

卢佩章 许国旺 著

气相色谱专家系统

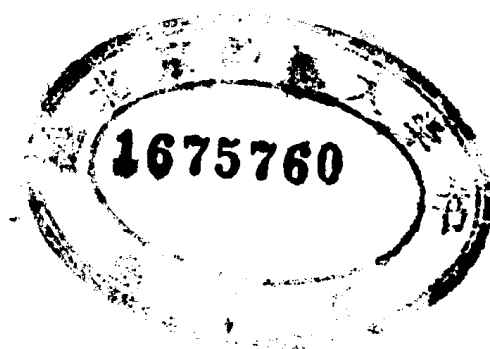
最佳柱系统的
选择及其软件

山东科学技术出版社

气相色谱专家系统

最佳柱系统的选择及其软件

卢佩章 许国旺 著



山东科学技术出版社

鲁新登字 05 号

气相色谱专家系统

最佳柱系统的选择及其软件

卢佩章 许国旺 著

*

山东科学技术出版社出版

(济南市玉函路 邮政编码 250002)

山东省新华书店发行

山东新华印刷厂印刷

*

787×1092 毫米 16 开本 20 印张 4 插页 446 千字

1994 年 5 月第 1 版 1994 年 5 月第 1 次印刷

印数:1—1000

ISBN7—5331—1343—8
O·56 定价 38.00 元

内 容 简 介

本书由浅入深地对气相色谱专家系统的理论和实践依据及其计算机软件作了系统和深入的介绍。对目前国际上尚未彻底解决的两个问题,即“什么样品应采用气相色谱法分析?”、“应采用哪种气相色谱柱系统?”作了系统回答,并给出了很多实际谱图和保留值数据。

这是国际上第一本系统介绍气相色谱专家系统的专著。在给出大量色谱知识的同时,还公布了气相色谱专家系统的源程序。因此,本书可供化学、化工、石油、地质、环保、医学、卫生部门的色谱工作者及科研人员、分析测试人员参考,也可作为大专院校有关专业高年级学生、研究生参考。同时,对于有志于计算机人工智能应用和研究的科技人员来说,本书无疑也是很有参考价值的。

2011/37/07

山东省泰山科技专著出版基金会

名誉会长 赵志浩 宋木文 陆懋曾 伍 杰
卢鸣谷 董凤基 宋法棠

会 长 陈光林 石洪印

副 会 长 宋桂植 何宗贵 吕可英 车吉心
孙肇琨 王为珍(常务副会长)

秘 书 长 王为珍(兼)

副秘书长 尹兆长

理 事 (以姓氏笔画为序)

王为珍 王凤起 尹兆长 刘韶明
李道生 李德泉 张传礼 陈 刚
蒋玉凤

评审委员会 (以姓氏笔画为序)

卢良恕 吴阶平 杨 乐 何祚庥
罗沛霖 高景德 唐敖庆 蔡景峰
戴念慈

山东省泰山科技专著出版基金会赞助单位

山东省财政厅

山东省出版总社

山东省科学技术委员会

山东科学技术出版社

山东泰山酿酒饮料集团总公司

董事长兼总经理 张传礼

山东金泰集团股份有限公司

董事长兼总裁 刘黎明

我们的希望

进行现代化建设必须依靠科学技术。作为科学技术载体的专著,正肩负着这一伟大的历史使命。科技专著面向社会,广泛传播科学技术知识,培养专业人才,推动科学技术进步,对促进我国现代化建设具有重大意义。它所产生的巨大社会效益和潜在的经济效益是难以估量的。

基于这种使命感,自1988年起,山东科学技术出版社设“泰山科技专著出版基金”,成立科技专著评审委员会,在国内广泛征求科技专著,每年补贴出版一批经评选的科技著作。这一创举已在社会上引起了很大反响。

1992年,在山东省委、省政府的支持下,在原“泰山科技专著出版基金”的基础上,由山东省出版总社、山东省科学技术委员会和山东科学技术出版社共同成立了“山东省泰山科技专著出版基金会”,并得到企业界的热情赞助,为资助学术专著的出版提供了更加可靠的保证。

但是,设基金补助科技专著出版毕竟是一件新生事物,也是出版事业的一项改革。它不仅需要在实践中不断总结经验,逐步予以完善;同时,也更需要社会上有关方面的大力扶植,以及学术界和广大读者的热情支持。

我们希望,通过这一工作,高水平的科技专著能够及早问世,充分显示它们的价值,发挥科学技术作为生产力的作用,不断推动社会主义现代化建设的发展,愿基金会支持出版的著作如泰山一样,耸立于当代学术之林。

泰山科技专著评审委员会

1992年12月

前 言

随着信息革命全面而深入地进入到我们的生活,我们深信已经广泛使用并且还将进一步广泛使用的气相色谱法,将和信息社会同步发展。计算机在分析仪器中的应用已取得了很大的成就,目前任何一台分析仪器已离不开计算机,气相色谱仪也如此,在温度控制、数据处理等方面已经取得了广泛的应用。

我们在 1974 年就开始了将计算机应用在气相色谱中的研究,不幸中途夭折。1978 年重新提出要开展智能色谱的研究。在国家科委专项支持下,智能气相色谱列入国家重点项目。智能与否目前已成为所有分析仪近代化的一个标志。“智能”促进了分析仪的销售,也造成了一些误解,因而我们在 1985 年开始将智能色谱的研究明确为发展色谱专家系统。因为随着气相色谱的广泛使用,仍然需要气相色谱的专家针对具体的样品推荐合适的分析方法,解决定性定量中的一系列困难。而随着计算机技术的发展,这些问题有可能形成专家系统让计算机来回答。经过 7 年的努力,已初步实现了上述目标。这期间先后有卢佩章、李浩春、邸凌等同志,陈孟尝、林炳承、林彬生、许国旺、张祥民和徐云等研究生参加了专家系统的研究工作。

我们最近几年来的研究,已初步形成了气相色谱专家系统,但由于各方面原因,要写全有困难。由于气相色谱柱系统的推荐,在国际上从来没有人写过,从来没有人讲过,更没有软件,因此把它总结成本书。书中公布了世界上第一个气相色谱最佳柱系统推荐软件的全部源程序,希望能促进这方面的工作。这是一个在科学上证明可以应用的源程序,但要商品化还要进一步改进才能满足各方面要求。我们在这儿发表了,希望大家使用,并请提出宝贵意见,以便提供更先进的学术思想,产生更先进、更可靠的软件。

由于本书系统地回答了有关色谱柱系统推荐的问题,提供了系统的理论知识和广泛的实践和基础数据,因此,对于欲应用和掌握气相色谱专家系统软件的读者,是必备的参考书。

本书在编写过程中,承蒙国家色谱中心领导张玉奎教授的热情指导和大力帮助,大连化学物理研究所有关领导和同志们也提供了帮助。在此一并致谢。

著 者

1993. 8

符 号 表

除正文特别注明者外,本书所用符号具有如下物理意义:

A_1	碳数规律常数
a	常数
a_1	常数
a_2	常数
b	常数
b_1	常数
b_2	常数
c	组分在气相的浓度
C	代表“碳”
$CCCL$	相应碳链长度
d	柱内径
d_t	液膜厚度
GI	基团增量
H	理论板高
H^1	$k' = 1$ 时的理论板高
H^∞	$k' = \infty$ 时的理论板高
ΔH_v	蒸发热
H_{eff}	有效理论板高
h	峰高
I	保留指数、电离能或噪音
K_1	总分离效能指标
K_3	峰高分离度
k	Boltzmann 常数或 Planck 常数
k'	容量因子
L	柱长
M	分子量
M_g	敏感度或检测限
N	柱效
n	碳数,尤指正构烷烃
M_1	一阶原点矩
M_2	二阶中心矩
P	蒸气压或峰容量
P'	碳数间峰容量

P_2	峰高法定量分析误差
R	气体常数
r	分子间有效距离
S_d	柱系统的选择性指标
T	柱温(K)
t_0	死时间
t_b	沸点
t_r	保留时间
u	线速
V	体积
V_R	保留体积
V'_R	调整保留体积
V_g	气保留体积
V_w	van der waals 体积
W_L	固定液重量
$W_{1/2}$	半宽
X	分子间作用力位能场
α	相对保留值或分子极化率
α'	保留时间之比
β	相比
μ	偶极矩
ρ	密度
σ	均方差
τ	拖尾因子
Φ	峰高比
χ	分子连通指数

常用角标

A	溶质
B	溶剂
l	液态
n	正构烷烃

目 录

符号表	(1)
第一章 绪论	(1)
§ 1.1 什么是气相色谱专家系统	(2)
§ 1.2 气相色谱专家系统在国内外的的发展情况	(4)
§ 1.3 气相色谱专家系统应选用的设备	(7)
§ 1.4 气相色谱专家系统要解决的几个主要问题	(8)
第二章 判别样品可用气相色谱法的依据及其软件	(15)
§ 2.1 样品可否采用气相色谱法的判别	(15)
§ 2.1.1 采用气相色谱法的两个基本条件	(15)
§ 2.1.2 气相色谱柱分离极限的预测	(17)
§ 2.1.3 气相色谱柱流出极限的预测	(28)
§ 2.2 限制组分采用气相色谱法的原因	(41)
§ 2.3 某些不稳定或反应性组分的气相色谱分析	(47)
§ 2.3.1 含卤素腐蚀性无机气体的特殊性	(48)
§ 2.3.2 含卤素腐蚀性无机气体的色谱分离	(49)
§ 2.4 推荐是否采用气相色谱分离模式的软件	(52)
第三章 最佳气相色谱柱系统选择的依据及其软件	(81)
§ 3.1 最佳气相色谱柱系统选择的色谱理论和实践	(82)
§ 3.1.1 同系物分离的最佳气相色谱柱系统	(82)
§ 3.1.2 异构体分离的最佳气相色谱柱系统	(97)
§ 3.1.3 沸点在 25℃ 以下的气体组分分离的最佳气相色谱柱系统	(111)
§ 3.2 最佳气相色谱柱系统推荐的软件	(129)
§ 3.2.1 有机组分分析时最佳柱系统的推荐软件	(130)
§ 3.2.2 无机气体和轻烃分析时最佳柱系统的推荐软件	(157)
§ 3.2.3 气相色谱柱系统推荐专家系统的总体结构	(175)
§ 3.3 验证气相色谱柱系统推荐的色谱谱图库	(186)
§ 3.3.1 谱图库的建立	(186)
§ 3.3.2 典型谱图的选取	(190)
§ 3.3.3 谱图库的应用	(192)
§ 3.4 常见气相色谱柱系统推荐方法介评	(208)
§ 3.4.1 固定相的分类	(208)
§ 3.4.2 固定相的选择	(210)
§ 3.5 新型填充柱——毛细管柱双柱系统	(213)

§ 3.5.1	最佳双柱匹配方案的理论预测	(213)
§ 3.5.2	双柱系统的仪器装置	(214)
§ 3.5.3	双柱系统的应用	(216)
附录 1	谱图库中的部分谱图	(228)
附录 2	气相色谱法部分重要固定相商品名对照	(303)

第一章 绪 论

气相色谱法是目前使用得最广泛的分析方法之一,这种趋势还将进一步保持下去。随着近代气相色谱仪的发展,分析方法确定后,分析一个样品所需时间一般只需几分钟、几十分钟即可完成。但对特定样品分析应选用的色谱方法、色谱柱、最佳的操作条件等的确定以及对最后所得谱图的定性、定量结果的获得,仍然需要由高水平的色谱专家来参与。因此,气相色谱日益广泛使用后,有经验的色谱专家将越来越需要。另一方面,已进入信息社会的今天,人工智能技术发展后,很自然的发展趋势就是利用计算机技术把专家的知识总结归纳,用专家系统技术来实现和代替专家的经常性劳动,因此,气相色谱专家系统的发展是必然趋势。^[1~14]

由于专家系统是在累积各种专家知识的基础上发展的,因此它能更全面地反映各方面的意见,可正确地回答有关问题,具有很强的实用性。比如,在实践中,大量的实际问题都难以建立精确的数学模型,也不宜用常规程序求解,而往往是要专家的独到经验。对于这类问题,正是专家系统研究的主要内容。同时,尽管获得科学的应用它的能力的培养很难,但人的生命有限,而且,人类专家的数量总是有限的,培养他(她)们不仅花费高昂,而且需要时间。如果有了专家系统则可捕捉专家的特长,把他们的知识存储起来,不仅可以调用,而且可以永远留存。同时,智能计算机发展后,专家系统软件还可继续发展,不仅如此,人脑的功能将得到发展、加强。例如,如果没有计算机,要对气相色谱曲线进行全谱图拟合是不可能的,而用了计算机后,不但可能而且可进一步对重叠峰进行解析^[15,16]。由于这一特性,使得专家用了专家系统能发挥更大的作用。不过,创造性的劳动还是要靠人,专家系统的进一步发展也需要人。

到目前为止,很多分析方面的专家系统还没有发展到很成熟的阶段^[11,17~30]。一般是采用事实归纳法^[18~20],只有零碎事实的归纳,缺乏强有力的理论基础。这样发展出的专家系统有局限性,应用时也会有困难。而我们现在能发展出国际上第一个气相色谱专家系统^[31~37],就是在强有力的理论研究上,系统地总结了大量实践,并运用了先进的计算机技术^[34,38~41]。例如样品可否采用气相色谱法来分析,一般只认为样品只要有挥发性即行^[42~45],而没有考虑到底是溶质蒸气压不够检测器测不出还是采用的气相色谱柱保留值太大而限制了气相色谱法的采用。实际上,我们对溶质的蒸气压、保留值等进行研究后发现,只要能以合适的保留值从柱子中流出,溶质的蒸气压满足检测器 FID 的要求是没有问题的^[34,46]。且在判断哪一个流得出、哪一个流不出、在什么柱子上什么组分能流出时,如果没有我们提出的相应碳链长度(CCCL)概念^[47],就不易解决。而对于哪些物质不能用气相色谱分离这个本质问题,也只有经过系统的研究后才能确定^[48]。例如用目前已发展出的气相色谱毛细管柱,使得分子式有差异的组分均可分开,因此问题的焦点在异构体的分离。当柱系统峰容量不足和链状化合物位置异构体中八位以后的相邻位置的位置异构体须分开时,气相色谱法还不能解决。其它情况的样品,原则上可用气相色谱法通过调节

极性或使用特殊选择性柱子来解决。这些结论都是我们在系统研究后在国际上首次提出的。因此专家系统要真正地实用,没有强有力的理论基础不行。

为了应用这个气相色谱专家系统并进一步发展和改进,我们编写了这本书,使读者对本系统所依赖的理论和实践有全面的理解。本书主要是对专家系统的色谱基础和计算机技术作了系统总结,同时也发表了国际上第一个气相色谱专家系统源程序。我们很希望我国能有更多的同志参加这项研究工作,使现在在国际上已处于领先地位的研究,以后能继续领先下去,为国家作些微薄贡献。

§ 1.1 什么是气相色谱专家系统

所谓气相色谱专家系统,它是一个智能程序系统,具有大量的专家级的气相色谱领域的专门知识及深厚的理论基础,并且运用人工智能的理论和技巧,根据一个或多个色谱专家作决定的过程,来解决需要色谱专家才能解决的气相色谱方法发展以及对色谱图的定性定量问题。也就是说,色谱专家系统是一个程序系统,可以以专家的水平完成专门的、一般是困难的专业任务。它把一般的、模仿人类的解题策略与这个问题所特有的大量理论知识和经验知识结合起来,使得色谱专家也可借助于气相色谱专家系统获得知识的延伸,发挥出更大的作用。与所有其它专家系统一样,色谱专家系统也是知识基系统。

专家系统与其它系统不同。首先,专家系统不同于一般的程序系统。一般的应用程序不注重知识、知识的表达和结构。知识在一般的程序中是以隐式形式存在的,一旦程序完成之后,其功能就确定下来,一般不易修改。而专家系统是以知识结构为基础来组织系统的,通常具备若干卓著的特色,其中包括符号的表达、符号推理和启发式搜索,知识库和程序部分是分开的,即独立于程序,因而易于增删和修改。第二,专家系统与数据库系统不同。数据库系统是对大量的数据进行组织和管理,提供检索策略。一般情况下其检索侧重的是对数据的查寻而不注重推理本身。在专家系统中,除了使用传统的数据库概念及技术外,需要将人类知识组织起来。在知识表示的设计中,要与推理网络的设计联系起来。第三,专家系统与一般的模拟程序也不一样。一个模拟系统或仿真系统,采用设想的模型来模拟客观的过程,而专家系统则着眼于解决问题所依据的机理表示,尤其是对那些不准确的机理表示。

它主要由以下几部分组成:

1. 知识库

它是气相色谱领域知识的存贮器,包括两类。第一类是领域的事实。如保留指数、基团增量、固定相特性参数等,这是广泛共有的知识。也就是写在书上的知识或常识。第二类是产生式的知识。如色谱专家系统的规则(详见第三章),它是色谱领域中正确的实践和判断的知识,很大一部分是凭经验得到的,是专家经过多年工作而掌握的“善于猜想的艺术”。

知识的确定性与完善性决定于领域专家,而知识的可用性既与领域专家有关,又与计算机工作者有关。因为即使一条可用的知识,由于计算机工作者处理不当,存放进计算机之后,用起来非常困难,这就表示可用性差,也就是说,如何表达专家的知识(称为知识表

达)是一项非常重要的工作,是专家系统建造中的一个关键问题。在我们的气相色谱专家系统中采用了三种表达方式:

(1)产生式表达法

它将专家的知识由称作产生式的规则集来表示,每一产生式构成一个知识模块的一条规则,可以写成 $a \rightarrow b$ 或 IF a THEN b 的形式,其中 a 称为前提部分,b 称为结论部分。例如,气相色谱专家系统中有这么一条规则:

规则	意义
(if (optical-isomers)	如果 分离光学异构体,且
(hydrogen-bonding-substrates)	其衬底能形成氢键
(then (optical-active-hydrogen-bonding-phase)	则 选择光活性氢键相

由于诸产生式之间是独立的,所以用产生式表达的系统修改、扩充都特别有利。这也是使产生式成为专家系统中用得最多的一种知识表示方式的原因。

(2)框架表示模式

框架就是一种描述立体形态的数据结构,它由一组槽所组成,一般框架有如下形式:

《框架名》

《槽名 1》《侧面名 11》(值 111,值 112...)

《侧面名 12》(值 121,值 122...)

⋮

《槽名 2》《侧面名 21》(值 211,值 212...)

《侧面名 22》(值 221,值 222...)

⋮

⋮

一个框架可以有任意有限数目的槽,一个槽可以有任意有限数目的侧面,一个侧面可以有任意有限数目的值。由于在我们的专家系统中要确定物质的保留值,用到的很重要的数据是基团增量值,且不同的基团在不同的位置于不同的柱子上数值不一样。为了将这一事实清楚地表达出来,一方面利于计算机处理,另一方面也要使修改和添加方便,框架表示方式显然是最佳选择^[35],详情请参看第二章。

(3)过程表示方式

过程表示是把有关领域的知识包含在若干过程中,它便于表示启发式知识,在模拟人类的假推理(default)和或然性推理(Plausible)等非形式推理方面具有一定能力;这种表示方式特别适于那些对领域知识了解比较透彻的情况。比如,柱系统推荐时,样品情况的问诊知识就是用过程式表示的。

2. 推理机

它是一组程序,用来控制、协调整个系统的。它根据当前输入的数据(如待分离样品的物化性质等),利用知识库中的知识,按一定的推理策略,去解决当前的问题。推理方法有:

(1)正向推理

即由原始数据到结论(称为数据驱动)。

(2) 反向推理

即先提出假设,然后去找支持假设的证据(称为目标驱动)。

(3) 正反向混合推理

即先根据数据库中的原始数据,通过正向推理,帮助系统提出假设,再运用反向推理,进一步寻求支持假设的证据。究竟采用何种控制策略,它和具体的应用方式和功能有关,在实际使用中还可以采用几种方式结合的方法以获得满意的控制效果。比如,在气相色谱专家系统中,先以正向链推出所选用的柱系统^[32],再以逆向链检索出谱图库中类似样品的谱图^[41,49~51]。

3. 人机接口部分

专家系统的建造靠专家,而使用则可以是专家,亦可以是一般工作人员,因而人机接口部分很重要。这部分主要包括:①数据和输入,比如用户对分离问题的描述等;②推理过程的解释,如由原始数据来得到结论的依据,对用户询问的实时回答等;③结果的输出部分。

4. 数据库

用于存贮该领域内初始证据和推理过程中得到的各种中间信息,也就是存放用户回答的事实、书籍的事实和由推理而获得的事实。总之,数据库存放的是该系统当前要处理对象的一些事实。如在柱系统推荐时,有一个数据库可用于记录已向用户提问过的事实,以阻止对同一事实向用户反复提问,而其它一些数据库用于记录真事实,并贮存推理过程中获得的中间结果,如最大最小组分结构碎片、最大最小组分保留值等,它们都以表的形式存放着。

5. 知识获取部分

知识获取部分也称为学习功能,它为修改知识库中原有的知识和扩充知识提供手段。在专家系统的交互过程中,认为哪些知识不合理,甚至错了,或者需要加进新的知识,都要利用这部分。它一般具有如下两个功能:①能删除知识库中原有的知识;②能将新知识加入知识库中。

§ 1.2 气相色谱专家系统在国内外的的发展情况

一般认为,分析化学的专家系统有三大类:①实验设计;②实验结果解释;③实验室仪器控制和故障诊断。最多的为第②类,发展了质谱^[1,2,52,53]、红外光谱^[54~58]、NMR^[57~59]、X射线荧光^[60]和粉末射线仪^[61]等数据的解释系统。但最有名的为斯坦福大学 Feigenbaum 等^[53]建立的 Dendral 和 Meta-Dendral 系统。后者由于具学习功能,因此它能自动发现、抽取这一领域的专业知识。迄今为止已在 5 个分子结构簇发现了分子断裂规则,其中 2 个是从前未报道过的新发现^[62]。而且大连化学物理研究所发展的 ASES/MS 系统,使用了以基本质谱数据库为基础的计算机学习方法来总结质谱解析规则,首次实现了单独利用质谱中性丢失信息推导未知物亚结构的方法,系统达到了能提供日常分析使用的水平^[63,64]。在仪器控制和诊断方面,建成了过程控制的专家系统^[65],也发展了质谱真空系统

故障诊断的专家系统^[66]。在实验设计领域,所做的工作主要集中在对给定样品选择最合适的分析方法上。出现了对药物样品分析方法^[67~69]和对各种样品表面分析方法提供指南的专家系统。

色谱专家系统是在 80 年代中期变得十分热门的^[6,10,11,17]。较有影响的为 Varian 公司的 ECAT(Expert Chromatography Assistance Team)^[18],欧共体的 ESCA(Expert System in Industrial Chemical Analysis)^[21~26,28,29]及中国科学院大连化学物理研究所的 ESC(Expert System for Chromatography)^[31,39,40,70]等。ECAT 始于 1984 年,目的是欲在人类专家的水平上,用设计成的专家系统完成 HPLC 的方法发展。目前已作了大量工作,到 1990 年 2 月时的完成情况见图 1-1。从样品基体和类型可获得柱和流动相的推荐。关于分离模式(GC/LC?)的推荐^[18]至今未见到实际工作,同时从其知识库来看,似乎文献实例的归纳比较偏重。

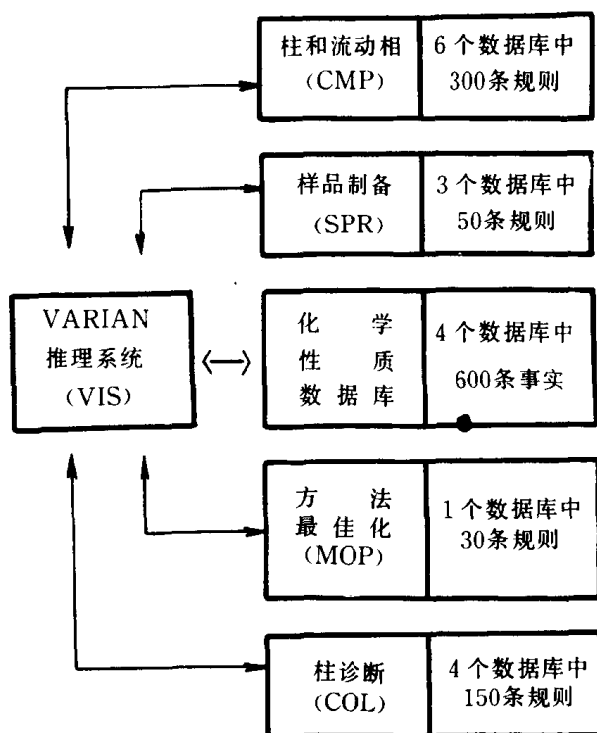


图 1-1 ECAT 专家系统方块图

欧共体的 ESCA 系统开始于 1987 年,主要目的是试验专家系统在 HPLC 应用于药物分析时发展色谱方法的可能性^[23],对象是中枢神经系统活性药物的纯度检查,系统结构见图 1-2。目前 ESCA 项目已结束,正在做模块的集成工作^[26]。

大连化学物理研究所的色谱专家系统有气相和液相两大部分(图 1-3),各有 5 个模块组成:①分离模式和柱系统的推荐;②样品预处理方法和检测器的推荐;③操作条件的最佳化;④峰的在线定性定量;⑤硬件故障诊断。各部分已均具规模。在液相色谱方面^[33,71],根据保留值方程 $\ln k' = a + b \ln C_B + c C_B$,发展了 HPLC 柱系统推荐软件,允许用户从分子结构和样品名通道进入以获得分离模式、样品预处理方法、固定相、流动相及检