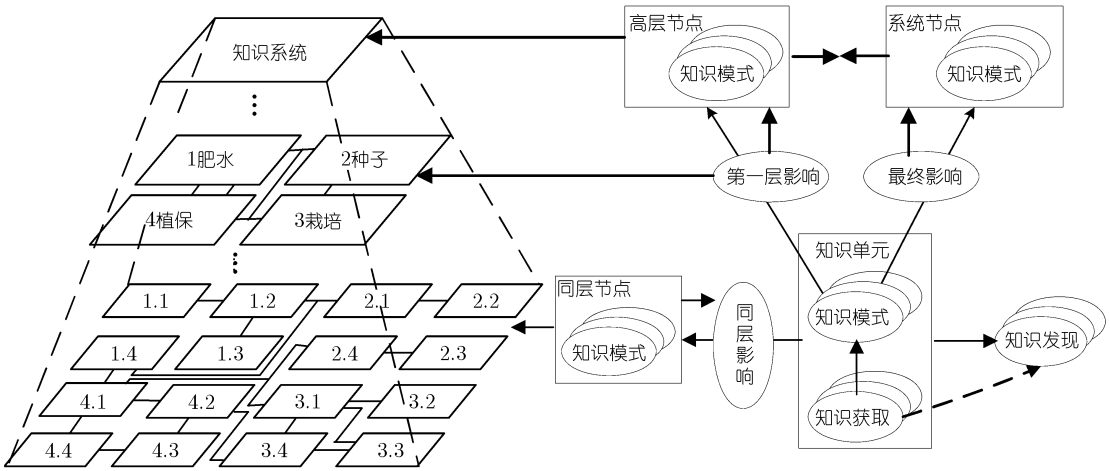


图 3　立体型的农业知识工程结构

到后来遗传算法、进化计算、模糊系统、免疫算法等智能技术, 以及传统“硬计算”技术的不断加入, 随着混合智能系统研究的不断升温, 也吸引了大量的学者投入到混合智能系统的研究中来.

1995 年, 著名学者 Medske 出版了名为“混合智能系统: Hybrid Intelligent systems”一书, 提出了混合智能系统的概念, 并介绍了联合使用模糊系统、神经网络和遗传算法的实例. 2001 年 2 月, IEEE 的 Computer society 在澳大利亚的阿德莱德 (Adelaide, Australia) 首次举办了第 1 届混合智能系统国际会议, 此后每年举办一次, 最近第 9 届会议于 2009 年 8 月在沈阳召开. 除此之外, IEEE, ACM, AAAI 等举办的国际会议也纷纷把混合智能系统作为一个重要的研究领域, 开展专题讨论. 同时, 许多杂志用专刊对混合智能系统的研究进行了支持. 其主要的研究工作, 集中在以下 3 个方面:

1) 在理论研究方面重要集中于: 混合类别、构造方法、系统评价等 3 个方面.

2) 在应用技术研究方面主要集中于: 模糊数学、神经元网络、遗传算法、专家系统. 近些年来, 随着一些新方法的引入, 也使混合智能系统的技术研究范围不断扩展并且进入了新的阶段.

3) 在应用研究方面主要涉及的领域有: 科学管理 (股票预测 [12]、金融决策 [13])、工业控制 (故障诊断 [14]、工业数据分析 [15])、信息系统建模预测 [16].

混合系统在农业领域的专业研究很少, 大部分是采用一种方法的. 本文在中国知网学术数据库上以各种智能技术加相应的农业应用进行 1979~2010 年的论文“全文”检索. 根据检索结果可知, 目前基于神经网络、遗传算法、专家系统、数据挖掘、3S 技术和模型优化决策等的智能技术在农业方面的应用和研究比较多, 在作物生产管理与施肥优化、节水灌溉、品种优化与选育、病虫害管理、温室智能控制与管理、畜牧/水产养殖环境控制与牛奶生产管理、水土保持管理、农产品食品加工、粮食储备与安全、生态环境污染控制、森林火灾控制、市场分析、农机装备智能化与故障诊断等方面都有比较多的研究和应用; 基于案例推理、模糊计算、进化策略、蚁群算法、粒子群算法、免疫算法等的智能技术在农业方面的应用和研究则相对较少, 有的甚至没有; 而基于上述智能技术的混合技术在农业方面的应用与研究, 则只有在农业病虫害预测、施肥预测与信息系统建模等方面有少许应用, 且多采用遗传算法、遗传规划、神经网络和专家系统等四种智能技术进行混合. 这在一定程度上说明智能技术在农业的应用现在还以单项技术的独立研究与应用居多, 见表 1.

农业知识工程不像专业学科那样注重单项技术及方法的成功, 它主要是综合运用和组装已存在的

表 1 1979~2010 年国内开发各类智能农业系统的统计

	作物管理	节水灌溉	品种选育	温室管理	畜牧水产	食品安全	生态环境	市场分析	农机装备
神经网络	5418	16761	7095	2372	4989	5381	10157	17359	3655
遗传算法	19280	7434	2913	894	1851	2091	4101	7923	1614
案例推理	1255	334	215	96	174	138	270	521	508
数据挖掘	16223	3660	2801	522	1513	1790	3152	7635	1210
模糊计算	480	183	67	30	58	35	110	204	150
进化策略	1335	483	226	87	161	194	371	555	330
蚁群算法	1458	493	209	50	116	194	371	555	330
粒子群算法	1029	405	108	36	79	69	179	390	324
免疫算法	652	197	89	34	48	58	153	338	149
专家系统	48722	17538	10022	3008	7063	5581	10536	15144	8818
3S 技术	5123	3111	383	216	1322	1412	5460	8750	3169
模型优化决策	13004	4890	4928	11110	17104	2796	4131	1074	8300
智能混合技术	41	0	0	0	0	0	0	0	0

单项农业科技成果, 以发挥其整体效益. 例如要获得粮食的高产稳产, 就需要把作物区划、种植制度、优良品种、栽培技术、病虫防治、经营管理、收获加工、运输储藏等各个专业环节协调组成一个整体, 才能达到目标. 否则, 即使其中某些环节优化了, 整体效果也不一定最佳, 而各专业学科往往强调的正是单项环节的优化问题, 有时甚至过分强调了某些环节, 以至于造成整体效益并不是最好.

通过使用混合智能技术的方法, 来解决农业知识工程的建设, 可以抛弃使用单独智能技术的智能系统的缺陷, 以期“克服单个技术的缺点, 而综合运用各种技术优势所得到的集合体”, 见图 4. 复杂智能系统就是在这样的背景下被提出的, 并被人们所关注的. 目前混合智能农业技术的研究仍在发展之中, 该领域的具体问题的解决还需要进一步地开发和研究.

4 农业知识工程在智能农业系统的重要作用

农业知识工程的研究与人工智能的研究密切相关. 1970 年美国农业部农业研究署 Baker 和 Lambert 等领导了各领域 20 多位科学家组成的小组在 SIMCOTI(Ducal, 1972) 和 SIMCOT Ⅱ (Jones, 1974) 的基础上, 花费百万巨资和 10 多年时间, 于 1983 年研制成的第 3 代棉花模拟模型 (GOSSYM: gossypium simulation model), 该系统进一步完善了棉花系统动力过程的模拟 [17]. 1984 年 GOSSYM 模型开始推广应用. 1985 年研制成解释 GOSSYM 的专家系统 COMAX(Lemmon, 1985), 诞生了著名的 GOSSYM/COMAX 棉花专家管理计算机模拟系统. 随后日本、埃及、欧洲等国家陆续开展了农业专家系统的研究与应用工作. 新疆农业大学在国家“九五”国家科技攻关项目资助下, 曾于 1996 ~ 1998 年在北疆 3 个试点进行实验 [18].

COMAX 系统以 GOSSYM 棉花生长模型为基础. 而 GOSSYM 本质上是棉花生长系统动力学模型 (表达植物根际土壤中水分和氮素与植株内碳和氮的物质平衡模型). 由于使用该系统时需要输入土壤养分、水分以及太阳辐射、昼夜温差等参数, 作为棉花生长系统中的驱动变量. 所以该系统在我国的推广应用存在一定的困难 (没有充分的数据支持). 同时它也不是严格意义上的专家系统, 而是一个计算机模拟系统.

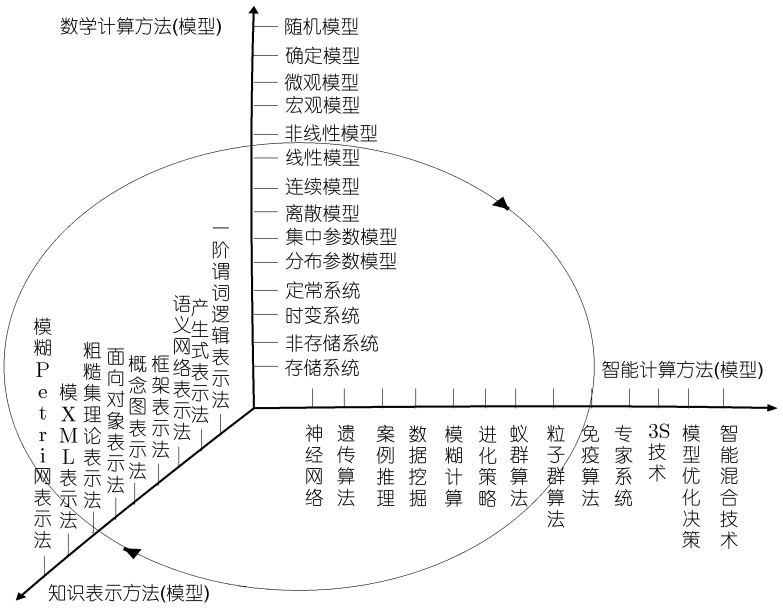


图 4　混合型的农业知识工程构造模型

COMAX 系统对我国农业知识工程的影响很大. 但是如上所述, 它在我国使用很困难. 中国农业科学院计算中心引进了美国的 CERES(也是植物动力学生长模型), 进行了汉化以及输入因子的简化工作, 并应用于水稻、玉米、小麦等多种作物的模拟生长过程.

1990 年, 成立了全国农业知识工程学会. 1992 年在中国召开了农业专家系统研讨会. 2000 年第 1 届“智能化农业信息技术国际学术会议 (ISIITA)”在北京召开, 至今已经举办了 5 届.

1991 年国家基金委设立了“农业知识工程的应用基础研究”、“大型智能系统与工具中多种技术集成的研究”、“基于数据库与知识库的知识发现及其农业应用系统的研究”等研究课题. 启动了我国农业知识工程的前期基础性研究工作. 1998 年国家科技设立了“智能化农业信息技术示范工程”, 支持了 5 个“农业专家开发平台”, 相继在安徽、北京、云南等地设立了 23 个省级示范区, 集中了我国 200 多位高级专家、近万名农技人员参与了示范区建设和农业专家系统的应用推广, 将我国农业知识工程的研究的工作从实验室, 推向了基层农村、田间地头.

以农业知识工程为理论基础的“农业专家系统平台”, 探索了农业计算模型与人工智能的符号模型系统融合的问题, 总结出了“规则 + 规则体”、“框架 + 模式”、“加权模糊逻辑产生式规则”等多种复杂农业知识表示方式, 以计算机“总线”的思想, 将逻辑知识与数理统计知识统一形成为“知识库”系统, 随后又将表格结构的数据库系统纳入知识模型中, 形成了双库、多库协同工作的知识工程系统.

农业知识工程的研究工作, 随着信息技术在我国广泛应用的大形式, 在应用需求的推动下, 取得了很好的成绩, 并且正在向着更深的研究层次挺进. 这些成果集中表现在以下方面: 利用智能计算优良的自组织、自学习和自适应能力, 初步解决了专家系统知识获取的“瓶颈问题”; 利用透明知识解释机制能较好地解释了农业生产各因素之间变化行为; 通过多种数据与知识的融合技术进行知识描述与推理, 明显地改善了专家系统的性能.

5　计划开展的研究工作

通过上述的分析, 本文计划以复杂系统科学理论为指导, 以混合智能计算为基础, 将智能农业系

统“知识工程”中的主角 (人、机) 概念互为融合, 开展相关的理论技术研究以及科学实践研究.

5.1 4 个方面理论技术研究

1) 互动性的语言动力学人机交互与信息抽取: 语言动力学研究的实质是如何通过词计算在语言层次上动态地有效利用信息, 解决复杂系统的建模、分析、控制和评估问题. 如何建立连接人类的语言知识表示与计算机的数字知识表示的桥梁, 成为下一代智能化人机交互和知识工程体系的理论基础之一. 科学问题举例: 聚类分析方法词化; 文本分类的特征提取方法; 农业领域知识模糊建模技术; 基于网格技术的农业复杂异构数据集成与共享技术.

2) 系统性的复杂知识结构设计与模型构造: 复杂系统所展现的问题空间通常都极其复杂且动态变化, 不确定性很强. 如何将复杂系统中的决策资源进行有效的管理, 建立体系组成子系统间的位置层次和功能映射关联模型, 进而进行体系的总体效能评估, 并不断优化改进现有体系的结构及性能, 从而灵活、快速地适应复杂多变的决策环境, 已经成为复杂系统问题研究的关键内容之一. 科学问题举例: 复杂农业知识模型的分层、分类方法; 复杂农业知识层次间的嵌套处理; 复杂农业知识模型推理技术的协调与控制问题.

3) 机理性的领域知识关键因素提取与分析: 农业生长过程中存在诸多影响因素, 一方面, 作物的阶段发育、光合生产、器官建成、同化物积累与分配等生理生态过程比较复杂; 另外, 外界的环境因素 (气候、土壤等) 也会对作物的生长过程造成很大的影响. 作物生产系统的多因素性和随机性导致其显著的区域性和时空变异性, 从而使得作物生产管理专家难以综合考虑多因子互作、预测作物生长趋势并量化栽培技术措施, 因此, 对农业生产过程中机理性的领域知识的分析和研究是农业知识工程领域最基本的研究内容之一. 科学问题举例: 领域知识知识库模型的构建; 系统数据库的构建; 系统方法库、关键技术库的构建; 农业领域知识关键因素提取与分析方法.

4) 开放型的农业生产模型扩展与改造: 农业生产模型是应用数学模型方法和计算机技术, 分析影响农业生产的主要因素 — 气候、土壤、作物、社会、经济等, 从而对农业生产进行定量研究. 通过模型研究, 可以在较短时间内, 用较少的人力、物力得到可靠而优化的本地化模型. 科学问题举例: 知识模型实例化; 生产知识结构模板化; 农业生产模型的结构优化.

5.2 6 个方面科学实践研究

1) 病虫预报: 病虫害预测预报工作内容中存在数据采集、数据报送与管理、数据处理与预测预报、病虫害预报信息发布这些环节, 其中各环节与信息技术中数据获取、数据传输数据处理和数据应用等技术相对应, 可以说信息技术能够为病虫害监测预警工作提供完善的技术支持. 科学问题举例: 病虫害发生环境数据采集技术; 基于 3S 技术的病虫害空间动态监测技术; 专家预测预报模型集成技术.

2) 肥水运筹: 土壤是作物生长发育所需水、肥、气、热等诸多因素的载体, 在气候因素相似的情况下, 土壤类型、质地和肥力水平就成为影响作物品质的重要因素. 在影响作物品质的诸多人为因素中, 对改善作物品质有明显效果的措施主要有施肥和灌水. 科学问题举例: 农田小气候影响条件下的肥水耦合模型; 调控过程中的模型 (数据) 修复、弥补、挖掘技术.

3) 模式栽培: 对于不同的作物, 都有很多不同的栽培模式, 每一种栽培模式都有其适宜的栽培环境. 根据栽培模式所适用的栽培环境, 指导作物的栽培. 作物类型矩阵为 X, 环境矩阵为 A, 栽培模式矩阵为 M, 则其对应关系为 $X * A \rightarrow M$. 由于作物及栽培环境各异, 造成栽培模式的多样性、复杂性、异构性. 科学问题举例: 利用数据信息网格技术对栽培模式进行有效管理, 结合栽培环境, 指导作

物的种植.

4) 风险提示: 农产品生产风险因子众多, 不同产品既面临有相似的风险, 也存在着其特殊的风险. 生产风险因子识别技术就是要对各类因子进行甄别和筛选, 对其危害性进行分析, 测算出主要风险因子的影响度. 科学问题举例: 生产风险因子早期识别技术; 主要风险因子的判别技术.

5) 价格分析: 农产品价格分析是农业信息科技的重要研究领域, 在农业生产、经营和管理中起着越来越重要的作用. 应用农业信息智能分析技术对农产品价格进行分析, 提高农业生产调控能力、应对风险与突发事件能力、促进生产与消费协调能力等, 是当前农业信息科技领域的重要科学实践问题之一. 科学问题举例: 农产品市场价格短期预测技术; 复杂农产品价格影响因素的筛选和提取技术; 复杂农产品价格模型的构建; 复杂农产品价格模型推理的协调技术.

6) 供需匹配: 以网络化的计算机技术和现代化硬件设备、软件系统为基础的物流管理除了完成传统的采购、进货、运输配送等基本职能外, 需要提供更多的增值服务内容, 网上结算以及市场预测等服务. 其中智能化的配送方案又包括最佳供需匹配、配送日程调度、配送车辆的合理利用等问题的决策. 科学问题举例: 现代农产品供求网络系统的建立; 农产品供应和需求信息的采集、清洗和发布、供需匹配模式的建立及优化.

6 结束语

面对复杂的农业生产系统, 农业知识工程的理论体系还很不完善, 智能农业知识工程建设还存在着许多不确定因素, 需从以下几个方面进一步深入研究: 立足于广义人工智能, 以模拟自然智能为目的, 从思维科学的角度出发, 研究人类认知过程中抽象思维与形象思维在功能、知识表达、思维活动等方面的内在关系; 相互理解和转换的知识表示方法, 以及接近于人类思维模式的知识表示机理; 将已取得的农业知识工程的研究成果向其他智能形式的综合集成中推广, 从而进行更广泛的智能系统集成方法研究.

参考文献

1 钱学森. 创建农业型的知识密集产业 – 农业、林业、草业、海业和沙业. 农业现代化研究, 1984, 5: 6–7

2 石元春. 现代农业. 中国农业科技导报, 2002, 4: 8–10

3 李明树. 以农业专家系统为突破口推进智能化农业信息技术的应用示范工程. 计算机与农业, 2000, 5: 1–2

4 赵春江. 农业智能系统. 科学出版社, 2009. 118–120

5 曹卫星, 朱艳. 作物管理知识模型. 中国农业出版社, 2005. 236–239

6 袁媛, 李淼. 基于智能计算的施肥模型算法研究. 农业工程学报, 2008, 24: 116–119

7 王飞跃, 史帝夫·兰森. 从人工生命到人工社会 — 复杂社会系统研究的现状和展望. 复杂系统与复杂性科学, 2004, 1: 37–39

8 王成红, 王飞跃, 宋苏. 复杂系统研究中几个值得关注的问题. 控制理论与应用, 2005, 22: 605–607

9 于景元, 周晓纪. 综合集成方法与总体设计部. 复杂系统与复杂性科学, 2004, 1: 16–18

10 王人潮, 史舟, 黄敬峰. 论农业信息系统工程的建设. 浙江农业大学学报, 1999, 25: 111–116

11 Minsky M. Logical vs. analogical or symbolic vs. connectionist or neat vs. scruffy. Artif intell MIT Expand Front, 1991, 12: 218–223

12 Aburto L, Weber R. Improved supply chain management based on hybrid demand forecasts. Appl Soft Comput, 2007, 7: 136–144

13 Goonatilake S. Intelligent hybrid systems for financial decision making. In: Proceedings of ACM Symposium on Applied Computing, 1995. 471–476

14 Angeli C. Integrating symbolic and numerical features for fault prediction and diagnosis by an expert system. Expert Syst, 1999, 16: 233–239
15 El-Fergany A A, Yousef M T. Fault diagnosis in power systems - substation level - through hybird artificial neural networks and expert system. In: 16th IEEE/PES Transmission and Distribution Conference and Exposition, 2001. 207–211
16 唐丽珏, 李淼, 张建. 混合 GP-GA 用于信息系统建模预测的研究. 计算机工程与应用, 2004, 40: 44–46
17 Lemmon H, 李位鑫. 棉田管理专家系统. Comax. 北京农业科学, 1991, 5: 29–31
18 赵黎, 钟骏平. 棉花生长模拟模型 GOSSYM 及其在北疆棉区的应用. 新疆农业大学学报, 1999, 22: 46–48

Research on complex system theory and hybrid calculation in construction of agricultural knowledge engineering

LI Miao, ZHANG HaoDong, ZHANG Jian, HU ZeLin & GAO HuiYi

Hefei Institutes of Physical Science,Chinese Academy of Sciences, Hefei 230031*, China*

Abstract Agricultural production is a complex life system, and a kind of knowledge-intensive industry as well. This paper focuses on the analyses of the limitation, vulnerability, isolation and monotony of the orderliness, performance, accuracy and timeliness in the acquisition, expression, application and management of agricultural knowledge engineering of traditional agricultural production control and intelligent agricultural control in the complex system — agricultural production, and analyzes the unsolved critical problems, such as controlling situation timely, controlling risks effectively and improving management efficiency. Guided by complex scientific theory of systems and based on hybrid intelligent calculation, this paper proposes four theories of technology, focusing on the human-computer interaction and information extraction of interactive language dynamics, constructional design and model construction of systematic complex knowledge, recognition and analysis of key factors of mechanical domain knowledge, expansion and transformation of open agricultural (large) production model，and also proposes six researches on scientific practice, including, pest forecast, fertilizer logistics, model cultivation, risk presentment, price analysis and sales matching.

Keywords agricultural knowledge engineering, complex system, hybrid intelligent calculation