

· 研究论文 ·

# 催化剂分子设计的通用数据库与用户界面<sup>①</sup>

方宁 孙杰 江扬波 易军 廖代伟<sup>②</sup>

(厦门大学物理化学研究所, 厦门大学化学系 固体表面物理化学国家重点实验室 厦门 361005)

**摘要** 用 Visual C++ 语言编写了催化剂分子设计专家系统的通用数据库, 应用面向对象编程技术, 使用主数据库结构信息封装类, 子数据库结构信息封装类和数据文件封装类封装了数据库信息。数据的输入输出则是通过将重载操作符设定为数据封装类的友元函数来实现的。这一方法实现了数据库的程序无关性, 方便了数据库的扩充与更新, 并设计成窗口式通用界面, 提高了程序的易用性。

**关键词** 数据库, 催化剂分子设计, 界面, 类

中图分类号 O 643.3 O 23

## UNIVERSAL DATABASE AND USER INTERFACE FOR CATALYST DESIGN

FANG Ning SUN Jie JIANG Yang-Bo YI Jun LIAO Dai-Wei

(*Institute of Physical Chemistry, Department of Chemistry, State Key Laboratory for Physical Chemistry of the Solid Surfaces, Xiamen University, Xiamen 361005*)

**ABSTRACT** An universal database for expert system of catalyst design was developed using Visual C++ language. Based on the object-oriented programming technology, the information of database was encased into the structure-information-encapsulation class of main database, the structure-information-encapsulation class of sub-database, and the encapsulation class of data files respectively. The over-loading operator was defined as a friend function of the database encapsulation class for the practice of data input/output. It is ease-to-use and convenient for the expansion and updating of database via the general window-type user interface.

**KEYWORDS** Database, Catalyst design, Interface, Class

利用计算机大容量信息处理和人工智能的功能, 结合催化科学的特点, 根据催化反应机理和活性结构的系列模型, 模仿专家进行过程设计的思维方式, 实现催化剂的计算机辅助设计, 是催化领域的前沿课题。它不仅可以用于指导实用催化剂的开发, 而且可以解释或验证催化机理和活性结构的合理性。计算机辅助催化剂分子设计的这种导向作用已引起了催化界的广泛重视。70年代以来, 国内外许多学者<sup>1-11</sup>都致力于催化剂设计的研究, 力图改变炒菜式人工调试催化剂的传统方法, 以缩短催化剂的开发周期。人工智能中的专家系统方法为计算机辅助催化剂分子设计提供了一个很好的工具。

我们<sup>15-11</sup>在这一方面已做了大量工作, 并针对甲烷氧化偶联催化反应初步开发出工作于DOS环境下的小型催化剂分子设计专家系统(ESMDC)。最近, 我们又使用Visual C++5.0程序语言, 实现了数据库的程序无关性, 极大地方便了数据库的扩充与更新, 并设计成窗口式通用界面, 提高了程序的易用性。

## 1 设计思路

专家系统的基本设计结构与一般应用程序存在着根本性的区别。一般应用程序将其知识组织成两级,

收稿日期: 1998-12-11

① 国家自然科学基金资助项目(编号: 29773037)

② 通讯联系人

方宁, 男, 22岁, 导师: 廖代伟, 化学专业

即数据级和程序级。专家系统则将知识组织成三级，即数据库级、知识库级和控制级。专家系统的主要特征是有一个巨大的知识库。我们的设想是设计出一个具有相当通用性的系统。虽然，针对不同类型反应的知识库的具体条件可能完全不同，但是，每个知识库的组织方式却存在着相当多的共同点。因此，只要有足够完善的数据库，再配合适当的搜索规则，就可能实现一个通用系统。根据这一思路，我们首先设计出一个数据库系统，使其能方便地实现数据的添加与修改，为实现一个较为完善的数据资料库做好软件准备。而后设计了知识库生成系统，可根据用户的要求生成知识库，由此实现对不同类型反应的催化剂设计。最后是设计一个知识库综合设计系统，根据用户生成的知识库设计出符合条件的可能催化剂，并进行综合调整和评估。本文将主要介绍通用数据库的结构及其用户界面。

## 2 数据库结构信息封装类

主数据库体系由一系列子数据库构成，每个子数据库则由数据结构信息和数据文件组成。我们使用 Visual C++ 5.0 程序语言，应用面向对象编程技术，用类 (Class) 封装了数据库信息。主数据库结构信息封装类如下所示：

```
class CMainLib; // 预定义
ifstream & operator >>(ifstream &, CMainLib *);
ofstream & operator <<(ofstream &, CMainLib *);

class CMainLib // 类名为 CMainLib
{
    CString m - LibName; // 子数据库名
    CString m - FileName; // 子数据库存档文件名
    CString m - Description; // 子数据库描述

public:
    void SetLibName(CString s) // 读取及设定数据库信息的
    void SetDescription(CString s) // 函数组
    void SetFileName(CString s)
    CString GetLibName(void)
    CString GetFileName(void)
    CString GetDescription(void)

    friend ifstream & operator >>(ifstream &, CMainLib &);
    // 输入重载 >>
    friend ofstream & operator <<(ofstream &, CMainLib &);
    // 输出重载 <<

    CMainLib( ); // 构造函数
    virtual ~CMainLib( ); // 析构函数
};
```

子数据库结构信息封装类如下所示：

```
class CSubLib
{
    CString m - strItemName; // 数据项名称
    CString m - strDescription; // 数据项描述
    CString m - strItemUnit; // 单位
    CString m - strLength; // 数据总长度(字符串)
    int m - intLength; // 数据总长度(整数)
```

```

CString m -strDec;           // 小数点位数(字符串)
int m -intDec;              // 小数点位数(整数)
CString m -strType;        // 数据类型 (字符串)
int m -intType;            // 数据类型 (整数)

public:
void SetItemName(CString s) // 以下为读取及设定数据库
                           // 信息的函数组

void SetItemUnit(CString s)
void SetDescription(CString s)
void SetLength(CString s)
void SetDec(CString s)
void SetType(CString s)
void SetIntType(int i)
CString GetItemName(void)
CString GetItemUnit(void)
CString GetDescription(void)
CString GetLength(void)
CString GetDec(void)
CString GetType(void)
int GetIntType(void)
int GetIntDec(void)
int GetIntLength(void)

friend ifstream & operator >>(ifstream &, CSubLib &);
                           // 输入重载>>
friend ofstream & operator <<(ofstream &, CSubLib &);
                           // 输出重载<<

CSubLib();                  // 构造函数
virtual ~CSubLib();         // 析构函数
};

```

单个数据记录用类 CData 封装, 再用类 CDataLib 封装一个数据库中的所有数据记录. 数据文件封装类 (CData 类、CDataLib 类) 的说明如下:

```

class CData
{
public:
// 成员函数
int * m -pLength;
CString m -strALine;

// 成员函数:
CData( );                  // 构造函数
virtual ~CData();         // 析构函数

friend ifstream & operator >>(ifstream &, CData *); // 输入重载>>
friend ofstream & operator <<(ofstream &, CData *); // 输出重载<<
};

```

```

class CDataLib
{
// 成员函数
    int m - intLength;           // 单个数据记录的长度
    CString filenameDat;        // 数据库存档文件名
    CString m - strLibName;     // 数据文件名
    CData * m - data[ 2000];    // 数据记录
                                // (假定记录数不大于 2000)
    int m - intDataNum;         // 数据记录个数
    int m - intItemNum;        // 数据项个数
    int m - intPosI[ 100];     // 每个数据项的位置
    int m - intTypeI[ 100];    // 每个数据项的类型
    int m - intLengthI[ 100];  // 每个数据项的长度
    CSubLib * subLib[ 100];    // 数据库结构
                                // (假定数据项数不大于 100)

// 成员函数:
    BOOL InitData( void);      // 初始化
    CString GetALine( int);    // 取得某一数据记录
    void SetAStrItem( int, int, CString); // 写某一数据记录
    CString GetAStrItem( int, int); // 取得某一数据记录的
                                    某一数据项(字符串)
    int GetAnItem( int, int); // 取得某一数据记录的
                                    某一数据项(整数)

    CDataLib( );              // 构造函数
    virtual ~CDataLib();     // 析构函数
};

```

数据库数据有字符串 (CString)、整数 (Integer)、浮点数 (Float) 和布尔型数 (Boolean) 等 5 种类型, 根据不同需要可选用不同的数据类型。数据项长度对字符串而言即为其实际长度, 对单字符而言为 1, 对整数为其位数, 对浮点数为其整数及小数位数之和并包括小数点, 对布尔型数也为 1, 数据项描述是对该数据项的说明性文字, 小数位数只对浮点数有效。

### 3 数据的输入输出

数据的输入输出是通过将重载操作符  $\gg$  与  $\ll$  设定为数据封装类的友员函数, 从而实现数据的读取。例如:

```

class CMainLib           // 类名为 CMainLib
{
    ..... // 无关项省略
public:
    friend ifstream & operator >>(ifstream &, CSubLib &); // 输入重载 >>
    friend ofstream & operator <<(ofstream &, CSubLib &); // 输出重载 <<
    // friend 表示将函数设定为友员
};
ifstream & operator >>(ifstream & inst, CMainLib * t)
{

```

```
char s[2000] = "\0";
```

```

instr.getline(s, 199, '@'); // 将“@”设定为边界
if ((s[0] < 20) && (s[0] > -1)) {
    s[0] = '\0';
    t->SetLibName(s);
    return instr;
}
t->SetLibName(s);
instr.getline(s, 199, '@');
t->SetFileName(s);
instr.getline(s, 199, '@');
t->SetDescription(s);
instr.getline(s, 199);
return instr;
}

ofstream & operator << (ofstream & outstr, CMainLib * t)
{
    outstr << t->GetLibName() << '@' << t->GetFileName()
        << '@' << t->GetDescription() << '@\n';
    return outstr;
}

```

其中, `ofstream` 为输出流类, `ifstream` 为输入流类, 主数据库和数据结构数据库的数据间以“@”间隔, 具体数据库的数据间以“|”间隔。

## 4 用户界面

选择 Windows 主界面 File 菜单中的选项 DataBase 项进入数据库操作。主数据库储存数据库的总体信息, 包括数据库的名称、描述及储存的文件名。主数据库界面中, Pre/Next: 上一个/下一个数据库。Append: 在主数据库末尾添加一个数据库。Insert: 在主数据库当前位置添加一个数据库。Delete: 在主数据库当前位置删除一个数据库。Save Change: 保存修改过的主数据库。Modify: 修改当前数据库的名字、描述及所指向的文件。Edit Item: 修改当前数据库的数据结构及添加修改数据记录。Close: 关闭本窗口。

数据结构的修改界面中, ItemList: 从此列表中选中待修改的数据项, 再进行添加、删除或修改操作。ItemUnit: 数据的单位。Type, Length, Dec: 类型, 长度, 小数位数。Undo: 取消所做的操作。Open Lib: 浏览数据, 进行添加、修改。

在数据浏览界面中, 单击“Modify Data”则进入数据修改、添加界面。其中, Input Data: 显示、输入数据的窗口。Select Item: 选择数据项。|<, >|: 移动到第一个/最后一个记录。<<, >>: 快速移动, 向前/后移动 10 个记录。<, >: 向前/后移动 1 个记录。Move to: 在方框中填入要移往的记录号, 再按此按钮即可移动指针, 其他情况下此处显示当前的记录号。

知识库的组织方式与数据库类似, 并可由使用者动态创建和修改, 详细情况将另文报道。

### 参 考 文 献

- 1 廖代伟. 催化剂设计浅谈. 化学通报, 1982 (3): 44-51 & 38  
Liao D W. Introduction to catalyst design. Chemistry, 1982, (3): 44-51 & 38
- 2 Trim D L. Design of Industrial Catalysts. New York: Elsevier Sci, 1980
- 3 Hegedus L L. ed. Catalyst Design—Progress and Perspectives. New York: John Wiley & Sons, 1987
- 4 Graziani M, Rao C N R. ed. Proceedings of the Workshop on Catalyst Design Advances in Catalyst Design. New Jersey: World Sci Pub, 1991
- 5 Liao D W, Huang Z N, Lin Y Z, Wan H L, Zhang H B, Tsai K R. Computer-aided molecular design of catalysts based on mechanism and struc-

- ture. *J Chem Inf Comput Sci*, 1996, 36 (6): 1178-1182
- 6 廖代伟, 龚文华, 李堂秋, 林银钟, 郭峰, 蔡俊修, 万惠霖, 蔡启瑞. 金属氧化物催化剂分子设计的物理化学判则和数学模型. 厦门大学学报 (自然科学版), 1994, 33 (Sup): 190-197  
Liao D W, Gong W H, Li T Q, Lin Y Z, Gou F, Cai J X, Wan H L, Tsai K R. Physical and chemical rules and mathematical models for molecular design of metal oxide catalysts. *J Xiamen University (N S)*, 1994, Sup: 190-197
- 7 黄遵楠, 廖代伟, 张鸿斌, 万惠霖, 蔡启瑞. 催化剂分子设计专家系统 ESMDC 及其在甲烷氧化偶联催化剂设计中的应用. 计算机与应用化学, 1994, 13 (3): 167-171  
Huang Z N, Liao D W, Zhang H B, Wan H L, Tsai K R. Expert system of molecular design of catalyst (ESMDC) and its application in design for oxidative coupling of methane. *Comput Appl Chem*, 1996, 13 (3): 167-171
- 8 黄遵楠, 廖代伟, 万惠霖, 张鸿斌, 蔡启瑞. 计算机辅助催化剂组分子设计的思想和实现. 计算机与应用化学, 1997, 14 (1): 12-16  
Huang Z N, Liao D W, Wan H L, Zhang H B, Tsai K R. Computer-aided molecular design for catalyst components. *Comput Appl Chem*, 1997, 14 (1): 12-16
- 9 廖代伟, 黄遵楠, 林银钟, 龚文华, 万惠霖, 张鸿斌, 蔡启瑞. 甲烷氧化偶联金属氧化物催化剂组份分子设计程序. 厦门大学学报 (自然科学版), 1996, 35 (1): 133-136  
Liao D W, Huang Z N, Lin Y Z, Gong G W, Wan H L, Zhang H B, Tsai K R. Molecular design program for catalyst composition of oxidative coupling of methane. *J Xiamen University (N S)*, 1996, 35 (1): 133-136
- 10 黄遵楠, 廖代伟, 林银钟, 张鸿斌, 万惠霖, 蔡启瑞. 催化剂分子设计专家系统中推理机的设计. 厦门大学学报 (自然科学版), 1996, 35 (3): 389-392  
Huang Z N, Liao D W, Lin Y Z, Zhang H B, Wan H L, Tsai K R. Realization of inference engine in expert system for molecular design of catalyst. *J Xiamen University (N S)*, 1996, 35 (3): 389-392
- 11 黄遵楠, 廖代伟, 林银钟, 张鸿斌, 万惠霖, 蔡启瑞. 催化剂分子设计专家系统中数据库和知识库的建立. 厦门大学学报 (自然科学版), 1996, 35 (5): 739-744  
Huang Z N, Liao D W, Lin Y Z, Zhang H B, Wan H L, Tsai K R. Construction of data base and knowledge base in expert system for molecular design of catalyst. *J Xiamen University (N S)*, 1996, 35 (5): 739-744

(上接第 104 页)

- 11 姚建华, 袁身刚, 陈海峰, 郑崇直, 杨铄. 三维分子结构检索系统的结构索引与匹配. 计算机与应用化学, 1999, 16 (2): 97-100  
Yao Jinhua et al. Structure indexing and matching in 3-Dimensional structure searching system. *Computers and Applied Chemistry*, 1999, 16 (2): 97-100
- 12 Yang Yuliang, Zhang Hongdong. Monte Carlo methods in polymer science. Shanghai: Fudan University Press, 1993
- 13 Yuan Shengang, Luo Shiwei, Cai Guoqin, Yao Jianhua, Chen Haifeng, Zheng Chongzhi. From pharmacophore to leads: Bioactive compounds discovery with computer-aided techniques. *J Chinese Chemistry*, 1998
- 14 Liu Yong, Kang Lishan, Chen Yuping. No-numerical Parallel Algorithms (No. 2) -Genetic Algorithm. Beijing: Science Press, 1995
- 15 The database of no cultivated compounds was developed at the laboratory. It has collected 25 000 3D structures of recently reported new compounds. For each compound only one favorite conformer with lower energy was stored. Most of compounds are from nature products.
- 16 Wang Shaomeng, Zaharevitz Daniel W, Shama Rajiv, Marquez Victor E, Lwein Nancy E, Du Linh, Blumberg Peter M, Milne G W A. The discovery of novel structurally diverse protein kinase C agonists through computer 3D-database pharmacophore search molecular modeling studies. *J Med Chem*, 1994, 37: 4479-4489
- 17 Wender Paul A, Cribbs Cynthia M, Koehler Konrad F, Sharkey Nancy A, Herald Cherry L, Yoshiaki Kamano, Pettiet George R, Blumberg Peter M. Modeling of the bryostatins to the phorbol ester pharmacophore on protein Kinase C. *Proc Natl Acad Sci*, 1988, 85: 7197-7201
- 18 Wang S, Milne G W A, Yan Xinjian, Posey Isadora J, Nicklaus Marc C, Graham Lisa, Rice G. Discovery of novel, non-peptide HIV-1 protease inhibitors by pharmacophore searching. *J Med Chem*, 1996, 39: 2047-2054
- 19 Hibert Marcel F, Hoffmann Remy, Miller Robert C, Carr Albert A. Conformation-activity relationship study of 5-HT<sub>3</sub> receptor antagonists and a definition of a model for this receptor site. *J. Med Chem*, 1990, 33: 1594-600