

**\*学科发展\***

# 计算机化学的发展和前沿

许志宏 周家驹 杨章远 董 茜

(化工冶金研究所)

**【摘要】** 本文从计算机化学的学科特征、国内外的动态、国内的研究基础以及科研政策等方面阐述了在我国建立和发展这一新兴分支学科的必要性。

## 一、引 言

电子计算机诞生至今,虽然只有 40 多年的历史,但已跨越了好几个阶段。硬件方面,从电子管、晶体管到中小规模集成电路以及大规模超大规模集成电路。软件方面,从机器语言、高级语言、到数据库技术、模拟技术、操作系统、网络技术和集成技术等。其效果是逐步把世界的科学信息资源变为可利用的公共财富。计算机的应用已渗透到各个学科,并从单一的进行运算发展到对整个行业和学科工作方法的根本变化,例如:计算机辅助设计、科研、教学、过程控制等,产生了巨大的效益。在研究的方法论上也不断有新的突破,改变着科学研究的面貌。

化学、化工是一门以实验为基础而发展起来的科学。化学信息资源正以空前的速度在增长,仅登录的化合物数目,目前已超过 1150 万种,并以每年约 60 万种新化合物的速度在增长。精确地测定了三维结构(分子中原子坐标)的化合物也已超过 10 万种。现代化学工业尤其是精细化工、石油化工、生物化工发展特别迅速,其开发、设计、生产、控制、管理大都以计算机为重要手段。在基础化学研究中,对分子水平上的化学结构研究,如量子化学、结构化学都已经完全离不开计算机技术。“实验”虽然仍是化学的基本方法,但“实验”一词的含意已经突破了传统概念,“计算机实验”亦属于广义的“实验”的概念,正被越来越多的化学家所接受。

随着计算机在化学化工中应用水平的提高,出现了一系列带有边缘性的,并具有独立特征的研究领域。如:计算方法,模型化方法,模拟技术,物质共性规律,结构和性质的关系研究,数据库技术,化学信息和化学结构的描述、处理、解析、分子图形技术,化工冶金系统流程模拟及计算机辅助设计(CAD)技术,化学化工中的人工智能和专家系统等。这些新兴领域就大体上形成了化学化工中的一个新的学科分支——计算机化学。

## 二、学 科 特 征

计算机化学是一门从生产实践中自然发展起来的新兴学科,它目前仍处在迅速发展的过程中,匆忙地、人为地划定其学科边界或给出简单的定义显然是不明智的。但另一方面,它作为一个学科确实已经初步成形。本文只就我们的认识,扼要讨论一下其学科基本特征。目的是为我国发展这一新兴学科的政策提供一些依据。

### 1. 计算机化学是一门交叉性很强的边缘学科

计算机化学的基础有三个方面：(1) 基础化学；(2) 工程化学；(3) 计算机科学。如果说基础化学各学科和工程化学各学科是其“经”，计算机科学中的若干新技术则是“纬”。通过交叉联系和渗透，某一学科领域内的新技术新成果可以通过计算机化学迅速传递到化学的各学科领域中去，而化学领域中的许多新理论、新成果可以通过计算机化学及时地在各其它学科或工程领域得到应用和发展。

### 2. 大量化学信息的计算机处理形成了计算机化学的鲜明特征

按照我国目前对基础研究范畴的提法，基础研究包括三个方面：(1) 基础理论研究；(2) 应用基础研究；(3) 基础科学数据的收集、整理、评价和有关规律的研究。和其它学科相比，按年度统计化学信息可占自然科学信息总量的 40% 左右。显然，基础化学数据(绝大部分是实验结果和多年积累的宝贵资料)的研究是至关重要的。化学家用通用的简洁语言——化学分子式、结构式、化学反应所表达的化学知识，化工过程中的各种数据和图形，这些都是化学信息的核心部分，把这些信息系统地计算机内表达和存贮，方便地实现各种指定目的的提取和加工是计算机化学各领域中的共性问题。尤其是在原子、分子水平上不同层次的化学结构信息处理方法，当然地成为计算机化学的研究重点。

### 3. 产生高技术软件产品是计算机化学的又一个突出特征

一般基础研究，多仅以文章形式发表成果，但计算机化学除了文章以外，还有系统的软件产品。以化工流程的模拟技术为例，一系列做为化工流程设计必备工具的软件产品如 ASPEN-PLUS PROCESS 等已做为商品出售。软件产业在全世界已经逐步形成。与化学、化工有关的计算机应用软件和开发工具是典型的高技术产品，在把科学技术成果转化为直接生产力的过程中起着关键的作用。由于计算机化学的研究工作和高技术软件产品的开发密不可分，使得这一新兴学科分支既具有基础研究的特色，又具有巨大的社会和经济效益前景。

以上列举了计算机化学的三个重要特征。认识到它的存在和发展前景，对制定学科发展的技术政策是重要的。

## 三、一些重要领域的国内外概况

### 1. 各类化学数据库的研制

质谱、色谱、红外谱、核磁谱、光电子能谱等在医学、药物、生物化学、地球化学和材料科学等方面的研究、分析中有广泛应用。这些谱图的数量极大，如何提高所得数据的使用水平；如何快速地、正确地解释这些数据，这是计算机应用的一个广阔领域。谱图库检索已成为一种常用的谱图解析方法，它根据存储在谱图库中的已知化合物的信息(标准谱图、分子式、分子量等)，用一定的程序进行检索，来判别未知化合物的结构。存于谱图库中的所有参考谱图都必须按照一定的格式输入。检索时未知物谱图按程序要求与库中的参考谱图进行比较，同时计算出它们的相似性指数，最后按最佳匹配顺序列出检索结果。我院上海有机化学所的红外光谱信息综合

处理系统 (CISOC-IR), 长春应化所的碳-13 核磁共振波谱数据库系统 (GIAG), 化学所的质谱数据库 (ICMSIS), 北京大学和我所与美国 NIST 合作的光电子能谱数据库 (ESCADB) 等在程序处理方面与系统功能上都有独到之处, 为我国科研、工业和国防部门提供了良好的查询和解释工作, 有着十分显著的实际效果。

化学物质中哪些是有毒的? 毒性如何? 这些都是人们密切关注的问题。据统计目前有毒化学物品年产量已达 400 万吨。人们迫切要求制定污染物的排放标准、卫生标准、防治技术和管理措施。毒性数据库的建立可以提供有毒化合物的监测数据、分析数据、毒性化合物的环境标准和法规等。同时通过化学结构的数据计算和处理, 可为解释化合物的致毒机理提供科学的数据, 可以部分地替代昂贵的动物实验。RTIECS 是美国职业安全与卫生研究所 (NIOSH) 根据美国 1970 年制定的职业安全与卫生法的有关规定, 从数千种科学文献中收集、整理、评价的毒性数据, 分别编制成文字版、缩微版和计算机磁带出版发行。从 1971 年开始, 每年修订增补一次, 每年增加 5 至 6 千种化学物质, 到 1984 年版已收集有毒性数据的化学物质 7.2586 万种。中国科学院生态中心在美国提供的 SANSS 与 RTECS 磁带作数据源的基础上, 建立了化学物质毒性数据库联机检索系统 (CTSS)。该库已在检测食品添加剂、农药等毒性中发挥重要作用。

热力学、相平衡与动力学是物理化学学科中的重要分枝, 长期以来各国化学家们在该领域测定的数据成千上万, 发表的理论与公式也极为丰富。因此建立该学科各分枝的数据库也是各国计算化学家们有兴趣的课题。他们除了收集与鉴审已发表的数据, 而且还利用计算机进行数据规律性的探索和物性估算模型的研究。有机化学中 UNIFAC 模型的创建, 是以德国的 Dortmund 大学的相平衡数据为基础建立的功能团活度系数模型, 它在化工热力学研究中具有划时代的意义。此类数据库的建立, 存贮与检索数据仅是其基本的工作, 但更多的是计算程序部分。例如在多相多组元复杂体系化学平衡计算问题上, 没有计算机时只能处理一些简单的体系。而现在已有了许多新的计算方法, 能够处理真实体系中各种复杂状况下的化学平衡计算问题。在相图计算研究中, 利用二元相图计算出活度系数模型参数, 进一步推算出多元系相图, 也有了很好的苗头。中国科学院化工冶金所自 1979 年以来已经先后建立了无机热化学数据库 (ITDB)、非电解质汽液相平衡数据库 (NEDB)、水溶液热力学数据库 (ATDB) 等。并且已在科研、生产、流程模拟以及教学等方面起到很好的作用。

## 2. 分子设计的研究

人们期望制造出具有某种指定性质的新分子、新农药和新材料。利用计算机对复杂结构的物质进行计算机辅助分子设计是 80 年代以来引人注目的课题。由于科学家们的卓有成效的工作使得计算机的功能迅速增强, 许多工作证实了计算机在辅助分子设计方面起着颇为重量的作用。人们所期望的 CAD/CAM (即计算机辅助设计和制造) 应用于医药生产和化合物制造的条件也日渐成熟。在药物作用研究的理论中, 应用量子化学的电子结构理论来研究药物分子外围的电子云分布等, 可以帮助分析药物的构效关系。随着分子轨道理论进一步有效的近似和电子计算机的应用, 量子药理学在短短十余年来发展迅速, 对于新药设计起到了重要的指导作用。北京大学、吉林大学、北京医科大学和中科院的上海药物所、化冶所、大连化学物理所、上海有机化学所等单位对分子静电势、生物活性分子的构象、药物的定量构

效关系和大分子的量子化学计算等研究方面已有很好的工作。用于分子力学计算的 MM2 和 MMP2 程序包,用于从头算分子轨道的 Gaussian90 程序包,目前都很通用,是分子设计计算中的重要组成部分。Gaussian 程序包从 1970 年到现在已有五个版本。70 年的版本只能限制处理 70 个轨道 35 种原子,无 d 轨道。76 年的版本就可以处理 d 轨道,它在不断地改进和完善,并增加新的内容和功能。1987 年北京大学和中国科学院化冶所联合引进了英国剑桥结晶数据库系统的磁带,已在 VAX-11/780 机器上运行,为我国分子设计的研究提供了一种重要的资源。1988 年起又引进了德国无机晶体结构数据库,这是进行材料设计的重要资源。

与此同时,计算机分子图形技术的研究也高速兴起。世界著名的德国的 MERCK 药厂,在 80 年代初期建立了“Merck Molecular Modeling System”。该系统利用计算机绘图技术可以帮助化学工作者绘制三维分子结构的图形、确定其最优构象,比较具有相似生物活性分子的几何形状的异同,从而为设计具有更好疗效的新药品提供模型并开拓新产品。又如可以支持麻醉镇痛药芬太尼类化合物的结构分析研究,从三维分子图形直观地了解和研究芬太尼类麻醉镇痛药各活性中心的结构特点,并进行多个具有不同活性药物分子结构的拟合,为活性与结构关系的分析研究提供有力的论据和基础。英国的 Chemical Design Ltd,和西德的 BEILSTEIN 研究所在计算机分子结构三维显示与分子设计的研究方面已有较成功的软件产品。据美国《科学新闻》报导:英国科学家使用计算机绘制了艾滋病毒产生的蛋白酶的三维结构,这种结构在研制新的抗艾滋病毒药物方面可能是重要的。中国科学院电子所在 IBM-PC/AT 机上设计并完成了高性能的分子模型微机系统,给化学、生物学、医疗等学科的科技工作者提供了简便而强有力的工具。东南大学、南开大学也都有类似产品研制成功。

### 3. 化学模式识别研究

模式识别 (Pattern Recognition) 是基于“物以类聚”的简单道理。目前已发展了一系列的数学方法来对信息进行处理,从而将物群进行分类,确定某物所属类别,了解各因素之间的关系,判别哪些因素对物群分类起着有效的影响。其最大的特点是可处理多因素问题。判别分类工作在科学研究和生产实践中是相当广泛的,计算机模式识别技术的发展能使人们在影响因素很多的情况下根据模式识别的法则对众多的信息进行处理、选择并应用决定性特征对物群进行分类。

国外模式识别用于化学主要是在分析化学和构效关系研究方面。我国在化学的另外一些领域中,也运用了模式识别这一有效工具。如:计算机辅助无机合成,对未知化合物实现了计算机预报,在预报基础上合成并发现了 LaPd、EuNi、EuFe、PrPd 等一批新化合物;用模式识别分析活检气相裂解物色谱、微量元素谱,从而辅助癌症早期诊断;用模式识别进行化合物分类,在有机化合物 C-13 核磁共振谱中应用;研究中国古瓷中的微量元素等。

中国科学院上海冶金所将计算机用于有色金属探矿、石油勘探,工业调优和宏观化学动力学研究等方面获得了许多研究成果,中国科学院长春应化所、中国科技大学等单位也开展了有关的研究工作。

### 4. 分析仪器智能化

各种分析仪器与计算机联接,进行实验数据的采集、处理、数模转换和模数转换,是分



析仪器的重要发展方向。各种分析仪器不尽相同,需完成的智能化程序显然不同,但其中不少数据处理的方法,如富里叶变换、解线性方程组和非线性方程组、拓扑法、相关与回归分析、偏最小二乘法、卡尔曼滤波方差分析检验以及最近发展起来的神经网络方法,都具有一定的通用性。近年来在国内外分析化学杂志上有关该方面的研究报告也很多。我国大连化物所在发展带有智能的色谱仪方面,已取得很多成绩。与此同时,一些谱图分析仪器与智能化的大型数据库联网也是一个重要的方面,这种将大型分析仪器联入计算机智能化系统,今后将会有更进一步的发展。

### 5. 计算机辅助化学教育

利用计算机图形学技术将各种高级分析仪器计算机化。如:质谱仪、红外谱仪、电子能谱仪等等,从而可以用来培训学生或有关操作人员进行模拟操作,以便尽量减少实地操作过程引起的事故。还有,利用计算机进行化学实验模拟,这是学生在进实验室实践之前进行的一种计算机模拟练习。还可以将学生经常接触到的概念、原理、计算等问题程序化,从而使学生通过上机可以主动地、逐步地建立概念,进行系统总复习等。此外,程序系统可以及时反馈出该学生在掌握基本概念、计算技能、综合应用基础知识等方面的情况并加以评价。此项工作在我国刚刚开始不久,但在第二次、第三次全国计算机化学会上,一些学校的计算机化学专家、教授,已经联合起来进行了开发。在化工方面,利用计算机培训工长,也是一个重要的发展方向,我国一些部门现在也已经开始进行这方面的工作了。

### 6. 化学中的人工智能、专家系统

人工智能(Artificial Intelligence)是用计算机模拟人类的智能行为,试图用计算机代替人的部分脑力劳动。专家系统(Expert System)是人工智能在专业领域中的表现形式,它是用计算机模拟某一学科的专家(或某一位专家)处理专业问题的思维、推理方法,作出专家水平的判断和决策,是试图用计算机代替或部分代替专家工作的软件系统。最早的专家系统是1969年由美国 B.Buchanan 提出的 DENDRAL 系统,该系统是用来帮助有机化学家通过质谱分析确定化学结构。近年来国外在合成路线设计、生物化学和生化系统模型、化工控制过程等专家系统方面均有不少新的研究进展。在1991年第四届国际化工大会上,专家系统的文章遍及各个行业分枝。

### 7. 模拟与优化

化学反应、化学热力学、化学动力学以及某化工过程的计算机模拟的研究工作在国内外已有较长的时间了。这为选择最佳实验条件、了解各元素或物质在反应过程中的行踪以及热力学可行性、经济可行性等研究提供了有效的工具。它对某一具体工业过程的系统模拟或现有生产体系革新挖潜;新过程开发及新设计研究;过程控制和局部系统及大系统优化研究都有帮助。它不仅对改造落后厂的挖潜有用,而且对新流程的改进也具有重要的作用。

模拟与优化的理论基础是建立在化学反应、分离工程科学之上的,但是没有模拟与优化方面的智能化程序及数据库这两座金桥,理论工作与实际应用是难以结合起来的。

近年来,在机器制造行业发展起来的 CIMS (管理制造系统一体化)技术,在化工过程工

程中也得到了广泛的应用,这意味着将在更为广义的体系内,实现化工过程的优化。

#### 四、若干前沿领域

根据国内外资料,当前计算机化学学科研究的前沿可大致归纳为十个方面:

1. 基础化学数据的采集、评价、检索和通讯网络;
2. 化学结构信息的计算机描述及处理方法研究;
3. 化学数据库;
4. 计算机化学中的人工智能和专家系统;
5. 结构—性质和结构—生物活性关系研究、分子设计、材料设计;
6. 3D-分子图形学;
7. 量子化学、分子力学、分子动力学计算方法;
8. 化学计量学和分析仪器智能化;
9. 化学工业中的过程模拟、控制和优化;
10. 计算机辅助化学教育。

就国内情况来看,有些方面我们已经做了不少工作;其中一些工作达到了国际先进水平;有些方面尚属起步阶段,正在做准备;另有一些方面尚未开展。因此,有必要针对我国国情,加强组织规划和协调,促进计算机化学学科在我国的稳步发展。一方面打好基础,一方面组织力量在国家特别急需的一些方面制定具体的有限目标,把研究成果和开发的软件产品推广到科研、设计和生产部门,使其尽快地在生产实践中发挥威力,促进我国现代化的进程。

#### 五、队伍组织和政策性建议

综上所述,我们认为把计算机化学做为一个交叉性边缘学科纳入院、科委和国家级的发计划是适宜的。

从 70 年代末至今,十多年来我国在该领域中的研究已有良好开端。中国化学会下属的计算机化学专业委员会于 1987 年成立。中国化工学会下属的计算机化工专业委员会于 1986 年成立。全国性计算机化学、化工学术会议已经各召开了三次,并于 1987 年 6 月在北京召开了第八届计算机在化学研究与教育中的应用国际会议(ICCCE)。与会成员包括 18 个国家的有关专家学者。1990 年在杭州召开的全国第三届计算机化学学术讨论会参加人数 220 人,论文 280 篇。无论是与会人数和论文水平都比前几次有大幅度增加和提高。由中国科学院化工冶金所和上海有机化学所联合成立的计算机化学开放实验室也于 1989 年正式成立,在奠定基础、培养人才、国际交流和开展研究方面都有一定进展。

从当前国内科研形势看,由于科研经费紧缺,如何稳定已有的研究基础,稳定目前已有一定水平和素质的队伍,特别是青年骨干力量是当务之急。

为此,我们呼吁院领导和科委领导在规划、组织方面对于这一新兴学科分枝不失时机地给予必要和可能的支持,由国家投资开发一批基础应用软件。这样做无疑对我国科研现代化将有重要和深远影响。