



Palaeoworld

This is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship.

PALAEOWORLD Editorial Office

State Key Laboratory of Palaeobiology and Stratigraphy

Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences

Beijingdonglu 39, 210008 Nanjing, PR China

e-mail: palaeoworld@nigpas.ac.cn

PALAEOWORLD online submission:

<http://ees.elsevier.com/palwor/>

PALAEOWORLD full-text (Volume 15 –) available at:

<http://www.sciencedirect.com/science/journal/1871174X>

用 FoxBASE+ 建造化石孢粉专家系统

陈民敏 孙志亮

(地质矿产部石油地质中心实验室)

宋之琛 黄 嫔

(中国科学院南京地质古生物研究所)

化石孢粉专家系统是一个具有大量化石孢粉专家知识和工作经验的计算机程序系统,因此,程序设计是系统的一个十分重要的组成部份。过去人们在软件系统的应用上,多选用 LISP 和 PROLOG 语言,这类语言具有较强的推理功能,但是,这类语言需要很大的内存,处理速度慢,并缺少良好的用户界面。随着专家系统的实际应用日益广泛,将面临着海量的数据管理和复杂的数据处理,就显得应付不了。于是人们探索将具有很强的数据管理功能的数据库管理系统和专家系统结合起来,把事实、规则的管理由数据库来完成,而在 AI 和 DB 之间建立接口,使 AI 和 DB 能互相传递数据。比如,有人以 PROLOG 为专家系统的宿主语言和 dBASE 数据库的连接方法,或以 dBASE 存贮 PROLOG 的事实,或以 PROLOG 直接读 dBASE 的数据库文件,或以 PROLOG+dBASE 构成演绎数据库等。

通过几年的探索和工作积累,我们采用 FoxBASE+ 的 DB 软件为核心,程序逻辑为辅的方法,较好地实现了化石孢粉专家系统。系统软件研制基本上是利用了 FoxBASE+ 软件提供的强有力的数据管理能力,编程能力,及先进的人机会话(即人机联作)功能,并利用 FoxBASE+ 运行速度快,会话管理菜单设计方便等特点,使用 RDB 技术及 FoxBASE+ 编程技巧,较圆满地实现了系统的查询,匹配,推理智能等活动。

现将我们用 FoxBASE+ 建造化石孢粉专家系统的体会及做法叙述如下。

一、系统最常用到的几个 FoxBASE+ 命令

系统最常用到的几个 FoxBASE+ 命令有:CREATE, LOCATE, CONTINUE, FIND, EDIT, APPEND, DISPLAY, LIST 等。

对于上述各条命令的格式和功能等的具体说明,请见有关教科书,本文不再重复。

二、建 库

关系型数据库系统的数据库 RDB 和专家系统的知识库 KB,尽管对数据(广义的)——包括数值数据和非数值数据——都有存贮、管理的功能,但两者毕竟不同,有人列表 1 表示它们之间的差异。

根据上述可以看出,数据库管理系统虽然处理大量商业数据和科研资料的能力很强,但对于专家系统中的基于规则的信息,如推理、解释、学习、联想、问题求解等,一般的数据库却难以

• 中国科学院南京地质古生物研究所现代古生物学和地层学开放研究实验室资助课题(893110)。

直接完成。为了使商业性的数据库系统能适应专家系统的推理需要,我们在数据库的基础上,增加了推理功能,把知识形式化、抽象化和数字化而组织到数据库中,从而构造了具备演绎推理功能的数据库,称之为演绎数据库(DDB),也称之为智能数据库。

我们在用 FoxBASE+ 建立知识库的过程中做了如下的改进和开发。

第一项改进是在库结构中增设了“权”字段。对化石属种各项特征在鉴定中不同的重要程度加以形式化和数量化。这个权值可以用数理统计方法自动给出,也可以由化石孢粉专家根据自己的经验人为地给出,我们是采用后者给出。这样的数据库就得到了强化,具备了动态的主动性特征。库里的知识就不仅仅作为一种数据信息被查询,而是具有了智能性。

表 I 数据库与知识库差异

数 据 库	知 识 库
存放数量大	存放规则量较小
数据类型较简单	数据类型复杂
规则性	不规则性
一般人员可以更新	专家更新
90%规则的存取	即席查询
模式修改少	经常修改模式
精确	不太精确
表示事实	表示规则
推导出知识	不能推导出数据

表 I 化石孢粉属种描述知识库结构

字段名	类 型	宽 度	小数位数
属种名	C
特征 1	C
权 1	N
⋮

权值的大小不仅表示该项特征对属种的重要程度。在推理过程中,也可以利用权来对推理起导向作用。比如某项特征的权值大于 0.7,而用户回答反对该项特征,这时可以认为用户对整体作出了否定。推理机将自动从头要求用户重新观察,或让 DBA(数据库管理员)在知识库中增加相应的知识。

第二项改进是如何用数据库严格的数据结构来存贮不规范,不精确的化石描述。由于化石属种的建立来自许多不同的作者,他们尽管也遵循若干描述格式,但不同的作者风格迥异,用语不同;有时也是由于化石保存缺陷。因此,化石描述的特征项和描述用语就有了千差万别。为了让 FoxBASE+ 的关系表能存贮这些不规范的知识,我们采用了如下方法:

1). 对描述工作进行规范和提炼,从原始数据中找到一个有规律的描述层次,对每一个从

类—属—种的形态分类系统落在一个树型结构上,使之能建立一离散的关系结构模型,描述上述的关系网络——一个树型的数据结构(插图 1)。

在树型结构中,各层的知识描述利用一个关系型表来存放。同时,表的第一个记录存放了下层表的指针。当用户输入相应的选择后,指针将自动向下移动,并将一个下层表打开,使指针从树顶(根)向终点结点(叶子)一步一步前进,从而获得结果目标,这就是导航式思想在本系统上的实现。

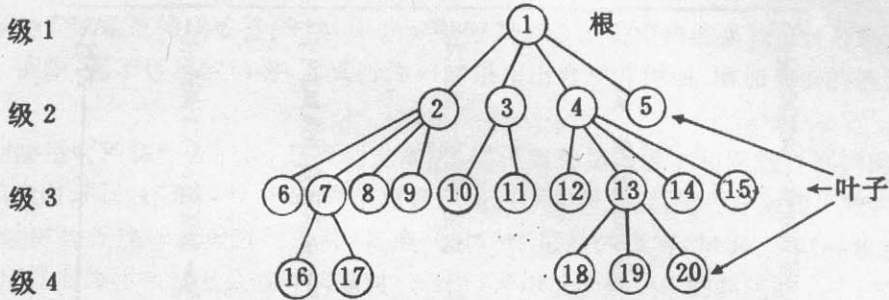


插图 1 树型数据结构模型

由于下层结点具有上层结点所选择后的事实,所以此树型结构具有非常严格的上下文关系和属性继承特性。

现以化石菌孢为例,将具体实现过程说明如下:

1. 建立正向推理知识库

1). 建立层次描述网,并归纳到一个树型结构上(插图 2)。

2). 建立各层描述知识库。

建立基本情况为: A. 描述名称; B. 情况; C. 权; D. 下级表名; E. 目标。

当 D 存在时, E 为空白。反之亦然。

如上述建立表 A, 表 A1, 表 A11, ……

3). 编写推理中的导航程序,利用表 A, 表 A1, 表 A11 等向用户提示,当用户选择后推理程序结合用户选择及指针情况,自动打开下层知识表。

2. 建立反向推理知识库

1). 用 FoxBASE+ 建库命令建立带有权值的知识库。

2). 将化石孢粉属种描述知识(带有权值)输入知识库中。

3). 利用属种名为关键词,将属种描述一个一个地向用户反问,用户回答肯定或否定,以及事实的可信度。

4). 系统作进一步工作。

三、推理机研制

推理功能是专家系统区别于其它应用程序主要特点之一。PROLOG, LISP 语言之所以能为人们选做专家系统的软件,就是因为它们具有较强的推理功能。那么 FoxBASE+ 的程序设计语言是否也能实现专家系统的各项推理、智能功能呢?如查询、模式匹配、回溯、切断等功能。我

们经过实践摸索,运用编程技巧,较好地解决了这些问题。我们用 FoxBASE+ 编写的推理程序也同样地满足了系统推理的需要。下面我们就查询、匹配、回溯、切断等专家系统必须具备的功能,介绍我们是怎样用 FoxBASE+ 实现的具体做法。

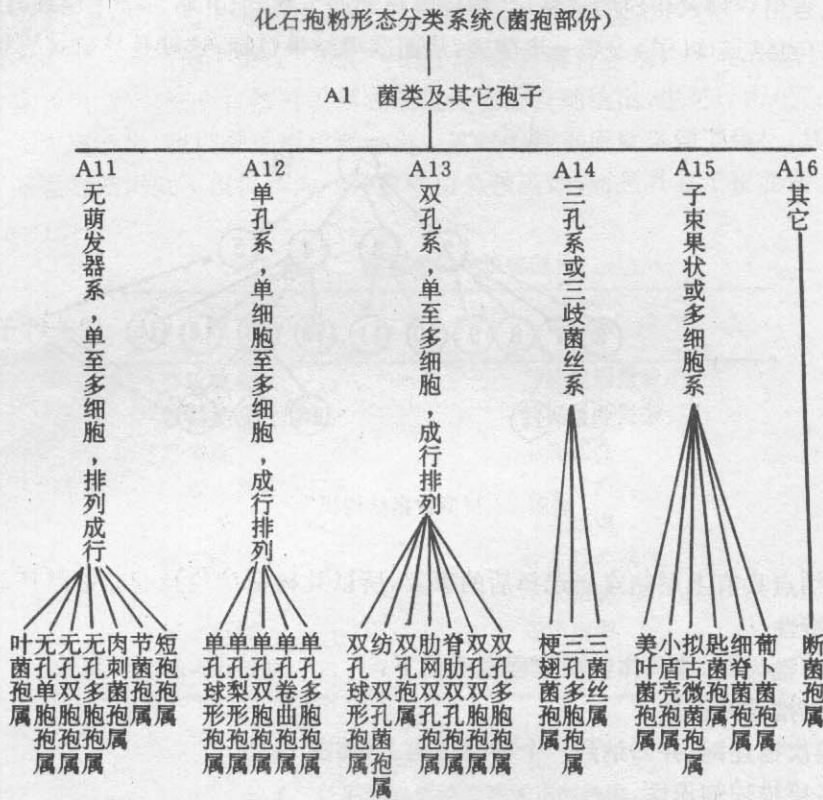


插图 2 化石孢粉形态分类系统树(部分)

1. 查询

系统可提供静态查询,动态查询,特殊查询和模糊查询等 4 种查询功能。简述如下。

1). 静态查询:由推理机向用户提供事实选择菜单(如上述的正向推理菜单),用户对菜单的提示给予肯定或否定的回答,同时当应答是否定时,用户对该事实的认识作进一步应答,推理程序决定程序走向并是否查询到相应结果。

2). 动态查询:由用户提供事实,推理机查找到所有满足用户提供的事实为前提的结果。

3). 特殊查询:根据工作需要,增设相应查询字段,以满足用户的特殊要求查询。

4). 模糊查询:所谓模糊查询即系统依据用户输入将最接近用户查询条件(模糊查询条件)的记录输出给用户。模糊查询采用组合条件查询的输入模式,处理时,对数值型字段取区间值,对字符型字段进行模糊匹配检查,如果输入值中有一定个数的非空字符(如 1/2 个数)可以找到,则认为模糊匹配而输出给用户。

2. 模式匹配

利用 FoxBASE+ 中的 FIND, LOCATE CONTINUE. 等语句中的条件表达式,利用布尔代数中逻辑运算关系,将用户提供的事实和知识库中的知识进行匹配,一旦条件成立,则自动将一个或多个匹配结果提供给本次匹配工作。如果无结果,则利用模糊数学的方法,并根据事实和

知识库中知识的权,动态决定取舍一些事实或是扩大某一事实的作用域。

3. 回溯

当搜索失败,退回上一级子目标的逆向运行的方式,称为回溯。由于系统采用了树型结构描述一个知识,恰恰严格地反映了知识库的上下文关系。当下级知识不能成立时,程序将逐级回溯到上一级知识,要求用户作出回答,如果用户提供了新的选择或认识,程序将自动向下继续搜索。否则,当回溯到最高层时(树根),程序将自动切断本次搜索,而终止程序的执行。

4. 解释接口

解释是专家系统的重要特征之一。所谓解释是指系统对用户或系统设计者所提出的问题给出一个清晰、易于理解的回答,对其行为和结果作出合理的说明,增加系统的透明度和可信度。

我们使用两种解释方法:一是预制文本法,即将属种描述原文和图版分别建两个库,系统在作出结论时,同时打开这两个库,向用户显示描述原文和图版,供用户对照比较。二是执行追踪法,通过对程序执行过程进行追踪,将推理过程中用过的事实,知识、规则存放在一个模块中,当用户提出为什么?怎么做等问题时,系统将调用这个模块中的相应的事实、知识、规则来回答用户的提问。

5. 学习功能

一个专家系统的功能高低,与知识库有很大关系,知识库中的知识越完全,越真实,则专家系统的水平越高,所以知识库的维护是十分重要的,要随时进行一致性检查,知识的扩充、删减与修改。

化石孢粉专家系统知识库的经常性维护工作主要是属种或描述的增加,修改,删除。这些知识一是来源于新发表的属种或文献,可及时在库中增加进去;二是在推理过程中,某条权值较大(比如 0.7)的规则被用户反对,这就表明有新的事实(属种)或新的规则(描述特征)发现,这就需要领域专家和知识工程师协商,在知识库中增加相应的新知识(属种或描述)。

上述学习机制是利用 FoxBASE+ 的库修改和维护功能的有关命令实现的。