



# 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](mailto: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>

# 化石孢粉识别专家系统 (BFES)

## 研究进展\*

陈民敏 孙志亮 宋之琛 黄 嫔

(地质矿产部石油地质中心实验室,无锡)

(中国科学院南京地质古生物研究所)

本文是作者在 1990 年文章《化石 (孢粉) 识别专家系统》的基础上, 经近 2 年的运行实践, 在以下几方面作了改进。

### 一. 系统的软 / 硬件环境

1990 年文章中, 系统使用的机种是 IBM4361, 使用的软件是 FORTRAN IV 嵌套 SQL/DS。考虑到现在国内微机较普遍, 用户也主要是国内孢粉工作者, 因此, 我们将系统移植到微机上 (IBM PC 及其兼容机, 如国产长城系列微机), 软件使用 CCDOS 操作系统支持的汉化 FoxBASE+, 从而大大有利于系统的普及、推广。

### 二. 系统结构

BFES 采用模块式结构, 由 3 级层次, 6 个模块组成, 3 级层次为: 知识库级, 事实级, 控制级 (推理机), 6 个模块是: 知识库, 推理机, 事实库, 学习机, 解释机, 人机接口 (插图 1, 图中箭头表示数据流向)。

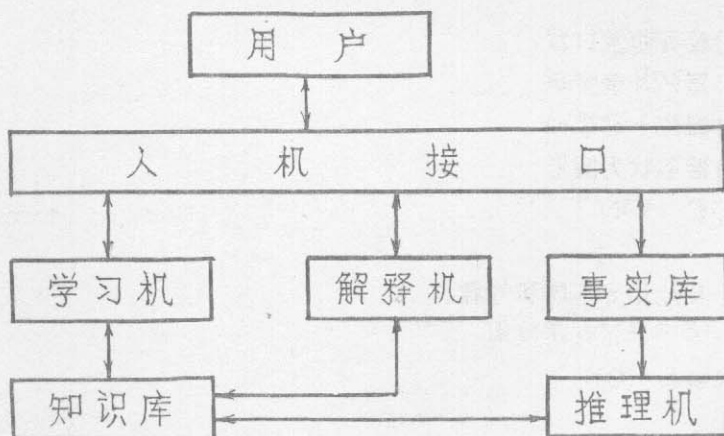


插图 1 BFES 结构示意图

这 6 个模块有的是一张表 (如知识库), 有的是一段程序 (如推理机), 有的只是一些

\* 开放实验室资助项目, 893110 项。

命令。一方面，每个模块都是独立的，一个模块修改，不影响其它模块，保证了各模块的相对稳定；另一方面，各模块之间通过某些接口命令相互调用，具有相当的灵活性，从而使系统成为既相对独立，又相互联系的整体。

### 三. 知识表示和知识库建造

化石孢粉属种是一个系统性很强的庞大分类群，属种识别有比较完整的、严格的鉴定特征。要准确地反映这样一个复杂的知识群（集），我们原来只用特性表一种知识表示方法是不够的，需要采用多种知识表示方法。现在 BFES 采用了产生式，特性表式和过程式等多种知识表示方法。下面简要介绍这些方法的应用及相互联系。

产生式方法 即规则表示法，其形式为：如果……则……，用于表示科（类）特征知识。如：

#### 规则 1

如果 孢粉粒有萌发口器  
且口器位于近面  
则 蕨类孢子 (JB)

#### 规则 2

如果 孢粉粒有萌发口器  
且口器区外壁变薄  
且口器位于远极面  
且口器形状为长形  
或口器中部有孔  
则 沟孔粉 (GK)

#### 规则 3

如果 孢粉粒有萌发口器  
且口器区外壁变薄  
且口器位于远极面  
且口器形状为圆形  
则 孔型粉 (KX)

#### 规则 4

如果 孢粉粒分化为本体和气囊  
或本体远极面有萌发沟  
则 具气囊粉 (QN)

#### 规则 5

如果 孢粉粒未见萌发口器  
且形状为圆形至椭圆形  
且表面无或具褶皱  
则 无口器粉 (WK)

#### 规则 6

如果 孢粉体由两个或更多的孢粉粒组成

### 则 多孢体 (DB)

**特性表方法** 是用二维表格表示属种特征知识, 因为这一类知识相对比较固定和雷同。

**过程式方法** 是用一段子程序表示某类知识或某项工作。这里用于表示知识查询、证据提问等方面的工作。如: 我们在知识库查询方面研制了数据查询程序来执行这一任务。

BFES 针对上述不同的知识表示, 采用相应的知识库建造方法。对产生式表示的逻辑关系用层次 (树) 模型组织 (插图 2); 对特性表表示的知识用关系模型组织 (表 I); 对过程式表示的知识用程序文件的方式组织。

这种树表结合的结构反映了化石孢粉分类系统, 将化石孢粉按科 (类) 分别建表, 运行时, 很快到达所需的表上, 识别就在这个表上进行。这样, 一个复杂的化石孢粉分类系统就被分解化简, 并使系统具有良好的可扩充性。

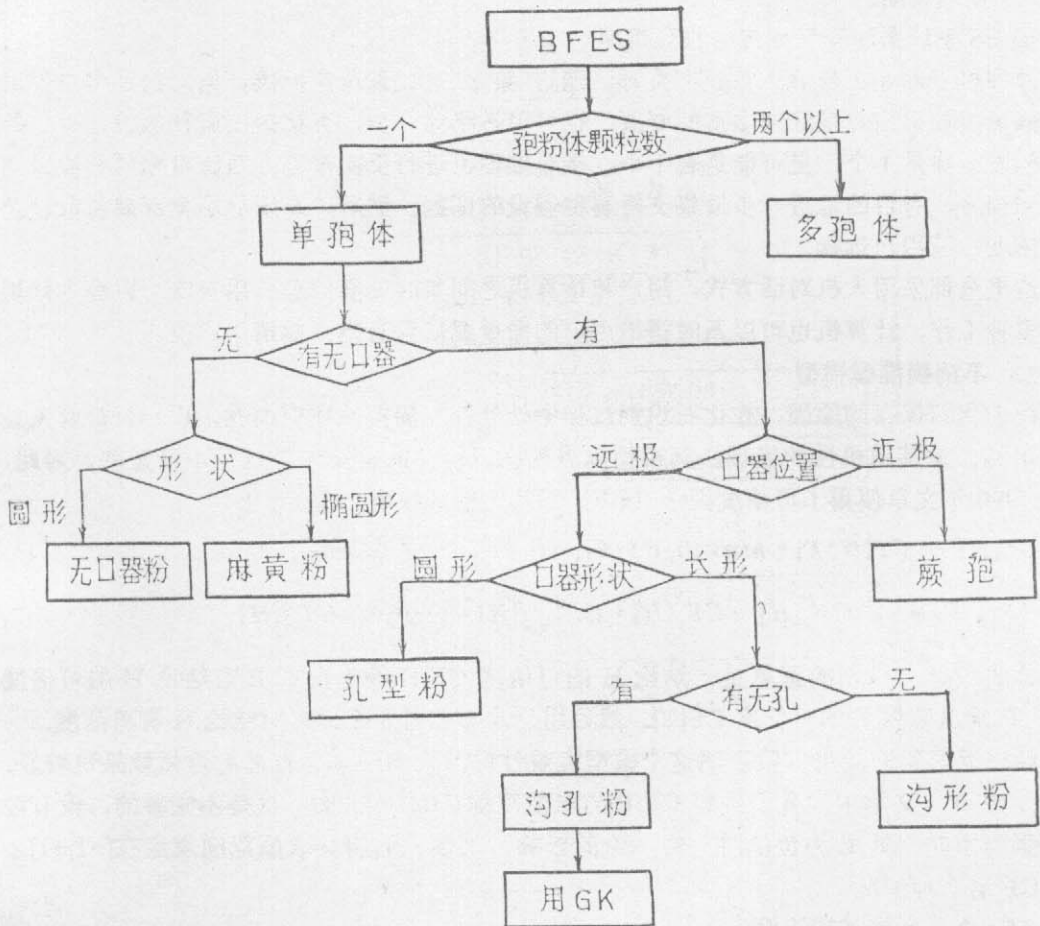


插图 2 BFES 知识层次 (树) 型结构关系示意图



表 I BFES 知识关系型结构示意图

表名: GK

记录号	属种名	特征 1	特征 2	...	特征 n
1	属种 A	特征 A1	特征 A2	...	特征 An
2	属种 B	特征 B1	特征 B2	...	特征 Bn
				...	
				...	
m	属种 k	特征 k1	特征 k2	...	特征 kn

## 四. 推理系统

### 1. 推理策略

BFES 系统采用正反向混合推理策略。

推理机一开始工作就进行正向推理, 通过“菜单”方式请用户提供证据, 帮助用户提出属种假设。如果用户是经验丰富的专家, 也可以省略这一步, 直接提出属种假设。得出的属种假设可能是 1 个, 更可能是若干个。接着推理机进行反向推理, 系统根据属种假设调用提问程序, 有目的地进一步搜集支持属种假设的证据。然后, 系统显示全部属种假设及其可信度, 请用户选择。

这里全部采用人机对话方式, 用户和计算机之间实时交换信息。用户既可以命令计算机做多种工作, 计算机也可以及时指出用户的命令或信息有错, 提请用户改正。

### 2. 不精确推理模型

由于化石保存的原因, 在化石识别过程中处处存在随机性和模糊性。识别特征有无是不确定的, 是其随机性的表现; 其模糊性表现在识别性的定义的外延是不清楚的。对此, 笔者 1990 年文章使用了可信度因子 (CF) 不精确推理模型, 即:

$$CF[H] = CF[H, E] \cdot \text{Max} \{0, CF[E]\} \quad (1)$$

$$CF_n[H] = \sum_{i=1}^n CF_i[H] - CF_1[H] \cdot CF_{i+1}[H] \cdot \dots \cdot CF_n[H] \quad (2)$$

公式 (1), (2) 的意思是: 结论 H 的可信度 CF[H] 等于证据 E 对结论 H 的可信度 CF[H,E] 乘以证据 E 的可信度 CF[E]。然后用公式 2 求出 n 个证据对结论 H 有可信度。

经过近两年的使用实践表明这个模型直观性好, 容易理解, 推理的近似效果也较好。但是, 1990 年文章中, 我们规定 CF[E] 的取值范围在 [0, 1] 之间, 这是不完善的, 没有考虑证据为假时, 即 E 为负值时, 对结论的影响。现在我们将其取值范围规定在 [-1,+1] 之间。CF 有 3 种情况:

CF > 0 证据支持假设;

CF < 0 证据反对假设;

CF = 0 证据没有出现。

当系统进行反向推理时, 提出 1 个证据要求用户证实。用户首先要在“是”、“不是”、

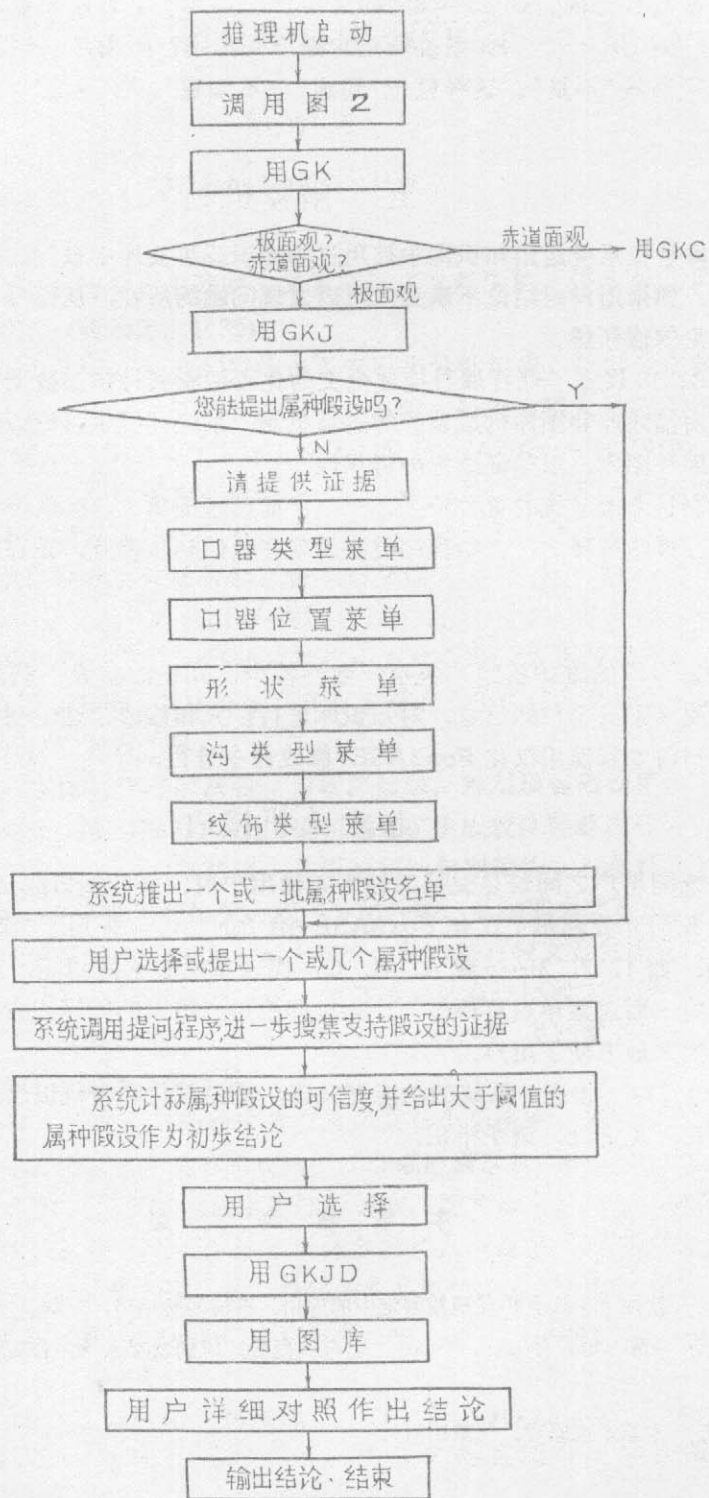


插图 3 沟孔型粉正反向推理框图

“不知道”三者之中择一。如果“是”，即  $CF > 0$ ，那就要在  $[0,1]$  之间给出证据的可信度是多少；如果证据没有出现，即  $CF = 0$ ；如果用户观察到的不是系统要求证实的证据，而是另外的证据，则  $CF < 0$ 。比如系统问用户是三孔沟吗？但用户观察到的不是三孔沟，而是三沟，则应回答“不是”。这样就把“知道”，“不知道”，“不确定”，“反对”等情况区分开来。

## 五. 解释机制

所谓解释是指系统运用知识库中被用过的知识或事实库中被用过的证据，回答用户对结论的质疑。如果用户对结论不满意，可以发现问题的所在，从而提高系统的透明度，增强用户对结论的信任感。

为此，BFES 设置了储存属种描述全文的化石孢粉属种描述数据库和储存化石图版照片的图库，由描述库和图库构成系统的解释机制。系统在提供结论以后，用户如果需要可以调用描述库和图库，以印证结论的可靠性。

解释机制是专家系统的重要特征之一，目前已经提出了许多解释方法，如预制文本和执行追踪法、策略解释法、自动程序员法。现在我们初步涉及，正在继续研制。

## 六. 学习机

是系统知识库的维护机制，包括一致性检查，知识的扩充、删减和修改。系统通过“学习”，接受专家或用户的信息，对知识库进行扩充和修改。这一功能由系统管理员或被授权的化石孢粉专家使用汉化 FoxBASE+ 修改命令进行。

## 七. 人机接口

这是系统同用户之间信息交换的输入/输出手段。好的接口应是尽可能地方使用户。在这方面，我们充分利用了汉化 FoxBASE+ 给予的功能，采用了“汉字菜单”技术，用最简明的代码，如 1, 2, 3……或 A, B, C, 代表系统的不同功能，用户根据需要提供输入代码，有如在餐馆面对菜单点菜那样，计算机即按用户指定的内容或方式去完成规定的工作任务，这就大大地方便了用户。

BFES 经过以上改进，系统的功能更完善，使用汉化和微机以后，也更有利于普及推广。但也存在不足之处，请予指正。

## 主要参考文献

- 陈民敏, 1985: 初探电子计算机在孢粉鉴定中的应用. 石油实验地质, 7 (4): 324—328.  
——、宋之琛、黄 焱、孙志亮, 1990: 化石(孢粉)识别专家系统. 古生物学报, 29 (1): 106—114.  
赵瑞清, 1987: 专家系统原理. 气象出版社.