

农业专家系统入门

石纯一 李明树 钱跃良 编著



清华大学出版社

<http://www.tup.tsinghua.edu.cn>

农业专家系统入门

石纯一 李明树 钱跃良 编著

清华大学出版社

(京)新登字 158 号

内 容 简 介

本书是一本面向有志于智能化农业信息技术的广大农村科技人员的科普读物。全书共有 4 章。第 1, 2 章介绍计算机、专家系统的基本知识。第 3 章介绍玉米、小麦、水稻专家系统。它们都在我国部分地区使用过。第 4 章介绍为方便地建造专家系统所提供的软件环境, 供开发专家系统或对专家系统做二次开发的科技人员阅读。

本书可作为广大农村科技人员学习操作专家系统的培训教材。

书 名: 农业专家系统入门

作 者: 石纯一 李明树 钱跃良 编著

出版者: 清华大学出版社(北京清华大学学研楼, 邮编 100084)

<http://www.tup.tsinghua.edu.cn>

印刷者: 北京人民文学印刷厂

发行者: 新华书店总店北京发行所

开 本: 850×1168 1/32 印张: 5.5 字数: 136 千字

版 次: 2000 年 4 月第 1 版 2000 年 4 月第 1 次印刷

书 号: ISBN 7-302-03849-X/TP·2248

印 数: 0001~2000

定 价: 16.00 元

序

历史表明,我国农业生产上的每一次重大突破,无一不和科技进步息息相关,依靠科技进步发展农业已成为当今世界农业的大趋势。党的十五届三中全会明确要求我们依靠生物工程、信息技术等高新技术,使我国农业科技和生产实现质的飞跃,逐步建立起农业科技创新体系。全国技术创新大会再一次要求我们加强科技创新,加速科技成果向现实生产力的转化,实现我国工业、农业和第三产业的跨越发展。

如果说生物工程是新的农业科技革命的核心,那么,信息技术是新的农业科技革命的先导。农业信息化是指充分运用信息技术的最新成果,全面实现各类农业信息(包括农业生产、管理、农产品营销、农业科技等信息)及其相关知识的获取、处理、传播和合理利用,加速传统农业的改造,大幅度提高农业生产效率、管理和经营决策水平,促进农业持续、稳定和高速发展的过程。

农业专家系统是我国农业信息化科技工作的突破口和切入点。它是一种智能化的农业信息系统,不仅可以作为农业现代化的载体,传播各类实用的农业知识和高新技术成果,而且拥有高层次、多方面农业专家知识,并能模仿人类的推理过程,对多种知识和多项成果进行综合集成,以形象、直观的方式向使用者提供各种农业问题的决策咨询服务。这种系统对于提高广大农业人员的科技文化素质,实现农业生产的跨越发展具有特别重要的作用。

1991年以来,在国家科技部(原国家科委)的领导下,国家“863计划”智能计算机系统主题专家组就把农业专家系统等农业信息技术列入国家“863计划”的重点课题,给予了重点支持。1996

年以后,专家组开始实施并在北京、云南、安徽、吉林建立了“智能化农业信息技术应用示范工程”的示范区,开展了智能化农业信息技术应用示范,并已取得显著的经济效益和社会效益。近年来,专家组又相继在黑龙江、甘肃、陕西、河北、山东、天津、山西、新疆、湖南、四川等地区启动了农业专家系统的示范项目。目前,在各级地方政府的大力支持下,农业专家系统的推广应用工作正在全国十几个省、市、自治区蓬勃展开。

为了对从事智能化农业信息技术的技术人员进行必要的计算机与专家系统的知识培训和技术指导,我们组织编写了《农业专家系统入门》这本书,面向有志于智能化农业信息技术的广大农技人员。该书属科普读物,内容由浅入深,深入浅出,且强调实用性。内容主要包括计算机的基础知识、农业专家系统及其开发平台的基本原理以及农业专家系统的应用实例,也总结了国家“863 计划”在开展智能化农业信息技术中所取得的经验。愿本书能在智能化农业信息技术的技术培训工作中发挥应有的作用。

高文 吴泉源

1999 年 12 月

前 言

这本小册子是在国家“863 计划”智能计算机系统主题专家组的倡导、支持下完成的，是一本面向有志于智能化农业信息技术的广大农技人员的科普读物。

专家系统是个计算机程序，是在有关应用领域的大量知识和数据的支撑下来完成问题求解的。如小麦专家系统是在收集、整理了小麦种植能手的经验知识以及小麦专家知识的基础上建造的计算机程序，进而用于指导小麦的种植过程。专家系统是人工智能学科的一个分支，善于知识处理而非数值计算，靠人们提供的启发式经验知识来求得问题的满意解。

本书共分 4 章，第 1、2 章介绍计算机、专家系统的基本知识，希望通过这部分的学习，读者能了解、掌握计算机的组成和运行过程以及专家系统是怎么回事。第 3 章介绍的玉米、水稻、小麦专家系统是在我国部分地区已使用的专家系统。分别由承担“863 计划”农业专家系统课题的吉林大学刘大有教授、北京市农林科学院赵春江研究员和合肥智能机械研究所李森高级工程师提供的，经编者整理写成。第 4 章介绍的是为方便地建造专家系统所提供的软件环境，称之为专家系统开发平台。对开发专家系统，或对专家系统做二次开发的农技人员值得一读。这部分是由吉林大学提供的。

在编写过程中，还得到了不少同志的帮助，一并表示谢意。

由于编者水平所限，又没有写过科普读物，不当之处，请批评指正。

石统一

于清华园

1999.12

目 录

前言	V
第 1 章 电脑入门	1
1.1 电脑的组成	1
1.2 二进制表示	3
1.3 门电路和布尔代数	8
1.4 2+3 是怎么做的	13
1.5 触发器和寄存器	16
1.6 译码器和计数器	20
1.7 控制器与时间的控制	24
1.8 一条指令的执行	27
1.9 软件系统	32
1.10 电脑应用举例	49
第 2 章 专家系统	58
2.1 概述	58
2.2 产生式系统	61
2.3 不确定推理	63
2.4 专家系统的开发过程	66
2.5 专家系统开发工具	67
第 3 章 玉米、水稻、小麦专家系统	69
3.1 玉米专家系统	71
3.2 水稻专家系统	90

3.3 小麦专家系统	124
第4章 专家系统开发平台	144
4.1 背景知识	145
4.2 平台的结构	149
4.3 构件的功能	151
4.4 平台的构件	153
4.5 平台的知识表示和推理	159
4.6 用平台建造农业专家系统	164

第1章 电脑入门

介绍专家系统以及专家系统是怎么帮助人们种田的是本书的目的。为此需要先学点电脑（计算机）的基本知识。专家系统是知识工程师设计的一个电脑程序，如某块田地的肥力、墒情、气候等条件已知，问何时种植小麦，播多少种以及预估产量是多少，这个程序依照农业专家预先整理好的知识，便可有序地一步一步运行，回答出这些问题。细想，这已涉及到什么是程序，程序的指令以及一条指令是怎么执行的等一系列电脑基本知识。这正是本章要介绍的内容。学习了这一章，你将会知道电脑怎么进行 $2+3$ 的计算；怎么把“中国”两个字输入电脑，又输出打印在纸上；电脑计算的速度有多快，是否还需更快的计算；存储数据或指令的能力有多大；电脑都会干些什么等电脑初学者有兴趣的问题。

1.1 电脑的组成

讲解电脑和介绍其它知识一样，要先将电脑分析一下，看它都由哪些部件组成，各部件的功能以及部件间的关系，然后再分析它是如何通过这些部件的协调有序的工作来完成问题求解的。电脑由处理器、存储器和输入输出部件组成，如图 1.1 所示。简单地说，当求解问题时，先将有关数据、程序通过键盘输入，由人逐一敲键暂存入处理器（记为 CPU），并在 CPU 的控制指挥下，写入存储器的指定地址或单元之中，进而启动已输入的程序，来对数据做处理，并将结果写入存储器。当需要显示处理结果时，仍需通过程序从存储器中调出有关结果，并命令输出打印机或显示

器给出结果。

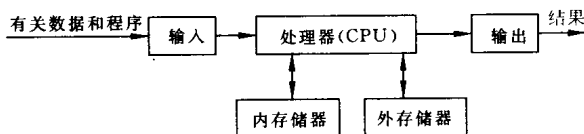


图 1.1 组成电脑的部件及相互关系

可见，存储器是被动的，在 CPU 的指挥下快速地完成存入或取出数据或指令（统称信息），而存储器存入的信息是长期记忆的，直到存储器的同一单元受命重新写入新的信息为止。另外，存储量大的软盘、硬盘也是存储器，称为外部存储器。它们与 CPU 相连，通过程序将有关信息调入内存储器后才能直接使用。现在回答 $2+3$ 是怎么做的。那是在先将数据 2, 3 以及计算 $2+3$ 的程序 A 预先输入、存于存储器的条件下，再启动 A 来实现的。这时程序 A 的运行是先从存储器取出 2（一条取数指令），再取出 3 同时与 2 相加（一条加法指令），相加结果得 5，送回存储器（一条送数指令），最后打印输出结果 5（一条打印指令）。其中 2 与 3 相加的过程是由处理器的运算部件完成的。由此可看出 CPU 是电脑的中心部件，存储器、输入输出部件都与它直接相连，并通过它实现存储器、输入、输出部件间的信息传递。

CPU 是由几个芯片（半导体集成电路）做的，输入键盘、输出打印机等设备通常是摆在桌面上的。而不容易想清楚的是 CPU 和存储器，其实存储器也是由大量芯片组成的。这里讲的电脑组成是指硬件说的，硬件是指那些看得见摸得着的元器件、部件等设备。我们买一台微型电脑时，首先要选的是 CPU 处理速度，每秒可做多少次加法运算。其次，要选存储器是 64K 还是多少 M（1K 指 1024 个存储单元（8 位二进制），1M 指 1024K），这是对硬件设备的要求。另外，还需考虑的是配备哪些软件。软件是各种

类型的系统程序和某些应用程序。没有较完备系统软件的电脑是难于完成一般问题求解的，更不用说复杂问题了。只有配备了较齐全的软件，电脑才能“活”起来，才能方便于人们的使用（详见1.9）。

1.2 二进制表示

1.2.1 二进制与十进制

电脑的功能是通过执行程序来处理数据的，而数据和程序的每条指令在电脑内部都是由二进制数来表达的。为什么不用人们熟悉的十进制？道理比较简单，就是二进制的电脑硬件容易用电路制造出来。也可以说电脑的硬件用二进制表示比用十进制表示更为节省。我们日常生活熟悉的是十进制数，计算起来处理进位方便。十进制数的特点是每位只见0, 1, ..., 9, 不见10, 而是逢10进位。多位十进制数如256可以看作是 $2 \times 10^2 + 5 \times 10^1 + 6 \times 10^0$ ，是以 $(10^2, 10^1, 10^0)$ 的结构来表示的。相应二进制数表示是每位只见0和1, 不见2, 是逢2进位。多位二进制数如101可以看作是 $1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$ ，对应的十进制数为5。

当我们只谈一位数的表示时，若用十进制，就需要一种电路能区别开0, 1, ..., 9, 十种可能的状态。若用二进制，就只需一种能区别开0, 1两种可能状态的电路，这可简单多了。如以电灯亮表示1, 以电灯不亮表示0, 那么，一个简单的类似电灯电路就可以表示一位二进制了，这是说表示。再说计算，为实现两个一位十进制数的相加，就更不简单了。仍用电路来完成这加法的话，要求这个电路知道 $0+0=?$, $0+1=?$ 直到 $9+9=?$ 例如要知道 $7+8$ 是15, $5+6$ 是11, ..., $9+9$ 是18。这如同小孩子初学乘法时要背会的九九歌一样，要让电脑完成，就不甚容易了。然而

对两个一位二进制数相加，用相应的电路来表示，就容易多了。这时，对电路的要求就降低了好多！只需会 $0+0=?$ ， $0+1=?$ ， $1+0=?$ 和 $1+1=?$ 就够了。当然，你也会问，对同一个数来说，用二进制表示要比十进制表示位数来的长？是的，例如十进制一位数 9 可用二进制四位数 1001 描述，意思说，表达同一数，两者从长度看顶多差 4 倍。但总的花费算起来，还是二进制表示划算。下面的问题是电脑用二进制表示信息，在进行计算时，并不要求人们自己将日常的十进制用语也改成二进制，怎么办？这个改制的任务要由电脑本身通过程序来完成。即若是十进制数输入，要先经电脑的程序化成二进制（十化二）再去执行计算；若要求计算结果给人观看时，再由电脑程序来个二化十（二进制转换成十进制）。而十化二，二化十，对电脑讲是很简单的计算程序，做起来也很快，人是感觉不到花了时间的。这么说来，电脑用二进制就是合理的选择了，而且电脑硬件确实是这么做成的。

十进制数与二进制数的对照：

9——1001	4——0100
8——1000	3——0011
7——0111	2——0010
6——0110	1——0001
5——0101	0——0000

具体地说，十进制→二进制，二进制→十进制的转换，若由人手工计算如下（实际上由电脑完成，不需人来做）。

二进制→十进制，如

$$\begin{aligned}
 10101101 &= 1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 \\
 &\quad + 0 \times 2^1 + 1 \times 2^0 \\
 &= 128 + 32 + 8 + 4 + 1 \\
 &= 173
 \end{aligned}$$

即二进制 10101101 表示十进制数 173。

十进制 $\rightarrow$ 二进制, 如 $a=173$, 先用 2 的幂次, 如 $(2^7, 2^8)$ 区间来试探, 看 173 是否落在这个范围中。由于 $2^7 < 173 < 2^8$, 所以有二进制表示 $0 \times 2^8 + 1 \times 2^7 + \dots$, 进而做 $173 - 128 = 45$, 再对 45 试探, $(2^7, 2^6)$, $(2^6, 2^5)$, $\dots$, 这时 $2^5 < 45 < 2^6$, 便得表示 $1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + \dots$, 如此重复这个过程, 便得 $173 = 1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$, 即 173 的二进制表示为 10101101。

当然电脑有标准的计算程序进行十进制 $\rightarrow$ 二进制的转换, 不会像人那样试探地做。电脑表示数和指令常以八位二进制为单位。一个八位二进制称为一个字节 (Byte, 简记为 B)。我们说机器的存储器是由众多的单元 (地址) 组成, 每个单元 (或多个单元) 记忆存储一个数或一条指令或有关中间信息。而对存储器来说, 地址的号码也是用二进制编写的。另外, 电脑内部将 0, $\dots$, 9 表示成什么样的八位二进制码, 这另有约定。通常使用的 ASCII 码, 规定

0——x0110000

1——x0110001

...

9——x0111001

其中后 4 位就是 0, $\dots$, 9 的二进制表示, 而 “011” 三位表示的是数字, 以区别于英文字母、数学符号的表示。至于数字位 x, 可规定为 0, 也可规定为 1。或赋予其它含义均可。

1.2.2 二进制数表示的范围

还需指出:

对一位二进制数来说, 只能表示 0, 1 两种状态, 对应十进制数 0, 1。

对二位二进制数来说, 能表示 00, 01, 10, 11, 共计 4 种状

态, 对应十进制数 0, 1, 2, 3。

对三位二进制数来说, 能表示 000, 001, ..., 111, 共计 8 种状态, 对应十进制数 0, 1, 2, ..., 7。

对八位二进制数来说, 能表示 $2^8=256$ 种状态, 对应十进制数 0, 1, ..., 255。

对十位二进制数来说, 能表示 $2^{10}=1024$ 种状态, 对应十进制 0, 1, ..., 1023。

为了方便, 将 1024 用 1K 表示 (不是通常的 1000), 说一个存储器的容量是 64KB 就是说它有 64×1024 个存储单元, 每个存储单元是一个字节 B。说存储器是 64MB, 即有 $64 \times 1024 \times 1024$ 个存储单元。如要计算 $127+105$, 这是通过两个八位的二进制寄存器, 便可分别寄存 127 和 105, 进而通过加法器的计算得和 232, 仍可由八位二进制寄存器寄存。显然, 对八位寄存器来说, 凡是大于 255 的数, 不管是初始数据, 中间结果还是最后结果都是不允许的。若要寄存, 则需扩大寄存器位数。微机的 CPU 就是针对八位二进制表示信息的处理, 超过八位数的是通过程序分段计数, 将多个八位的计算结果连接起来, 求得最后结果。

为表示数 $3.1416, 2 \times 10^{33}$, 人们输入机器前需先写成

$$3.1416 = 0.31416 \times 10^1$$

$$2 \times 10^{33} = 0.2 \times 10^{34}$$

其中的尾数 (小数) 部分第一位要求不为 0, 直接由机器表示。而 10 的幂次 (称为阶数) 另外加以表示。这种做法表示数的范围扩大多了, 而且这种称作浮点 (小数点可浮动) 表示下的数的运算也是容易的。可想象两个浮点数的加、减, 要先对阶 (只有阶相同才能相加), 和的尾数是两尾数相加的结果, 和的阶码还需加以相应的调整。两数相乘先尾数相乘, 阶码也需加以相应调整。当然浮点表示和运算也是在化为二进制后进行的。另外还需要给出数的符号。可简单的以二进制数如 0 表示正, 1 表示负, 值得注意

的是，需要考虑到加法运算后的符号。

1.2.3 补码表示

引入补码，是为了将减法化为加法来做。对十进制而言，做 $6-2$ 可将正数 6 不动，与 8 ($8+2=10$) 相加（对十进制，称 8 是 2 的补码），有 $6+8=14$ ，若略去进位 10，得 4，这便是 $6-2$ 的结果。这样便将相减变为相加了。看一下 12 进制的手表，若现在是 7 点整，问 3 小时前是几点，这自然是逆时针退 3 小时，做 $7-3$ ，求得 4，若不然仍可顺时针做加法，因为 $7+9=12+4$ ，略去钟表的进位数 12，便得 4。相对十二进制来说，如 3 的补码 ($3+9=12$) 是 9，减 3 可用加其补码 9 来代替。同样的道理，对二进制来说，若有一个四位二进制加法器，这时是十六进制， $6-2$ ，也可由 6 加 2 的补码来完成。

$$6=0110$$

$$2=0010$$

先将减数 2 化成补码， $2+14=16$ ，而 $14=1110$ ，这样， $0110-0010=0110+1110=0100$ （略去进位） $=4$ 。有什么简单方法，从 0010 怎么直接求得其补码 1110 呢？这有一个规律，即将 0010 按位求反 ($1\rightarrow 0, 0\rightarrow 1$) 得 1101，然后再加 1 略去进位，便得 1110。

这里说的是负数求补码，而正数的补码规定为正数本身。

在两个二进制数求和的过程中，需要考虑的问题还有数的符号处理，以及补码相加后为负数时需将结果再求一次补码，返回原码的过程。

例 1 $X=+0.1010$ ($+10$)， $Y=-0.0101$ (-5)，求 $X+Y=?$

1. 先将 X, Y 的符号以 0 表示正，1 表负，改写为 $X=01010$ ， $Y=10101$ 。

2. 进而将被加数、加数求补码，得 $X(\text{补})=01010, Y(\text{补})=$

11011(符号位不动)。

3. 计算 $X(\text{补}) + Y(\text{补}) = 00101$, 已丢弃最左边一位的进位。

4. 对这个结果再求一次补码, 因为是正数, 仍是 00101, 便得 $X + Y = 00101(+5)$ 。

例 2 $X = +0101(+5), Y = -1010(-10)$, 求 $X + Y = ?$

1. 标明符号 $X = 00101, Y = 11010$ 。

2. 求补 $X(\text{补}) = 00101, Y(\text{补}) = 10110$ 。

3. 计算 $X(\text{补}) + Y(\text{补}) = 11011$ 。

4. 再求补得 $X + Y = 10101(-5)$ 。

例 3 $X = +1010(+10), Y = -0101(-5)$, 求 $X - Y = ?$

1. 因为减数为负数, 便知 $X - Y = X + (-Y) = +1010 + (+0101)$ 。

2. 求补, 因为正数求补为原数, $X(\text{补}) = 01010, Y(\text{补}) = 00101$

3. 求出 $X(\text{补}) + Y(\text{补}) = 01111$ 。

4. 再求补得 $X - Y = 01111(+15)$ 。

例 4 $X = -0101(-5), Y = +1010$, 求 $X - Y = ?$

1. 先判断 Y 为正数, 便知 $X - Y = X + (-Y)$, 对 X, Y 加上符号, $X = 10101, Y = 11010$ 。

2. 求 $X(\text{补}) = 11011, Y(\text{补}) = 10110$ 。

3. 计算 $X(\text{补}) + Y(\text{补}) = 10001$, 已丢弃左边最高位的进位。

4. 再求补, 得 $X - Y = 11111(-15)$ 。

1.3 门电路和布尔代数

电脑中信息的表示采用了二进制, 从而像 $2+3$ 的简单计算或更复杂的信息处理, 都是通过一个二进制表示的信息本身或多个二进制表示的信息间多次或大量的“运算”来实现的。说得更清

楚一些,对一位二进制数来说,运算结果无非是1转换成0,或0转换成1,或者是保持不变。在电脑中的信息运算除引入二进制的加法($0+0=0$, $0+1=1$, $1+0=1$, $1+1=0$)外,还需引入体现日常生活中常使用的一些连接词,其中有“与”,“或”,“非”(否定)等。如在一个房间里,有灯A、灯B,描述这两个灯的亮和不亮的状态,常使用灯A“与”灯B都亮了,灯A“或”灯B亮了,灯A没亮(含义是A亮的否定)。

我们称体现连接词“与”、“或”、“非”的运算为逻辑运算。令我们难以想象的是,只有加法和逻辑运算功能的电脑已足够用来解决复杂的信息计算和信息处理了。为表示一个数或多个数间的运算规律,在初中学习了初等代数,这里为对二进制信息表示和计算也使用一种代数工具,叫逻辑代数或布尔代数。若用A,B分别表示一位二进制数,则规定:

算术加法

表 1.1 算 术 加 法

A	B	和 (H)	进位 (C)
1	1	0	1
1	0	1	0
0	1	1	0
0	0	0	0

逻辑运算

表 1.2 逻辑运算

A	B	A 与 B	A 或 B	A 的否定
1	1	1	1	0
1	0	0	1	0
0	1	0	1	1
0	0	0	0	1