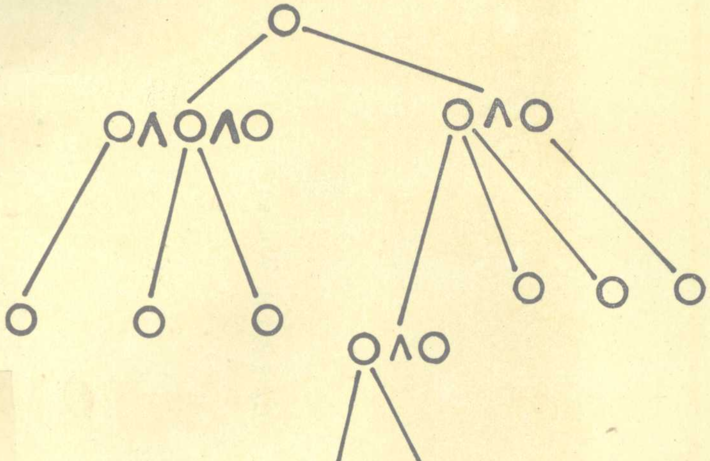
人工智能丛书

化学合成专家系统

ESCS

张经江 主编

国家自然科学基金与兵器科学研究院资助课题



西北工业大学出版社

人工智能丛书

化学合成专家系统 ESCS

国家自然科学基金与兵器科学研究院资助课题

张经江 主编

吴家榕 赵洁 副主编

张经江 吴家榕 赵 洁 顾企江 张 江 郑远祥 何 伟 著

江苏工业学院图书馆藏书章

西北工业大学出版社

1993年12月 西安

陕新登字 009 号

【内容简介】 本书共分三部分，第一部分为方法研究，详细叙述了化学合成专家系统的原理及设计思想，它包括：平面分子结构的线型表示及机内表示，合成树的构造，向前向后推理方法，分子的同构与相似度计算以及子结构的识别方法等关键技术。第二部分为系统设计，包括：系统总体结构及一些重要模块的结构，微机程序的使用说明书，以及用 BASIC 语言编写的源程序。第三部分为计算结果及展望，包括实例计算结果分析及发展前景。

本书是作者十年来工作的全面总结，适合于人工智能专业及计算化学专业的科研人员及大学本科生、研究生、教师参考书。

人工智能丛书
化学合成专家系统
ESCS
张经江 主编
责任编辑 柴文强
*
© 1993 西北工业大学出版社出版发行
（西安市友谊西路 127 号 邮编 710072）
全国各地新华书店经销
西北工业大学出版社印刷厂印装
ISBN 7-5612-0645-3/TQ·7
*
开本 850×1168 毫米 1/32 7.75 印张 157.5 千字
1993 年 12 月第 1 版 1993 年 12 月第 1 次印刷
印数：1—1 000 册 定价：20.00 元

序言

一个机器智能化的时代已经开始!

如果说经典物理学的形成和蒸气机的诞生导致了第一次科学革命和第一次产业革命，实现了人的部分体力劳动的机械化和自动化，使人类社会有了今天这样发达的物质文明和精神文明。那么正在到来的第二次科学革命的第二次产业革命则导源于广义信息科学的形成和计算机的出现，它正在实现人的部分脑力劳动的机械化和自动化。人们称能够部分代替人的脑力劳动的机器为智能化机器。正如蒸气机是一种比较原始的动力机一样，现有的电子数字计算机也只是智力机的雏形。人工智能的近期研究目标就是要提高现有电子数字计算机的智力水平，使其变得更加灵巧。人工智能的远期目标是探讨智能的基本机理，研究如何利用各种自动机来模拟人的某些思维过程和智能行为。智能科学有广阔的外延和丰富的内涵，她正处在蓬勃生长的青春期。为了增进读者对她的理解，在智能化的历史进程中有所作为，我们正编辑出版一套人工智能丛书，陆续阐述智能科学各研究分支的原理、方法、研究成果和发展动态。

张经江研究员主编的《化学合成专家系统 ESCS》是他所在的科研课题组十年研究工作的结晶，填补了我国在火化[redacted]专家系统研究方面的空白，反映了我国在这方面研究的最新成果。

化学合成是一个非良性结构问题，化学合成专家系统开发是一个举世瞩目的难题。在分子同构，分子相似度计算，分子结构的计算机表示和模拟，结构变换的向前和向后推理，子结构的匹配算法，环的识别算法，官能团的加成与替换条件，结构变换的

条件，由类反应变为预测反应的方法等等问题上，都存在许多尚待研究的课题。作者在他们的著作中，提出了一套系统实用的完整方法。如复合结点树构造法，路段演绎规则，路经优劣评价方法，有机结构线性表示法，专家系统 ESCS 中得到成功的应用。

作为我国第一本化学合成专家系统方面的专著，我们把她郑重地推荐给广大读者，希望能对您的专家系统研究，特别是化工类专家系统的研究带来效益。

何华灿

1993 年 12 月 20 日

此为试读，需要完整PDF请访问：www.ertongbook.com

前言

人工智能是计算机科学中最生动活泼的领域，而专家系统又是其中最光辉的一个方面。在某些学科中，科学家们对其内在的机理并不十分清楚，但是他们有着丰富的实践经验，这些正是专家系统大有用武之地的原因。譬如，化学是一个古老而又久经不衰的学科，它关系着人类生活的各个方面。在合成机理上化学家很难用一个数学公式去描述它，这给进行数值模拟带来了巨大的困难，这就是我们选择专家系统在这个问题上应用的原因。

关于专家系统的原理、构成、设计等一般描述的书籍已经很多，但是当一个计算机科学工作者要去设计一个专家系统时，一开始往往会觉得无从下手，他们很希望去了解一个具体的专家系统，以便能依葫芦画瓢模仿着去描述自己的专家系统。以往科学工作者常常去参考由斯坦福医科大学门诊药物系教授 Edward Hance Shortliffe 博士所著的书 MYCIN，这是一本治疗炎症疾病方面的一个专家系统，这导致了我国最早的一批医疗诊断专家系统的问世，由于这本书原文是英文，加上许多医疗方面的专业词汇，给人们阅读带来了许多困难。

我们编这本书的目的是给我国广大科技工作者，展示我们自己搞的化学合成专家系统的全貌。一方面，对于有兴趣致力于使用专家系统在化学合成方面的计算化学工作者，希望能推进这方面的工作，向更深的层次进军；另一方面，对于有兴趣致力于专家系统其他方面应用的计算机科学工作者能起一个抛砖引玉的作用。我们相信这本书对上述两部分人员会起到一些有益的作用。

本书是我们课题组十年来工作的描述，参加本课题组的工作

并参与了本书编写工作的同志还有：吴家榕、赵洁、顾企江、张红、郑远洋、何伟等同志。其中第一章、第二章、第四章、第九章以及附录部分由张经江编写，第三章、第八章由吴家榕编写，第五章及第六章由赵洁编写，第七章主要由赵洁、张红、张经江编写，4.11, 4.12 两节由顾企江编写。

本书大部分内容由贾裕民主任、杨泊清副研究员审核。

本书共分三个部分，第一部分是方法概述包括四章：绪言、ESCS 专家系统的设计思想、有机化学的线型表示以及第二代专家系统的设计思想。第二部分包括三章，程序系统的结构及技术，使用说明书以及程序模块 ZZ1，ZZ2。第三部分计算结果及展望，包括两章，实例结果分析及展望。附录部分包括国外两个专家系统的简介。

作者首先要感谢全国人工智能学会副理事长何华灿教授，本项科研工作正是在他的启发下开展的，在工作的各个阶段都得到了他的积极支持和肯定。

作者还要感谢，张明南教授、李福平教授、张厚生教授，他们在本课题的工作中都给予过指导和帮助。

作者还要感谢张懋森教授，于永忠教授、李伟民教授、蒋承伟教授、郝仲璋教授、李惠黎教授等，他们积极地肯定了该项工作的成绩，并提出过不少建设性意见。

作者对我所各级领导的支持表示感谢，他们是陈崇沼所长，顾汉全所长、卢炎处长、费家能处长、胡焕性主任、范士俊主任、杨培进主任。同时对我的同事们的关怀和鼓励表示感谢，他们是李秋年、杨正权、谢春生、张珊珊、邬志兴等高级工程师们。

作者对在各个方面对本课题工作有过帮助的同志表示感谢，他们是何国书，王国良、韩莉，还有很多的朋友，恕不能一一列出。

作者对于国家自然科学基金和兵器科学研究院综合处的支持表示深深的感谢，没有他们的支持该项工作是不能进行的。特别是张浅秋工程师独具慧眼，为该课题的进行做出了重大贡献!

本书能够得以出版，还应当归功于西北工业大学出版社的同志们，在目前出版界面临如此严峻的条件下，能够出版这本专著，实在需要拿出极大的勇气。

最后我要感谢我的父亲，他对我的要求十分严格。现仅以本书作为一个小小的礼物，献给他老人家。

化学合成专家系统是一个新的领域，由于作者水平有限，不当之处在所难免，望读者批评指正。

西安近代化学研究所　　张经江

1993.7.10

目　录

第一部分　方法研究 …… 1
第一章　专家系统在化学研究中的应用 …… 1
1.1　专家系统的产生与构成 …… 1
1.2　专家系统在化学科学中需要解决的问题 …… 2
1.3　国内外研究现状 …… 2
1.4　合成树与路段演绎法 …… 5
第二章　化学合成专家系统 ESCS 方法描述 …… 6
2.1　化学反应式的知识表示与化学合成树的构成 …… 6
2.2　路段演绎规则 …… 9
2.3　化学合成途径的寻求 …… 13
2.4　合成路径优劣评价方法 …… 20
2.5　优缺点评述 …… 23
2.6　结论 …… 26
第三章　有机化学结构的线型表示 …… 27
3.1　定序法则 …… 28
3.2　小结 …… 31
第四章　新型含能材料计算机辅助合成专家系统方法概述 …… 44
4.1　系统构成 …… 44
4.2　分子结构的线型表示及带有 R 部分的结构表示 …… 46
4.3　连结表表示法，线型表示与连结表表示的转换算法 …… 56

4.4 子结构集合及对分子结构的剖分 …… 61
4.5 分子的同构算法 …… 62
4.6 两个分子结构相似度计算 …… 63
4.7 用模型匹配方法寻找目的分子中的R和R′
并将类反应变为预测反应 …… 64
4.8 从已知结构到目的结构的向前推理 …… 68
4.9 分子结构的输入方法 …… 75
4.10 子结构的匹配方法 …… 80
4.11 官能团的加成与替换规则及条件 …… 82
4.12 结构变换规则及条件 …… 91
4.13 用官能团作基本单元存在问题及解决办法 …… 92
第二部分 系统设计 …… 95
第五章 程序结构及系统设计 …… 95
5.1 系统功能 …… 95
5.2 方法简介 …… 95
5.3 整体结构及各主要部分的结构设计 …… 96
5.4 程序实现 …… 107
5.5 程序设计技巧 …… 108
第六章 PC机程序系统使用说明书 …… 112
6.1 系统装机运行 …… 112
6.2 选择打印方式 …… 113
6.3 选择知识库 …… 114
6.4 建新库 …… 115
6.5 知识库管理系统 …… 115
6.6 进入运算 …… 120
6.7 打印输出 …… 136
6.8 预测方程的处理 …… 138

6.9 结束语 …… 139
第七章 微机程序系统 …… 140
7.1 微机程序系统模块 ZZ1 …… 140
7.2 微机程序系统模块 ZZ2 …… 164

第三部分 计算结果与展望 …… 209
第八章 实例计算结果分析 …… 209
8.1 知识来源 …… 209
8.2 对已知化合物计算实例结果分析 …… 210
8.3 预测方程实例计算结果分析 …… 213
第九章 展望 专家网络在合成中的应用 …… 216
9.1 专家网络 …… 216
9.2 用神经网络法计算各种含能材料的物理化学性能 …… 216
9.3 专家网络的应用 …… 221
附录 1 SECS 系统简介 …… 222
附录 1 CHESS 系统简介 …… 225
参考文献 …… 228

此为试读,需要完整PDF请访问: www.ertongbook.com

第一部分 研究方法

第一章 专家系统在化学研究中的应用

1.1 专家系统的产生与构成

计算机发明的直接目的是解决科学研究中大量而复杂的数值计算任务，后来人们又把计算机用于非数值计算，如：数据处理、数据库管理、办公室自动化、计算机语言翻译、自动控制等，这些工作大大地减轻了人的繁锁的重复的劳动，但是能否利用计算机代替人的某些高级的复杂的思维活动呢？回答是肯定的，这就是人工智能学科诞生的动机，专家系统是人工智能的分支之一。它的目的就是要弥补人们知识的不足，在已有知识的基础上，经过"学习"与"推理"得到新的知识，从而协助科学家更快地产生新的科研成果。

作为一个专家系统，首先它必须有一个知识库，对于某学科领域内知识要有一个形式化的知识表示方法。这是计算机进行推理的基础，这个知识库要尽可能地合理地搜集本领域的知识，并且能够通过不断"学习"与"实践"加以发展使其更加完整。目前知识表示的方法有：产生式法、框架法、谓词表示法、语义网络法、与或树法。其次必须建立一套机械化推理规则，推理规则也是一种知识，不过通常是过程性知识。第三建立如何选择推理规则进行推演的控制策略，即搜索，通常采用正向推理法和反向推理法两种策略。最后必须解决系统构成技术，编制出系统程序。

1.2 专家系统在化学科学中需要解决的问题

专家系统在化学领域内主要解决两类性质的问题：一类是未知有机化合物分子结构的确定问题；另一类是有机化合物的合成问题。

化学家在确定未知化合物分子结构的过程中，需要解决这样两个问题：第一，该化合物由哪些原子组成；第二，这些原子是怎样排列和联系的。前者对于化学家是容易解决的，即化学家能够写出化合物的分子式；而后者，由于一个给定的分子式其原子的排列和联系的各种可能性是如此之多，化学家往往感到非常棘手，人工智能试图在这个问题上对化学家有所帮助。

在有机合成中所需要解决的问题是，确定一系列化学反应，通过这些反应，用已有物质来合成得到某种物质。合成路线的确定需要考虑到反应的效率，所用物质的成本以及实验条件等等。在考虑合成路线确定过程中化学家和系统的关系时，有两种方法，一种是交互式方法，专家以交互方式对系统的产生过程不断予以指导。另一种方式是非交互式，即化学家不介入系统的运行。系统所用到的知识主要是化学反应的知识，这些知识的使用通常有两种方式：一种是从已有物质出发，进行化学合成得到新的物质，再对这些新的物质施加各种反应，这个过程一直进行到目的物质达到为止；另一种方式正好相反，从目标物质出发，考虑到可能产生目标物质的各种反应。这样问题就转变成为这些反应所用到的那些物质合成了，这个过程一直进行下去，直到反应需要的物质是原材料为止。

1.3 国内外研究现状

几十年来，专家们在解决有机化合物分子结构的确定问题方面做了大量的工作，1964 年斯坦福大学医院 Joshna Lederberg

提出了穷举所有环分子结构的算法，1965年与该大学计算机科学系人工智能科学家 Feigenbaum Edwerd A 合作开始研制 DENDRAL 系统，以后1976年，Brown Masinter Hjelmeland 又研制出了新的结构生成程序 CONGEN (Constrlneed GENerater)。1977年由 Engelmore a Mil 研制了 CRYSALLS 系统，它是人工智能在物质结晶学中的具体应用，它试图解决电子云密度图的解释问题，从而帮助化学家确定物质化合物的分子结构。

在有机合成方面主要有三个系统：

1. LHASA, 逻辑与启发，用于合成分析, 美国哈佛大学系 Corey. E, J 等研制于1969年。

2. SECS 系统，化学合成的模拟与计算，美国加利福尼亚大学化学家 Wipke.w.等研制于1977年。

3. SYNCHEM2 (SYNthetic CHEMistry)，有机化学, 美国纽约 State 大学 Gelernter 研制于1977年。

现将 LHASA 作一简单介绍, 它的系统结构如下图:

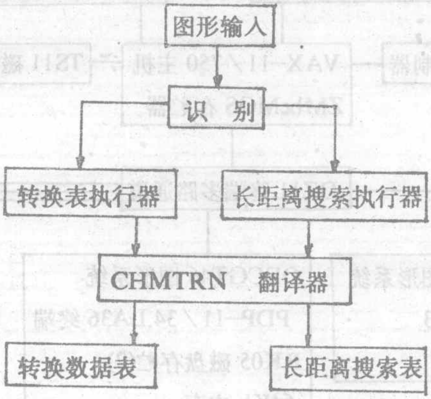


图1.1　LHASA 软件结构

LHASA 软件大部分是用 FORTRAN 语言编制，但是关于

有机反应知识与长距离搜索策略是用一种专门的化学语言CHMTRN编写的，这种语言很容易被化学家读写，而且也易于程序处理，上图指出了LHASA主要软件模块。

LHASA是一个使用图形输入／出的交互系统，用户给出一个需要合成的分子作为目标结构，LHASA产生为合成到目标的先驱物的图示建议，然后化学家选择出先驱物之一，LHASA返回到合成先驱物的建议，这样一步一步设计出完整的合成路径。图形输入／出使得化学家很容易且很愿意使用。系统具有增加新的信息的功能，以便不断完善反应知识库。下面是小型计算机基础上的LHASA计算机系统。

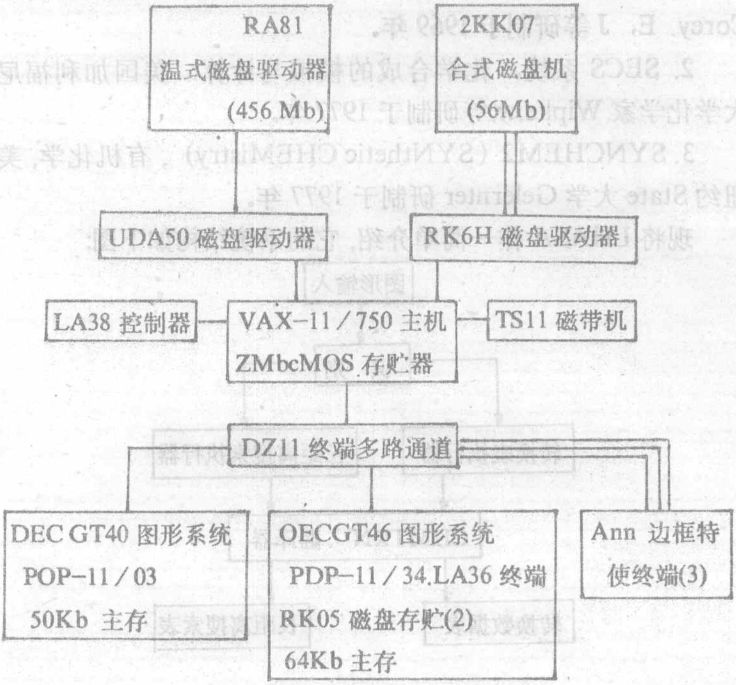


图1.2 LHASA 硬件系统

SECS 系统是 LHASA 系统的扩充，这些系统在解决实际问题时都是成功的。

1.4 合成树与路段演绎法

本书作者根据人工智能原理，设计出了一种复合结点树，或叫合成树。大意是这样，当要合成某种物质时，把此种物质当做根结点，从化学反应知识库中寻找能产生此物质的反应，每找到一条反应时，则从结点向下画一条线段，旁边标上此反应序号，线段下端标上此反应左边物质，叫做复合结点，搜索完后，再对下一层每一个复合结点内的每种物质重复此过程，直到所有复合结点内的物质为原材料为止，这样合成树就构成了。为了寻找合成途径，又规定了对所合成树进行路段演算的 13 条规则，这些规则不仅规定了演算结果，并规定了规则运用的先后次序，因之这 13 条规则也包括了控制规则。

本文作者对此种方法编制了 BASIC 程序，对于文中所举的例子，计算机都顺利的进行了演算，得出了预期的结果。

第二章　化学合成专家系统ESCS方法描述

为了合成某种物质，通常要知道关于此种物质的反应方程式，如果从原材料出发经过一次反应即可得到这种物质，那么其合成途径只是一个反应方程式，这种情况比较简单，只需建立一个化学反应方程式的知识库，然后在库内进行查找有关此种物质的反应方程式，找到后打印出来即可，但是如果从原材料出发，合成此物质需要多步反应，那么情况就比较复杂，要描述这种情况，需要对此种物质构造出一棵复合结点树，或叫合成树，并需对此树进行演算，从而得出从原材料到此种物质各种可能的合成途径。本章先描述合成树方法，合成路径演绎规则。然后举例说明使用方法，最后给出路径优劣评价方法与计算实例。

2.1　化学反应式的知识表示与化学合成树的构成

2.1.1　知识表示与知识库的建立

化学反应方程式在我们的知识表示中被描述为<映射式>，用算法语言来描述形式为:

<序号>:　<符号与式>→<被映射符号>

其中<序号>是一个整数，<被映射符号>即某一个化学物质的结构分子式，<符号与式>是若干个结构分子式，其间用“∧”号连起来，例如：a∧b∧c，c∧d，e等，而a，b，c，d，e代表不同化学物质的结构分子式。例如，反应方程式:

$$C+O_2 \rightarrow CO_2$$

此为试读，需要完整PDF请访问：www.ertongbook.com