

TRS 管理员工具

TRS Admin 6.5

用户手册

北京拓尔思信息技术股份有限公司

Beijing TRS Information Technology Co., Ltd.

版权说明

本手册由北京拓尔思信息技术股份有限公司（以下简称 **TRS** 公司）出版，版权属 **TRS** 公司所有。未经出版者正式书面许可，不得以任何方式复制本文档的部分或全部内容。

©北京拓尔思信息技术股份有限公司 版权所有。保留所有权利。

TRS 是北京拓尔思信息技术股份有限公司的注册商标。

关于本手册

TRS 管理员工具是 TRS 全文数据库的客户端管理工具，主要用于对各种 TRS 对象进行管理、设计和维护等，并具备完整的检索、浏览和维护功能。本手册详细地介绍了 TRS 管理员工具的使用方法，同时提供了有关检索表达式、数据库加载文件规范等方面的参考信息。

读者对象

本手册的主要读者为 TRS 全文数据库的系统管理员，但普通用户也可以使用。只要基本掌握 Windows 95/NT 操作系统的使用方法就可以很快地学会用 TRS 管理员完成各种 TRS 系统操作。

用户反馈

TRS 公司感谢您使用 TRS 产品。如果您发现本手册中有错误或者产品运行不正确，或者您对本手册有任何意见和建议，请及时与 TRS 公司联系。您的意见将是我们做版本修订时的重要依据。

目 录

第 1 章 总体概述	1
1.1 TRS全文数据库概述.....	1
1.2 TRS管理员工具概述及新增特性.....	2
1.2.1 界面风格焕然一新, 提供标签页式多文档管理, 新增可浮动的导航和属性面板.....	2
1.2.2 新的系统选项对话框, 采用属性表方式显示选项, 每一项均有概要提示	6
1.2.3 树形管理用户组、用户和数据格式, 可筛选所属对象.....	7
1.2.4 属性表式管理各种对象属性, 可直接修改更新, 结构和权限独立出来	8
1.2.5 采用新的文本方式导入/导出对象, 方便修改和将来扩展.....	11
1.2.6 支持生成多级导航树, 概览增加字段排序和记录标志, 细览实现图文混合显示和命中点定位	11
1.2.7 可自定义表单编辑记录界面, 下拉列表提供日期/字符型字段可能值选择	13
1.2.8 改进了装库控制文件生成和结果报告查看对话框.....	14
1.3 启动和退出TRS管理员.....	16
1.4 登录和退出登录	18
1.5 集成环境概述	19
1.5.1 对象管理窗口	20
1.5.2 记录维护窗口	26
1.5.3 文件编辑窗口	28
第 2 章 使用数据库对象	33
2.1 TRS数据库对象概述.....	33
2.2 TRS数据库字段对象概述.....	36
2.3 创建TRS数据库对象.....	40
2.3.1 TRS数据库结构设计	40
2.3.2 TRS数据库建库策略.....	40
2.3.3 TRS数据库创建.....	43
2.4 修改TRS数据库对象.....	50
2.4.1 修改数据库的常规属性	51
2.4.2 修改数据库的结构	54
2.4.3 修改数据库的访问权限	57
2.5 数据库重命名	58
2.6 删除TRS数据库对象.....	59
2.7 数据库/视图检索.....	60
2.7.1 一般检索.....	60
2.7.2 高级检索与排序	61
2.7.3 表达式生成器.....	62
2.7.4 选择排序方式对话框	64

2.7.5 查询索引.....	66
2.7.6 修改检索选项.....	68
2.7.7 检索附加参数.....	70
2.7.8 二次检索和排序.....	71
2.8 浏览和处理检索结果.....	72
2.8.1 概览窗口.....	73
2.8.2 细览窗口.....	74
2.8.3 表单窗口.....	76
2.8.4 查看检索结果属性.....	80
2.8.5 检索结果导出.....	81
2.8.6 发送检索结果.....	83
2.8.7 打印检索结果.....	83
2.8.8 分类统计结果.....	85
2.9 建立和使用导航.....	87
2.9.1 建立导航文件.....	90
2.9.2 查看源文件.....	92
2.9.3 清除导航.....	92
2.9.4 指定显示的概览字段和细览字段.....	92
2.9.5 指定显示的记录范围.....	93
2.9.6 改变记录维护窗口的标题.....	94
2.9.7 设置控制选项.....	94
2.9.8 建立和维护导航树.....	95
2.10 记录维护.....	99
2.10.1 添加记录.....	99
2.10.2 批量删除记录.....	100
2.10.3 批量更新记录.....	101
2.10.4 记录密码.....	103
2.11 维护TRS数据库对象.....	104
2.11.1 数据加载.....	104
2.11.2 结构导入/导出.....	110
2.11.3 数据备份.....	112
2.11.4 数据恢复.....	113
2.11.5 数据库的复制.....	114
2.11.6 复制记录.....	114
2.11.7 数据优化.....	116
2.11.8 索引维护.....	117
2.11.9 数据库解锁.....	118
2.11.10 自动分库.....	119
第3章 使用视图对象.....	121
3.1 TRS视图概述.....	121

3.2 TRS视图字段对象概述.....	122
3.3 创建新的视图.....	123
3.4 修改TRS视图对象.....	128
3.4.1 修改视图的常规属性.....	129
3.4.2 修改视图结构.....	130
3.4.3 访问权限.....	132
3.5 视图重命名.....	133
3.6 删除TRS视图对象.....	134
3.7 加入数据库.....	135
第4章 使用词典对象.....	136
4.1 TRS词典概述.....	136
4.1.1 分词词典.....	137
4.1.2 附加分词词典.....	138
4.1.3 停用词典.....	138
4.1.4 附加停用词典.....	139
4.1.5 稀疏词典.....	139
4.1.6 主题词典.....	140
4.1.7 同义词典.....	141
4.1.8 反义词典.....	142
4.2 创建新的词典.....	143
4.3 修改词典的属性.....	144
4.3.1 修改词典的常规属性.....	144
4.3.2 访问权限.....	146
4.4 词典重命名.....	147
4.5 删除TRS词典对象.....	147
4.6 词典的维护.....	148
4.6.1 索引词典的维护.....	148
4.6.2 主题词典的维护.....	149
4.6.3 同义词典的维护.....	151
4.6.4 反义词典的维护.....	153
4.6.5 保存词典查询结果.....	154
第5章 使用用户组对象.....	156
5.1 TRS用户组概述.....	156
5.2 创建新的用户组.....	156
5.3 修改TRS用户组属性.....	157
5.4 用户组重命名.....	158
5.5 删除TRS用户组对象.....	159

第 6 章 使用用户对象	160
6.1 TRS用户概述.....	160
6.2 创建新的用户.....	162
6.3 修改TRS用户属性.....	164
6.4 用户重命名.....	166
6.5 删除TRS用户对象.....	166
第 7 章 管理数据格式	168
7.1 TRS数据格式对象概述.....	168
7.2 创建新的数据格式.....	169
7.3 修改数据格式的属性.....	170
7.4 重命名数据格式.....	171
第 8 章 TRS管理员的设置	173
8.1 缺省登录.....	173
8.2 集成环境.....	174
8.3 检索维护.....	176
8.4 系统配置.....	178
8.5 知识词典.....	179
附录A 检索表达式构造	181
A.1 运算符.....	181
1、算术运算符 (作用于NUMBER字段).....	181
2、比较运算符.....	181
3、逻辑运算符.....	182
4、重复逻辑运算符.....	182
5、属性运算符.....	182
6、限定运算符.....	184
7、加权运算符.....	185
A.2 等价运算符.....	185
A.3 运算符与字段的关系.....	185
A.4 统计函数.....	187
A.5 检索函数.....	187
A.6 统计表达式.....	189
A.7 检索表达式.....	189
A.8 检索键值.....	192
1. DATE型字段上的键值.....	192
2. NUMBER型字段上的键值.....	194

3. CHAR型字段上的键值	194
4. PHRASE和DOCUMENT型字段上的键值.....	195
A.9 特殊键值	195
A.10 保留字	195
附录B 数据库加载格式规范	198
B.1 数据记录文件	198
1. 标准格式	199
2. 字段内部标识号标记格式	200
3. 字段顺序号标记格式	202
4. 无字段标记格式	203
5. 全文格式	204
6. ISO2709 国际标准格式	205
7. XML格式	205
8. 数据记录文件中注释	207
B.2 批处理数据文件	207
B.3 加载多媒体数据	208
1. 加载DOCUMENT字段的多媒体数据	208
2. 数据格式控制符	209
3. 存储方式控制符	210
4. 加载格式说明	211
5. 加载BIT字段的多媒体数据	213
B.4 控制文件	214
B.5 日期的入库格式	225
1. 通用格式	226
2. 英文格式	226
3. 无分隔格式	227
4. 中文格式	227

第1章 总体概述

1.1 TRS 全文数据库概述

TRS 全文数据库是由北京拓尔思信息技术股份有限公司开发的一个体系结构先进、功能强大的面向文本数据的数据库管理系统，广泛地用于各种企业信息数据库、以及企业信息门户的建设。TRS 全文数据库的核心技术为全文检索，它为各种格式文档的存储、管理和检索提供动力。它的主要功能和性能如下：

1. “零”空间膨胀率（-10%至 20%）。
2. 基于成本优化的查询算法(索引分区技术，多线程并行运算技术，面向全文检索的 Cache 技术)，使得 G 级数据库查询速度达到亚秒级，并支持大量并发用户的同时访问。
3. 多线程设计，支持 SMP 体系结构。
4. 支持跨平台计算：支持 Windows, Linux, Sun Solaris, IBM AIX, HP-UX。
5. 除了提供规范的 C/C++ API 外，还提供 ADO 和 JavaBeans 组件，可以使用符合工业标准的开发工具进行快速开发(ASP 和 JSP)。
6. 采用开放的数据存贮管理机制，通过 TRS for RDBMS Gateway，可以透明地连接 oracle、db2、sql server、sybase 和 informix 五个主流的数据库管理系统。它是电子商务应用中全文检索引擎的最佳解决方案。
7. 领先的中文信息处理技术：内嵌汉语自动分词系统，支持按词索引、按字索引、按关键词索引、二元组索引，大大提高检索的准确性和响应时间。
8. 检索信息快、准而且全：允许使用文中的任意字、词、句和片段进行检索，提供了基于文献内容而不仅仅是文献外部特征的全文检索手段。TRS 所提供的按词和按用户自定义关键词进行索引和检索，以及基于知识词典的扩展检索功能，满足了特殊应用领域的高查准率和高查全率的要求。
9. 检索功能强大：全方位检索手段，提供了多达 48 种检索运算符。包括外部特征与正文内容的各种逻辑组合检索、位置检索、二次检索、渐进检索、历史检索、词根检索、大小写敏感检索、概念检索、对检索结果按与检索表达式的相关性和重要性程度排序等。

10. 安全性好: 具有系统级、数据库级、记录级和字段级四级安全控制机制。
11. 提供功能强大的 TRS 管理工具: 数据库维护、数据库的安全性控制、数据的完整性和一致性、数据库备份和恢复、用户管理等功能。

1.2 TRS 管理员工具概述及新增特性

TRS 管理工具是 TRS 全文数据库的客户端管理工具, 主要用于对各种 TRS 对象进行管理、设计和维护等, 并具备完整的检索、浏览和维护功能。在上一版本的基础上, 又作出了以下几个方面的重大改进:

- 界面风格焕然一新, 提供标签页式多文档管理, 新增可浮动的导航和属性面板
- 新的系统选项对话框, 采用属性表方式显示选项, 每一项均有概要提示
- 树形管理用户组、用户和数据格式, 可筛选所属对象
- 属性表式管理各种对象属性, 可直接修改更新, 结构和权限独立出来
- 采用新的文本方式导入/导出对象, 方便修改和将来扩展
- 支持生成多级导航树, 概览增加字段排序和记录标志, 细览实现图文混合显示和命中点定位
- 可自定义表单编辑记录界面, 下拉列表提供日期/字符型字段可能值选择
- 改进了装库控制文件生成和结果报告查看对话框

1.2.1 界面风格焕然一新, 提供标签页式多文档管理, 新增可浮动的导航和属性面板

界面右侧是一个多标签页的窗口设置, 已打开的窗口以标签页的方式显示, 想要查看某一窗口, 只需要点击上部的窗口标签或者点击最右侧的“活动文件”按钮在激活窗口列表中选择一个窗口即可进行切换, 同时相应的菜单、工具栏和属性、导航窗口也进行相应的切换。在标签页顶部点击右键会弹出“关闭当前”和“关闭所有”两个菜单, 点击“关闭当前”即关闭当前激活窗口, 点击“关闭所有”则关闭除对象管理窗口以外的所有窗口。

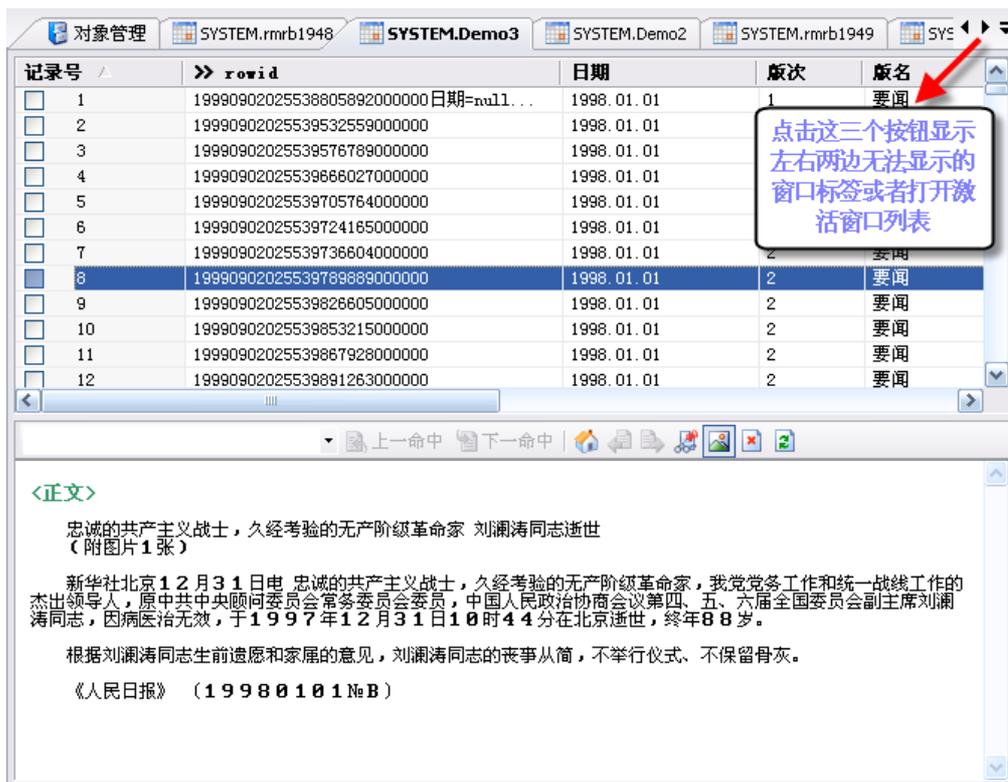


图 1—1 标签页式的多窗口显示

本版本中，增加了两个新的窗口——导航窗口和属性窗口。导航窗口用来显示用户及用户组列表、数据格式列表和某一对象对应的检索导航树。属性窗口用来显示各个对象，字段的具体属性，并能完成修改及更新功能。

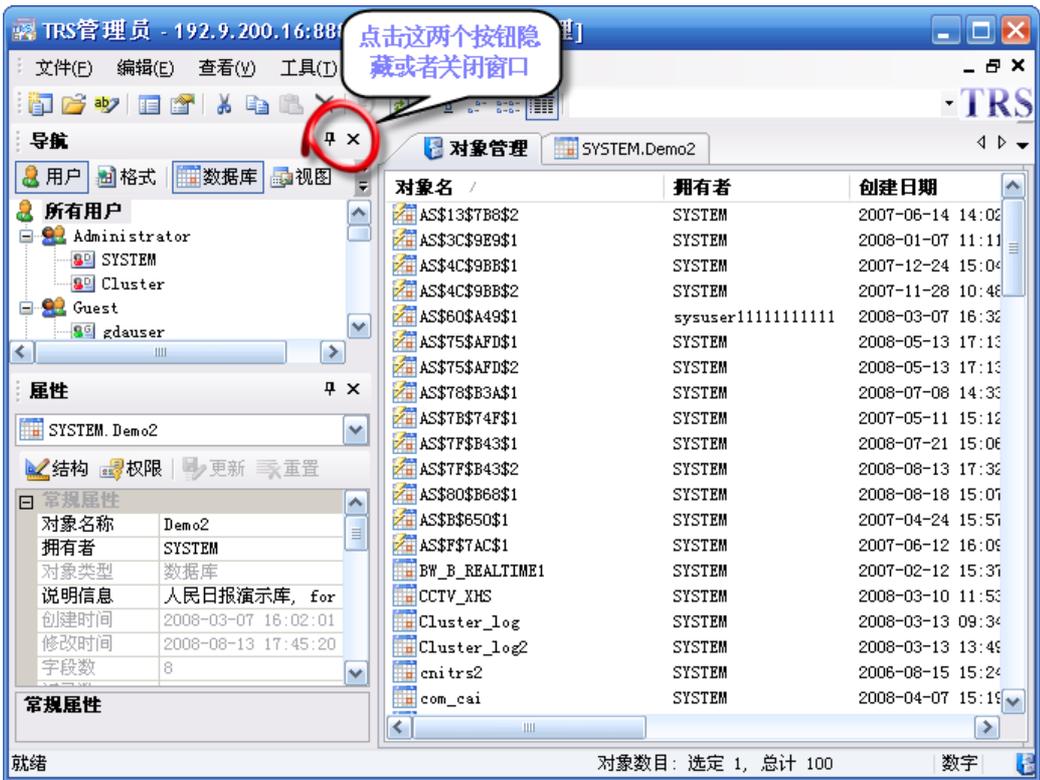


图 1—2 新的导航窗口和属性窗口示意图

这两个窗口采用了新的界面形式，点击隐藏按钮后窗口可隐藏在界面最左边，如下图：

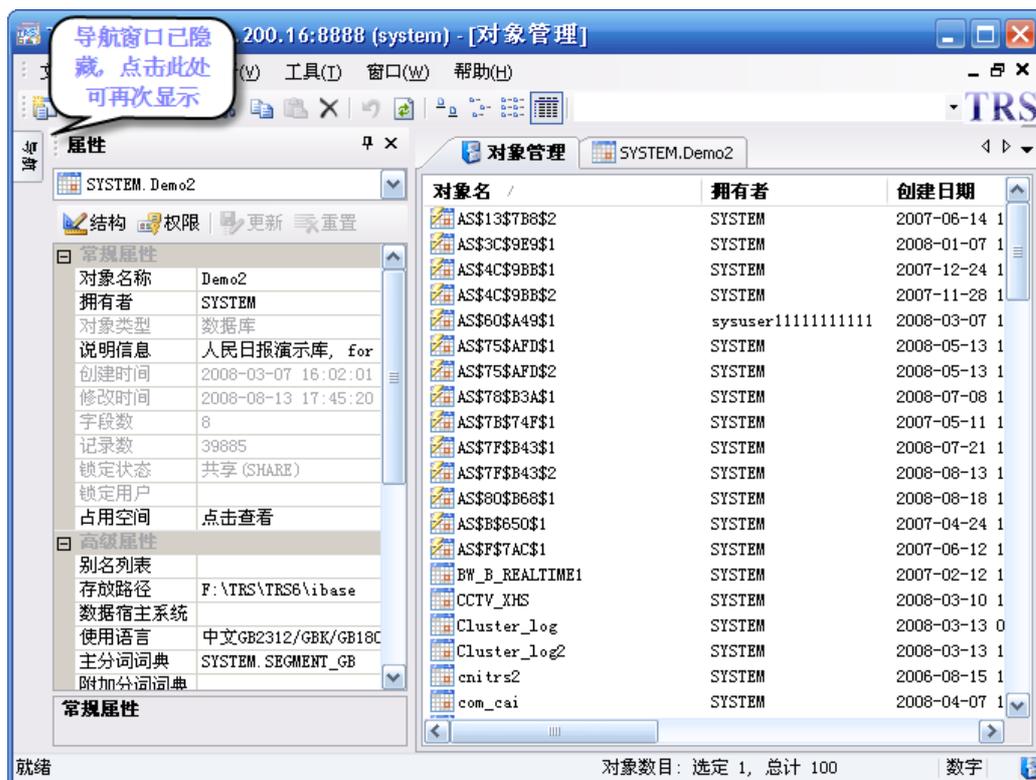


图 1—3 导航窗口隐藏状态

窗口既可以停靠在界面的四边，也可以悬浮于窗口的任何位置。如下图所示，属性窗口停靠在界面的下边，导航窗口悬浮于界面的右侧：

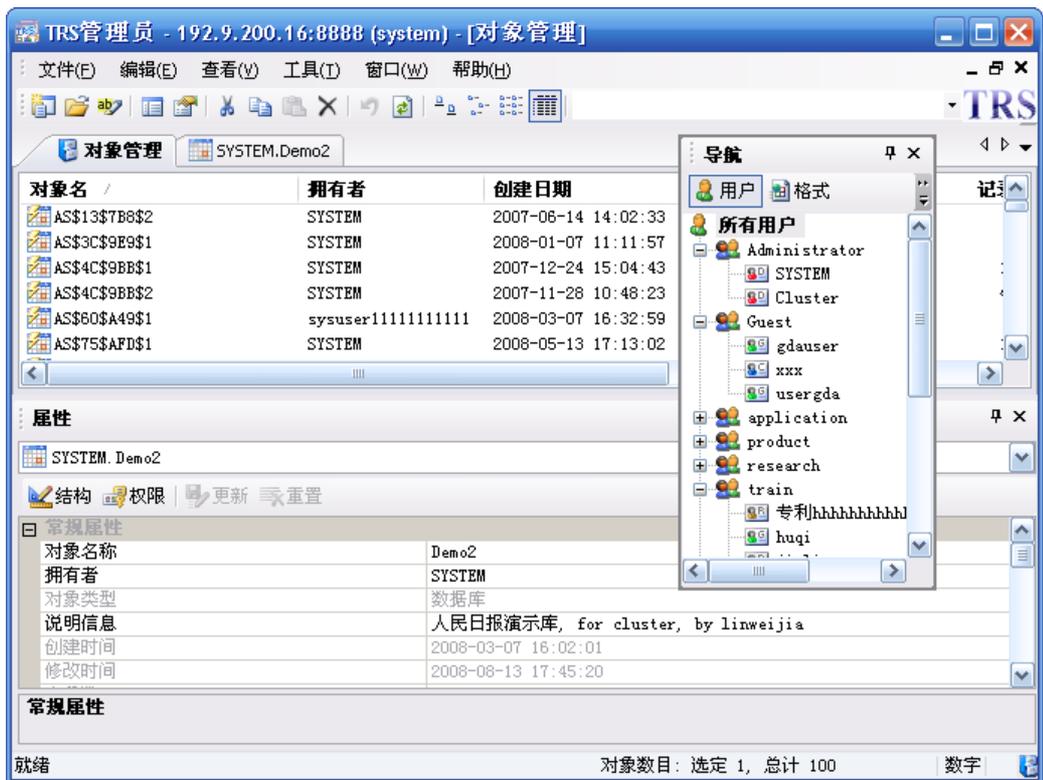


图 1—4 属性窗口停靠导航窗口悬浮

此外，当两个窗口关闭后，可以点击工具栏上  按钮和  按钮，再次打开这两个窗口，且所在位置与窗口关闭前的位置一致。

1.2.2 新的系统选项对话框，采用属性表方式显示选项，每一项均有概要提示

系统选择对话框也采用了新的属性框。单击对话框左部图片框中的图标可分别查看各页面的具体选项的设置。点击每一个属性项时，下方的注释框显示其相应的注释。

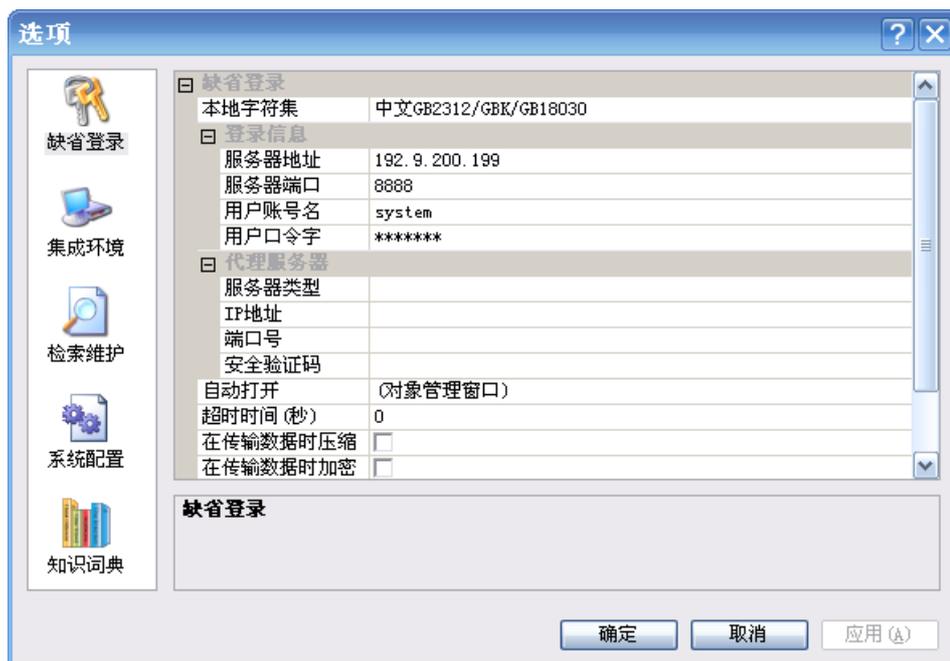


图 1—5 选项对话框

1.2.3 树形管理用户组、用户和数据格式，可筛选所属对象

本版本中，以树型结构的形式显示用户组、用户及它们之间的所属关系。用户组的各个子字节对应的用户属于此用户组，若父节点为根节点的用户表示不属于任何用户组。在此列表树中还可以通过拖拽将用户移动到别的用户组。



图 1—6 导航窗口中的用户列表树

此外，数据格式的类别也以树形结构的形式显示。点击各个类别，对象管理窗口即显示当前类别包含的所有数据格式列表。



图 1—7 导航窗口中的数据格式列表树

1.2.4 属性表式管理各种对象属性，可直接修改更新，结构和权限

独立出来

属性窗口采用新的属性框进行属性的显示。各属性可在其中直接修改，被改变的属性以粗体显示。点击各个属性项时，相应的注释将显示于下方注释框中。同时结构和权限被独立出来，作为弹出对话框显示，如下图：



图 1—8 使用属性框的属性窗口

属性框有下面几种控制项：

- ☆ 编辑框，直接在编辑框里输入内容

对象名称	Demo2
拥有者	SYSTEM

图 1—9 属性框中的编辑框

- ☆ 数字微调框，点击右边的微调按钮直接对数字进行更改，此外，需要输入数字的属性项有输入限制，非数字禁止输入。



图 1—10 属性框中的数字微调框

☆ 复选框，当点击出现对勾时，表示本属性项值为 TRUE，否则为 FALSE。



图 1—11 属性框中的复选框

☆ 下拉组合框，点击时，会显示下拉箭头按钮，点此箭头将会展开下拉框。



图 1—12 属性框中下拉组合框

☆ 弹出组合框，点击时，出现...按钮，点此按钮弹出相应的对话框。

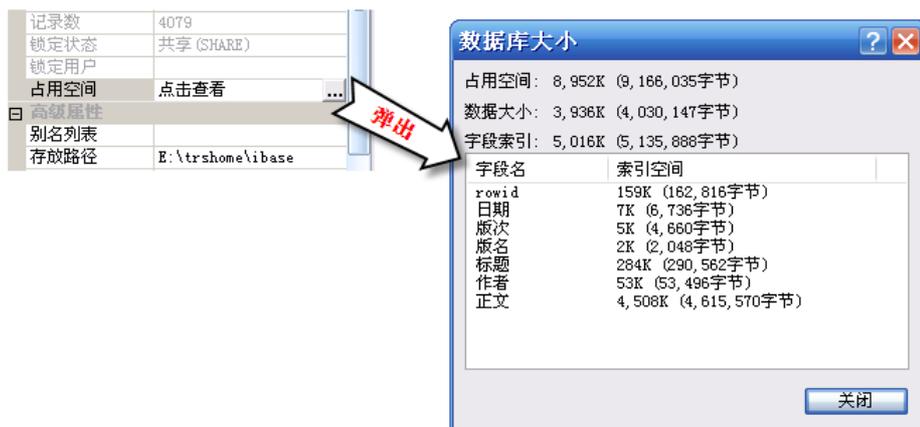


图 1—13 属性框中的弹出组合框

1.2.5 采用新的文本方式导入/导出对象，方便修改和将来扩展

本版本的 Admin 将采用新的“.tof”格式文件存储导出的对象信息和结构，此种格式的文件可以文本的方式打开，直接进行编辑和修改，并且可以将用户组与用户的信息保存到同一个文件中。

1.2.6 支持生成多级导航树，概览增加字段排序和记录标志，细览实现图文混合显示和命中点定位

记录维护窗口对应的导航树新增了检索条件具有蕴含关系的多级导航功能，即在继承根节点的检索条件下根据二级导航节点再次进行检索。以此类推，还可进行更多级的检索。但需要注意的是，此多级导航功能只有在使用“自动生成”功能的情况下才能生成，其他情况下生成的多级导航中，根节点和子节点代表的检索条件没有蕴含关系。

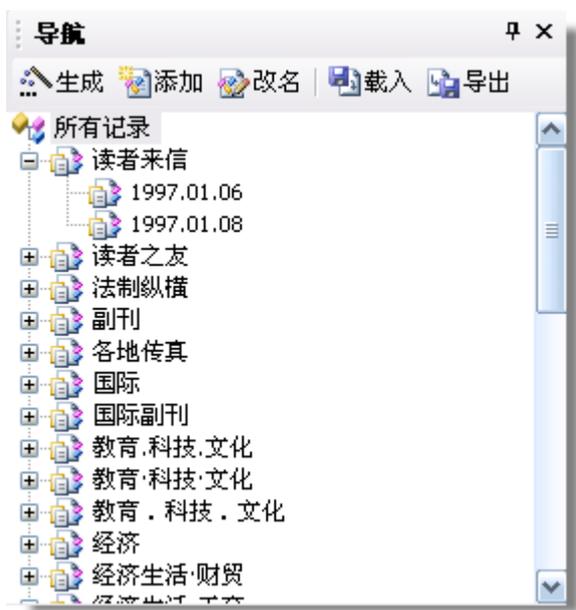


图 1—14 多级导航树

概览窗口中可按字段排序，点击字段名可按才字段升序或降序排列，字段名旁边的三角形尖角向上表示升序，尖角向下表示降序。同时，经过加密、修改或者删除等操作的记录前面会有相应的标志以示区别。

记录号	时间	版次	版名	>> 栏目 ▲	标题
<input type="checkbox"/>	1	1997.01.09	4	要闻	“三下乡”活动情暖万民心 西安创建：欢乐的村庄、希望的田野
<input type="checkbox"/>	2	1997.01.09	4	要闻	“三下乡”活动情暖万民心 北京组织：艺术家、书法家、科技和
<input type="checkbox"/>	3	1997.01.08	1	要闻	今日谈 『结穷亲』又有新发展
<input type="checkbox"/>	4	1997.01.11	1	要闻	今日谈 “卖掉奥迪发工资”联想
<input type="checkbox"/>	5	1997.01.04	1	要闻	今日谈征文 记民情日记好
<input type="checkbox"/>	6	1997.01.03	4	要闻	人民论坛 新年话信心
<input type="checkbox"/>	7	1997.01.06	4	要闻	人民论坛 论谋势
<input type="checkbox"/>	8	1997.01.08	4	要闻	人民论坛 精神文明就在身边
<input type="checkbox"/>	9	1997.01.10	4	要闻	人民论坛 戒『浮』
<input type="checkbox"/>	10	1997.01.06	1	要闻	社论 贵在行动 重在落实
<input type="checkbox"/>	11	1997.01.01	1	要闻	社论 把握大局 再接再厉 同心同德 开拓前
<input type="checkbox"/>	12	1997.01.10	1	要闻	现场短新闻 贫在深山有远亲
<input type="checkbox"/>	13	1997.01.07	4	要闻	现场短新闻征 特殊礼品赠教师
<input type="checkbox"/>	14	1997.01.02	1	要闻	现场短新闻征文 特别的新年晚会（现场短新闻征文）

图 1—15 概览窗口

下方细览窗口采用浏览器的方式实现了图文混合显示，同时支持命中点定位，利用上部的工具栏可定位到上一或者下一命中点处。

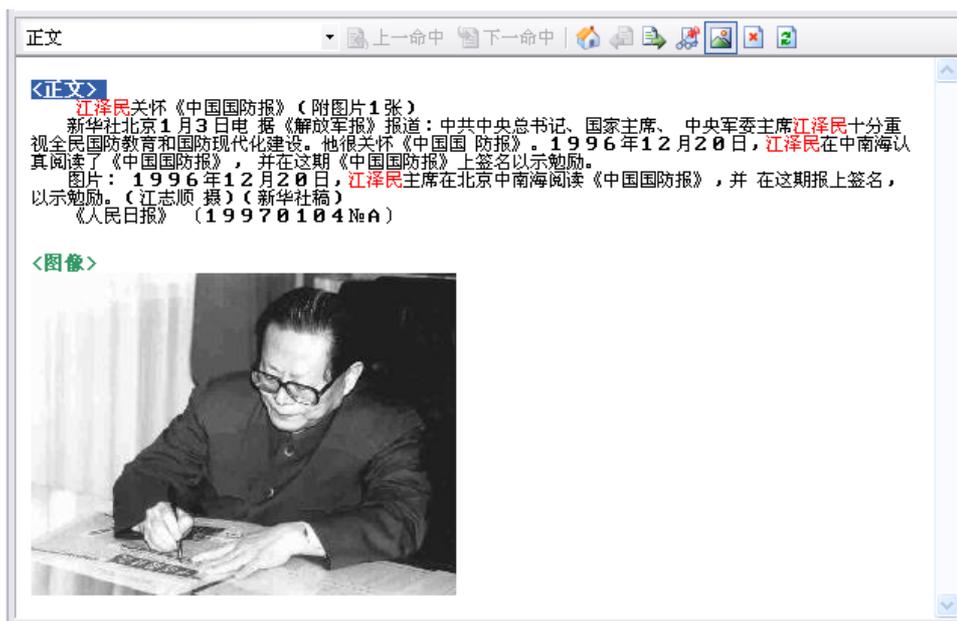


图 1—16 细览窗口

1.2.7 可自定义表单编辑记录界面，下拉列表提供日期/字符型字段可能值选择

本版本中的表单编辑界面可以自己对其编辑框的大小和位置进行调整，调整后点击  保存，即可保存所做修改，下次再打开此对象的表单窗口时，将自动恢复上一次对各编辑框的设置。下图即为控件设计状态，用户可自己对编辑框的位置和大小进行调整：

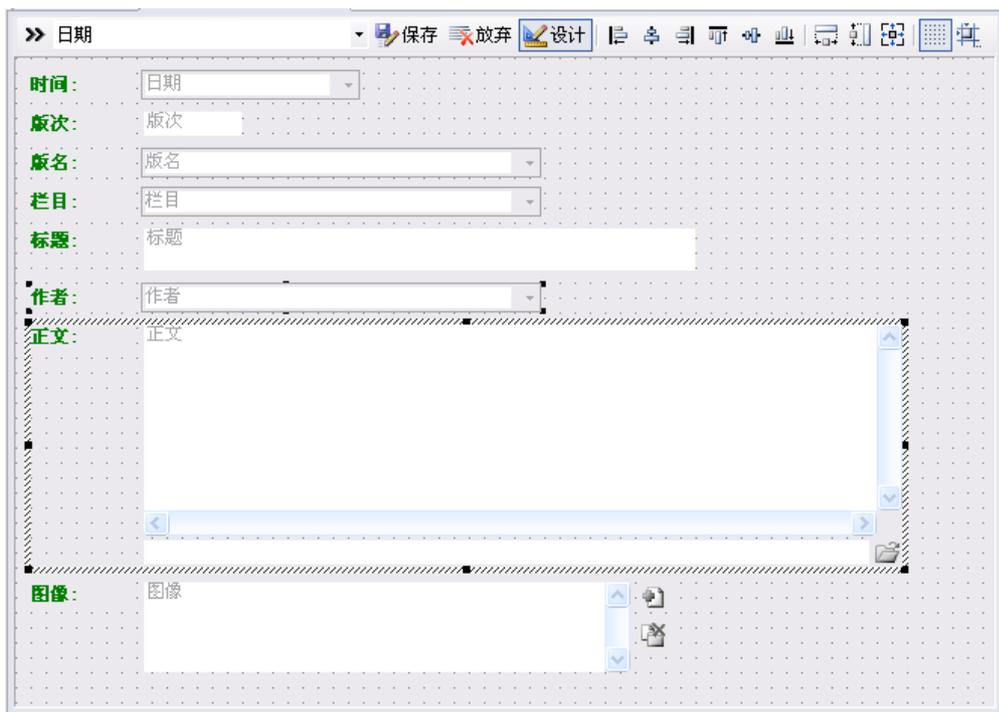


图 1—17 控件设计状态下的表单编辑窗口

同时，日期、字符型字段可以进行可能值的选择。



图 1—18 日期型字段的选择

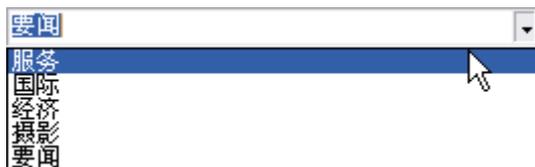


图 1—19 字符型字段可能取值的选择

1.2.8 改进了装库控制文件生成和结果报告查看对话框

装库控制文件生成对话框也采用了新的属性框进行各项属性的显示。

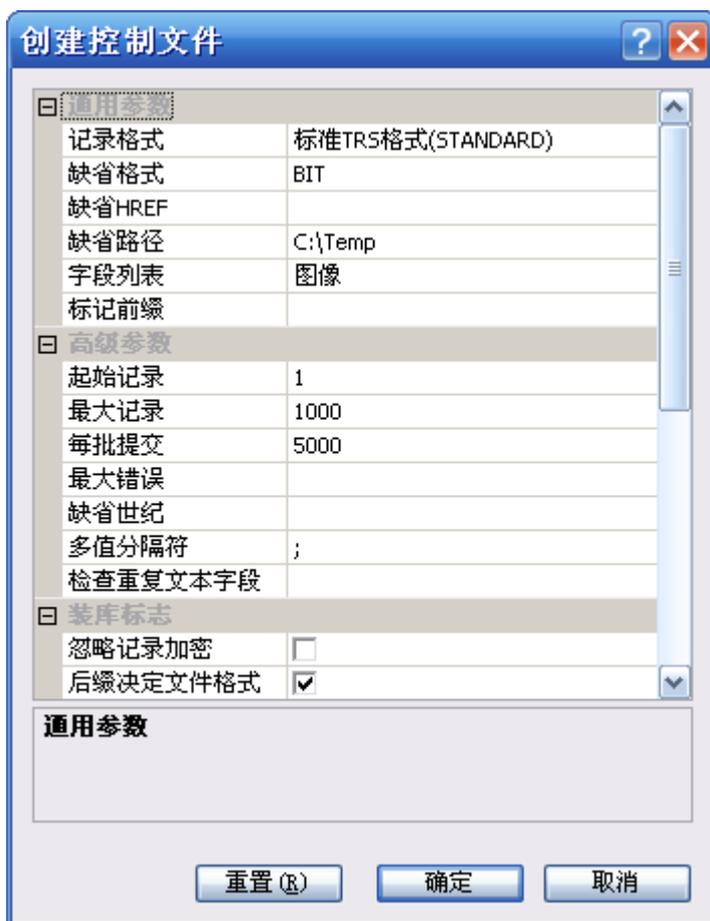


图 1—20 装库控制文件生成对话框

结果报告对话框添加了“生成报告”和“打开报告查看”的选项，在结果报告很大时，可以设置为只生成报告而不打开报告查看，节约了时间。

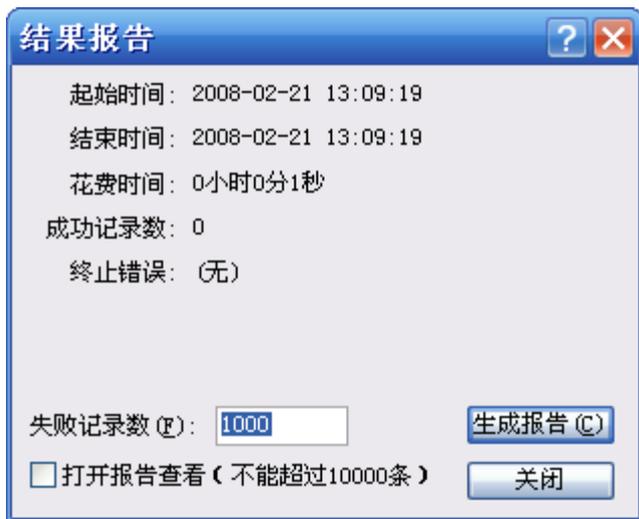


图 1—21 结果报告对话框

1.3 启动和退出 TRS 管理员

TRS 管理员运行于中文 Windows XP/2003 或以上版本平台，并且最好安装 Microsoft Internet Explorer 6.0 以上版本，否则你将无法浏览检索结果中的源文件以及联机帮助。如要登录到 TRS 服务器，必须保证该服务器的指定端口上安装了 TRS 全文数据库版本，并且该 TRS 服务进程处于启动状态。

双击安装时在桌面上生成的“TRS 管理员工具”图标，或在开始菜单中选择“TRS 全文数据库”程序组中的“TRS 管理员工具”菜单项，或以资源管理器、命令行或其它方式运行 TRSAdmin.exe 程序，即可启动 TRS 管理员。在欢迎信息框显示完毕以后，TRS 管理员将弹出如下“启动”对话框：



图 1—22 启动对话框

对话框中各控制的说明如下：

控制名称和取值	功能简介	
登录信息	服务器地址	服务器所在的主机，既可以是 TRS 服务器在网络中的 IP 地址（如“127.0.0.1”），也可以是服务器所在主机的名称（如“ALPHA”）。如果输入的是服务器主机的名称则要求该主机名是当前所使用的网络驱动系统的数据库文件 hosts 的一个入口项（如“127.0.0.1 alpha”）。服务器地址可以从 TRS 的系统管理员处获得。
	服务器端口	服务器端口既可以是 TRS 服务器的 INTERNET 端口（如“8888”），也可以是 TRS 服务器的名称（如“trsserver”）。如果输入的是服务器的名称，则要求该名称是当前所使用的网络驱动系统的数据库文件 services 的一个 TCP 协议入口项（如“trsserver 8888/tcp”）。服务器端口可以从 TRS 的系统管理员处获得。
	用户帐号名	用户的帐号名，必须是指定服务器中已经存在的用户名。用户的帐号由 SYSTEM 用户创建。
	用户口令字	用户的密码，区分大小写。为防止泄密，输入字符将用“*”号显示。用户的密码可由用户自己修改。
	保存密码	是否保存所输入的密码标志，密码会加密保存。
自动打开	对象管理窗口	登录成功后自动打开对象管理窗口，为缺省设置。
	数据库/视图	登录成功后自动打开指定的数据库/视图，如总是对某个数据库/视图进行操作可选择该选项。右边的组合框中列出的是以前打开过的最后十个数据库/视图名。
显示启动对话框	启动时是否显示“启动”对话框标志。如清除该选项，则下次启动时将不再显示本对话框。	
最大化窗口	是否最大化子窗口标志。如设置该选项，则所有子窗口都将占满整个主	

	窗口的空间。
--	--------

在登录组内依次填入服务器地址、服务器端口、用户帐号名和用户口令字，然后按确定，TRS 管理员会尝试与指定服务器建立连接。如连接成功，会打开用户在自动打开组内指定的窗口。输入的登录信息以及其它设置将会成为下一次运行时的缺省配置。

选择了“文件退出”命令（或直接用鼠标单击关闭按钮、按下 Alt+F4）后，如当前所打开的对象或文件中有尚未保存的，TRS 管理员将提示保存修改结果。保存完毕后，TRS 管理员将关闭所有打开的窗口，并切断与当前登录的 TRS 服务器的连接，最后关闭主窗口。

1.4 登录和退出登录

如想登录到其它服务器，或以其它用户的身份重新登录，可选择“工具|登录”命令，TRS 管理员将弹出如下“登录”对话框：



图 1—23 登录对话框

对话框中各控制的说明如下：

控制名称	功能简介
服务器地址	服务器所在的主机，既可以是TRS服务器在网络中的IP地址（如“127.0.0.1”），也可以是服务器所在主机的名称（如“ALPHA”）。如果输入的是服务器主机名称则要求该主机名是当前所使用的网络驱动系统的数据库文件hosts的一个入口项（如“127.0.0.1 alpha”）。服务器地址可以从TRS的系统管理员处获得。
服务器端口	服务器端口既可以是TRS服务器的INTERNET端口（如“8888”），也可以是TRS服务器的名称（如“trserver”）。如果输入的是服务器的名称，则要求该名称是当前所使用的网络驱动系统的数据库文件services的一个TCP协议入口项（如“trserver 8888/tcp”）。服务器端口可以从TRS的系统管理员处获得。
用户帐号名	用户的帐号名，必须是指定服务器中已经存在的用户名。用户的帐号由SYSTEM用户创建。
用户口令字	用户的密码，区分大小写。为防止泄密，输入字符将用‘*’号显示。用户的密码可由用户自己修改。

分别填写各项登录信息，然后按[确定]。TRS 管理员将首先关闭所有打开的浏览窗口，并从目前登录中退出，然后再尝试登录到指定服务器。如登录成功，主窗口的标题栏上会显示当前所登录的服务器的地址和端口号，以及登录的用户名。如对象管理窗口处于显示状态，则自动刷新对象列表。

当不再需要在当前登录的服务器上执行 TRS 操作时，选择“工具|退出登录”命令，系统将弹出一个提问对话框请求确认“是否决定退出登录？”。选择[是(Y)]，将切断与当前 TRS 服务器的连接，并关闭所有打开的浏览窗口，如对象管理处于显示状态则清除对象列表；选择[否(N)]，将继续保持当前连接。退出登录后，如要再执行 TRS 操作，必须重新登录到服务器。

1.5 集成环境概述

TRS 管理员是一个多窗口应用程序，主窗口主要执行一些全局操作以及管理各个子窗口。子窗口有三种不同类型：对象管理窗口、记录维护窗口和文件编辑窗口，并可同时打开多个记录维护窗口和文本编辑窗口，此外还有两个辅助窗口：导航窗口和属性窗口，主要用于显示上述窗口对应的不同对象列表和各自不同的属性。对应于上述三种窗口，导航窗口和属性窗口也存在一定的差别。

注意：各对话框在第一次打开时将采用各选项的缺省值，其后在没有重新启动 Admin 以前，已经打开过的对话框都将继承上一次设置的值。

TRS 系统中的对象包括：

- ◇ 用户组
- ◇ 用户
- ◇ 数据库和数据库字段
- ◇ 视图及视图字段
- ◇ 数据格式
- ◇ 词典

所有这些对象都具有一致的命名规则：

- ◇ 对象名可以是中文，英文或中英文的组合。
- ◇ 大小写无关。
- ◇ 只能包括A-Z，a-z，0-9，_，或汉字。
- ◇ 不能是TRS保留字(关键词)。
- ◇ 用户组 and 用户不能重名。
- ◇ 属于同一个TRS用户的数据库、视图和词典对象不能重名。
- ◇ 同一个数据库/视图的字段不能重名。

注意：建议不要使用数字开头，尽管这是允许的。

1.5.1 对象管理窗口

TRS 管理员是一个面向对象的管理工具，对象管理窗口里列出TRS的数据库、视图、词典和数据格式对象，对这四种对象进行的创建、修改、删除、重命名等操作均在此窗口完成，并可打开数据库/视图/词典进行浏览和维护操作。

左侧上部对应的导航窗口显示用户列表树或者数据格式列表树，左侧下部对应的属性窗口显示当前所选对象的具体属性。点击导航窗口工具栏中不同的对象，或者列表树的不同节点时，对象管理窗口将显示相应对象的列表，同时左下部属性窗口显示其属性。对象管理窗口如下图：

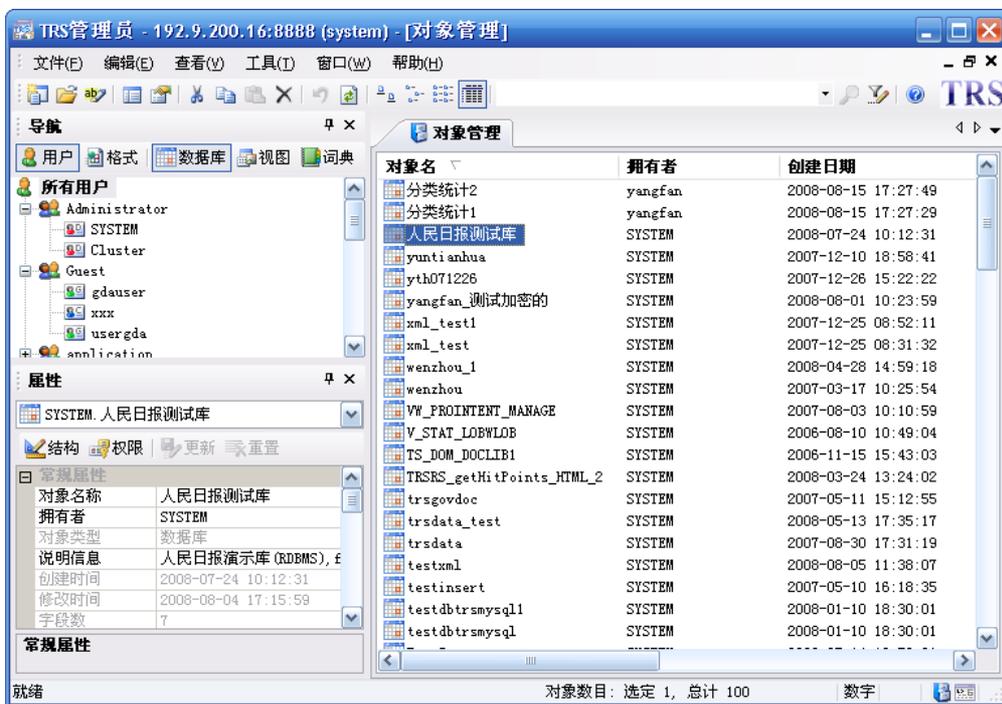


图 1—24 对象管理窗口

点击左侧导航窗口工具栏中不同的对象，对象管理窗口将显示相应对象的列表，同时左下部属性窗口显示其属性。对象列表可以下表中所示的四种方式显示(缺省情况下为详细信息方式)，选择“查看”菜单里相应的命令可方便地改变显示格式，也可使用相应的工具按钮改变对象显示方式。

	大图标:	用大图标方式显示对象，每个图标表示一个对象，并只列出对象的名字；
	小图标:	用小图标方式显示对象，每个图标表示一个对象，并只列出对象的名字；

	列表:	显示对象的列表，每个图标表示一个对象，并只列出对象的名字；
	详细信息:	显示对象的详细信息，对每个对象均显示其对象名、拥有者、创建日期、说明等信息。

选择“查看|排列图标”可按一定的顺序排列图标。排列方式有：按名称、按拥有者、按日期、按类型四种，并可选择当窗口大小改变时是否重新排列图标。如当前是以详细信息方式显示，单击栏目标题可设定按该栏目排列图标，栏目标题旁边的三角形尖角向上表示升序，尖角向下表示降序。选择“查看|对齐图标”可按行列对齐图标（只对以大图标或小图标方式显示对象列表时有效）。如想在窗口大小改变时自动对齐图标，请选中“查看|排列图标|自动排列”选项。

选择“查看|刷新”或按下 F5 键可刷新当前的对象列表。当对象列表可能被其它应用程序或其它 TRS 管理员实例修改时，最好使用本命命令刷新对象列表，以得到最新的对象信息。同时，相应的导航窗口和属性窗口也将被刷新。

对象管理窗口对应的工具栏上包含了一些常用的菜单命令，状态栏的中间部分显示当前的对象总数以及选定的对象个数。另外，对象管理窗口提供了一些快捷键和快捷菜单用于快速执行菜单命令。双击对象管理窗口的对象名可打开该对象，根据对象的类型分别进入记录维护窗口、词典维护对话框或对象属性对话框。

（一） 导航窗口

对象管理窗口对应的导航窗口可显示两种不同的对象列表树，即用户列表树和数据格式列表树。

1. 用户列表树是一个用树型结构显示各用户组及用户之间关系的列表。用户组的各个子字节对应的用户属于此用户组，若父节点为根节点的用户表示不属于任何用户组。点击各个节点，对象管理窗口显示当前用户或用户组拥有的相应对象列表，同时下方属性窗口显示其属性。用户列表树还支持节点的拖动，但只支持用户移动到别的用户组。操作方法是：按住所要移动的用户节点，并把它拖动到目标节点下，然后放开鼠标左键，如果移动合法的话，此用户将被移动到新的用户组中，同时用户的所属用户组属性也将改变。



图 1—25 导航窗口的用户列表树

各节点和前面图标的对应关系如下：

	用户组
	系统数据库管理员 DBA
	用户组数据库管理员 GDA
	系统资源用户 RESOURCE
	系统登录用户 CONNECT

用户列表树各个节点的维护通过快捷菜单来完成，选择相应节点并单击鼠标右键即可调出快捷菜单。如当前选择节点为根节点，快捷菜单如下：

- 新建用户组 创建新的用户组对象
- 添加导航 新建具有筛选功能的导航节点
- 粘贴 把剪贴板内容复制到当前服务器
- 查看源文件 查看当前用户列表树对应的导航文件

如当前选择节点不是根节点，快捷菜单如下：

- 新建用户 创建新的用户对象，若当前所选节点为用户组，则新建属于此用户组的用户。若当前所选节点为用户，则新建与此用户同组的用户
- 添加导航 新建具有筛选功能的导航节点
- 剪切 删除所选对象并把它放到剪贴板上
- 复制 复制所选对象到剪贴板上

粘贴	插入剪贴板内容
删除	插入剪贴板内容
重命名	重命名当前选中的对象
属性	显示选定中对象的属性，若属性窗口关闭，重新打开属性窗口

若在导航窗口的空白处点击右键，快捷菜单如下：

新建对象	创建新的用户组或者用户对象
添加导航	新建具有筛选功能的导航节点
刷新	刷新当前窗口

2. 显示数据格式列表树。这是一个用树型结构显示数据格式类别的列表。各个子节点表示当前服务器拥有的数据格式类别。点击各个节点，对象管理窗口显示当前类别包含的所有数据格式列表。同时下方属性窗口显示其属性。



图 1—26 导航窗口的数据格式列表树

 用户	此按钮按下时，显示用户列表树，同时右边“数据库”按钮自动选中，对象管理窗口显示所有的数据库列表
 格式	此按钮按下时，显示数据格式列表树，同时对对象管理窗口显示所有的数据格式列表
 数据库	此按钮按下时，“用户”按钮即自动选中，对象管理窗口显示所有数据库列表；当点击用户列表树的某一节点时，对象管理窗口显示当前所选用户组或用户拥有的所有数据库列表
 视图	此按钮按下时，“用户”按钮即自动选中，对象管理窗口显示所有视图列表；当点击用户列表树的某一节点时，对象管理窗口显示当前所选用户组或用户拥有的所有视图列表

	此按钮按下时，“用户”按钮即自动选中，对象管理窗口显示所有词典列表；当点击用户列表树的某一节点时，对象管理窗口显示当前所选用户组或用户拥有的所有词典列表
---	--

数据格式列表树各个节点的维护通过快捷菜单来完成，快捷菜单如下：

新建格式	创建新的数据格式对象，若当前所选节点为数据格式类别，则新建属于此类别的数据格式
属性	显示选定中对象的属性，若属性窗口关闭，重新打开属性窗口

(二) 属性窗口

对象管理窗口对应的属性窗口可以查看和修改当前选中对象的属性、结构和权限等信息。当点击对象管理窗口的对象或者导航窗口的节点时，属性窗口会显示所选对象的属性。如要对数据库、视图、词典的属性进行修改，则当前用户必须是 **DBA** 用户，或为选中对象的拥有者，或对选中对象有 **ALTER** 权限。只有 **SYSTEM** 用户才能修改用户组/用户的属性。

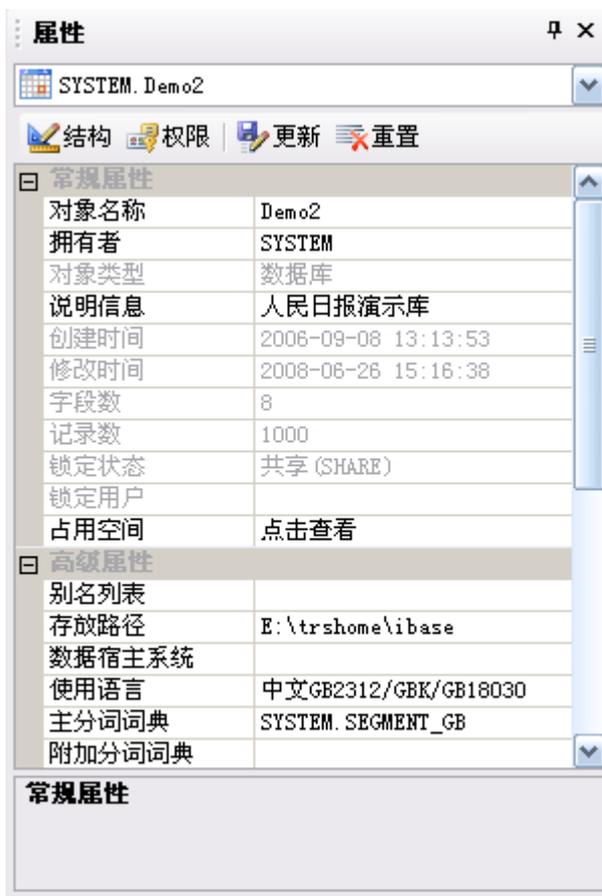


图 1—27 对象管理窗口对应的属性窗口

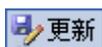
窗口最上部是一个对象列表框，其下拉框中列有最近曾经显示过属性的对象列表，选择相应的对象，下方属性框中将显示其属性。工具栏中的每个按钮的具体功能如下：



打开对象结构对话框



打开对象权限对话框



保存所修改的对象属性



重置当前对象的属性，即恢复到上次已存的属性

属性窗口除能显示各种对象的属性（见后面几章的具体描述），还能显示具有筛选功能的导航节点的属性。如下图：



图 1—28 导航节点的属性

属性框中各控制的说明如下：

控制名称	功能简介
筛选书签	定义当前筛选对象的节点名。长度在 1-31 个字符之间，不能包含半角冒号。
筛选条件	显示对象的筛选条件，可使用通配符：用'?'匹配单个字符，用'*'匹配任何字
可检索	如选中表示只显示可以检索的对象。
可更新	如选中表示只显示可以更新的对象。
可索引	如选中表示只显示可以建立索引的对象。
可修改	如选中表示只显示可以修改结构的对象。

1.5.2 记录维护窗口

记录维护窗口用于数据库/视图的检索、浏览结果和维护记录。右边部分可有四种方式：只有概览、只有细览、上下分隔、表单编辑，一般缺省为上下分隔方式（上方为概览窗口，右下方为细览窗口，边界可以上下调整）。左侧上部的导航窗口显示当前对象对应的导航树，左侧下部的属性窗口显示当前对象的检索属性和上方导航树的属性。下图为上下分隔时的记录维护窗口：

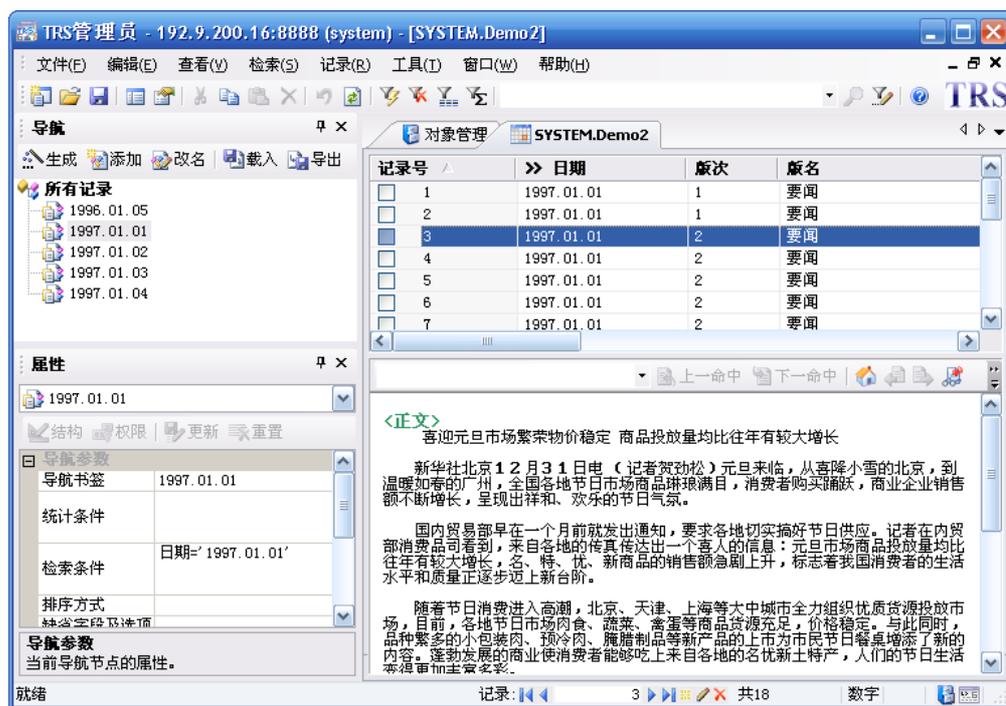


图 1—29 记录维护窗口

当右侧显示为记录维护窗口时，导航窗口显示导航树。导航树为一树型导航结构，每一个节点记录了检索条件、排序方式、统计表达式以及检索选项等信息，选中一个节点将根据这些信息重新进行检索，并按节点设置的显示方式显示检索结果。导航树的维护可直接在该窗口内进行，也可以在外部编辑一个导航文件，并把它载入到当前窗口里。同样，当前窗口的导航树也可以导出到一个外部的文件中。

概览窗口以表格方式显示用户指定的概览字段值。字段名显示在表格最上面的标题栏，点击字段名可进行排序，三角形尖角向上表示升序，尖角向下表示降序。第一列显示记录号和相关度（为简明起见，这里只显示近似的百分数，而不是精确的相关性数值）。表格最左边的复选框为选择记录之用，可选定部分记录进行删除、更新、保存等操作。当当前记录改变的时候，细览窗口的内容会相应地作出改变。显示的字段列表可以通过选择“检索|显示字段”命令或设置导航节点的显示方式来改变，每一栏的宽度可通过拖动边界或选择“查看|表格列”来设定，缺省宽度由数据库字段的显示宽度属性决定。

细览窗口以浏览器的方式显示用户指定的细览字段值，可直接显示记录所带的 jpg 或者 gif 类型图片，其他类型的附件在默认状态下以连接形式显示，单击带下划线的文字以打开或保存该文件或连接。但处于直接浏览记录附件状态（即  按钮按下）时，点击记录时将直接显示该记录所带的附件，或者弹出对话框询问用户是否打开或保存所带的附件，选择[打开]，将在细览窗口中直接打开此附件，选择[保存]，则可将此附件保存到指定

的目录中。显示的字段列表可以通过选择“检索|显示字段”命令或设置导航节点的显示方式来改变。

表单窗口以表单方式显示用户指定的字段值，每一个字段对应一个编辑框，可直接在编辑框里修改字段值。对于 DOCUMENT/BIT 型字段可点击右边按钮  或者  添加文件，点击按钮  删除文件。显示的字段列表可以通过选择“检索|显示字段”命令或设置导航节点的显示方式来改变。本版本的表单窗口还新增了编辑框位置和大小设计。点击  设计按钮，所有编辑框变灰，表单窗口处于控件设计状态。点击某一个编辑框，可对其大小进行改变或者随意拖动其位置。修改完以后点击  保存按钮即可保存这些设置，下次再打开此对象的表单窗口时，将自动恢复上一次对各编辑框的设置。

记录维护窗口对应的属性窗口显示当前导航节点属性和当前进行检索对象的检索信息。

记录维护窗口对应的工具栏上包含了一些常用的菜单命令，状态栏的中间部分为记录定位，显示当前记录号，并可快速定位到某条记录。另外，记录维护窗口提供了一些快捷键用于快速执行菜单命令，而窗口的每个组成部分各有相应的快捷菜单。

1.5.3 文件编辑窗口

文件编辑窗口用于创建和编辑文本文件。它的用法与 Windows 的记事本类似，功能有所增强，并能编辑大于 64K 的文件。TRS 管理员还提供了针对不同类型文件的向导，用于快速地生成导航文件、装库数据文件、批装库文件和装库控制文件的模板。左侧的属性窗口同时显示当前文件的属性。以下是打开了一个“.txt”文件的编辑窗口：

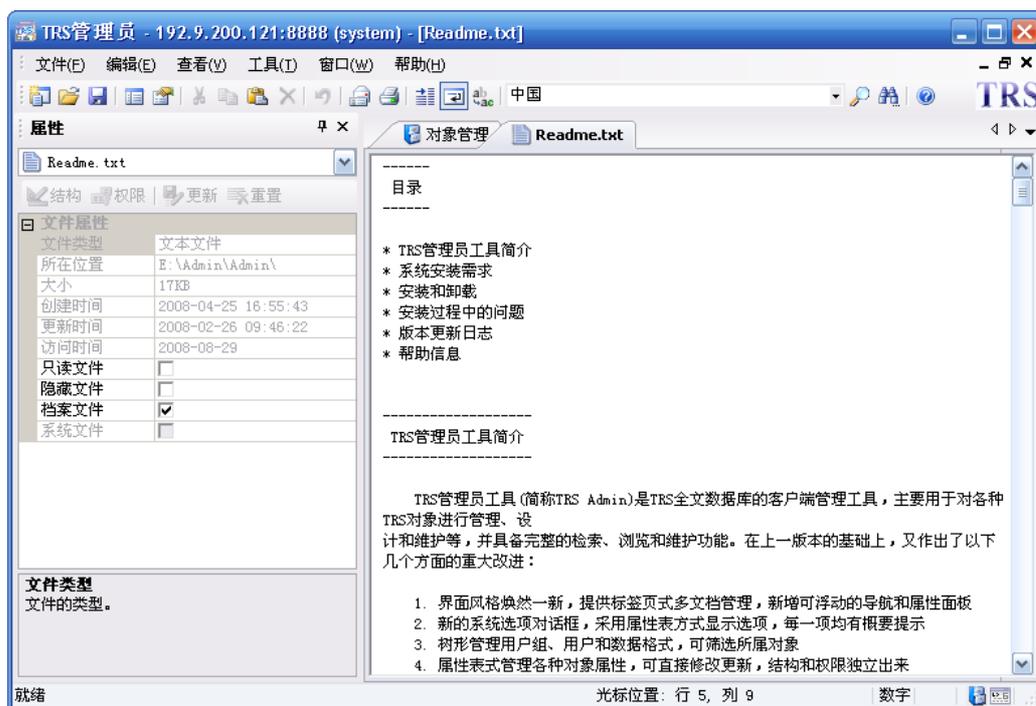


图 1—30 文件编辑窗口

文件编辑窗口对应的工具栏上包含了一些常用的菜单命令，状态栏的中间部分显示当前插入点所在的行和列。另外，文件编辑窗口提供了一些快捷键和快捷菜单用于快速执行菜单命令。在文本编辑窗口里单击鼠标右键即可调出以下快捷菜单：

- | | |
|------|-----------------|
| 撤销 | 撤销上一次操作 |
| 剪切 | 删除所选内容并把它放到剪贴板上 |
| 复制 | 复制所选内容到剪贴板上 |
| 粘贴 | 插入剪贴板内容 |
| 删除 | 删除所选内容 |
| 全部选定 | 选定所有内容 |

下面列举一些文件编辑窗口独有菜单命令：

制表位 修改编辑控制的制表符字符数，在 1-256 之间，缺省值为 8。所设制表位字符数将对以后打开的所有文件编辑窗口有效。

查找 在当前显示文本中查找指定的字符串。选择“查看|查找”命令后，TRS 管理员弹出如下对话框：



图 1—31 查找对话框

对话框中各控制的说明如下：

控制名称	功能简介
查找目标	要查找的文字。
全字匹配	选中表示只查找整个单词，而不查找包含此单词的长字符串。
区分大小写	选中表示只查找与输入文字大小写一致的字符串。

输入要查找的文字，设置好查找选项，按[查找下一个]，TRS 管理员将从文件编辑窗口的当前位置开始查找键入文字下一次出现的位置。由于查找对话框是一个无模式对话框，所以可以执行多次查找而不关闭对话框，并可切换到文件编辑窗口进行文本编辑。如要关闭对话框，请按[取消]。

替换 在当前显示文本中查找指定的字符串并用另一指定字符串替换。选择“查看|替换”命令后，TRS 管理员弹出如下对话框：



图 1—32 替换对话框

对话框中各控制的说明如下：

控制名称	功能简介
查找目标	要查找的文字。
替换成为	用于替代查找文字的新字符串。
全字匹配	选中表示只查找整个单词，而不查找包含此单词的长字符串。
区分大小	选中表示只查找与输入文字大小写一致的字符串。

输入要查找的文字或用于替换查找文字的字符串，设置好查找选项，按[查找下一个]，TRS 管理员将从文件编辑窗口的当前位置开始查找键入文字下一次出现的位置；按[替换]，将搜索查找文字下一次出现的位置，并把它替换为新的字符串；按[全部替换]，将把文件编辑窗口中所有的查找文字全部替换成新的字符串。由于替换对话框是一个无模式对话框，所以可以执行多次查找/替换而不关闭对话框，并可切换到文件编辑窗口进行文本编辑。如要关闭对话框，请按[取消]。

重复 重复上一次查找/替换操作。

编码 改变已打开文件的编码。有 GB18030、BIG5、UTF8、Unicode 和 Unicode(big endian)几种编码选择。

属性窗口

当右侧显示的是文件编辑窗口时，属性窗口显示的是当前文件的属性。

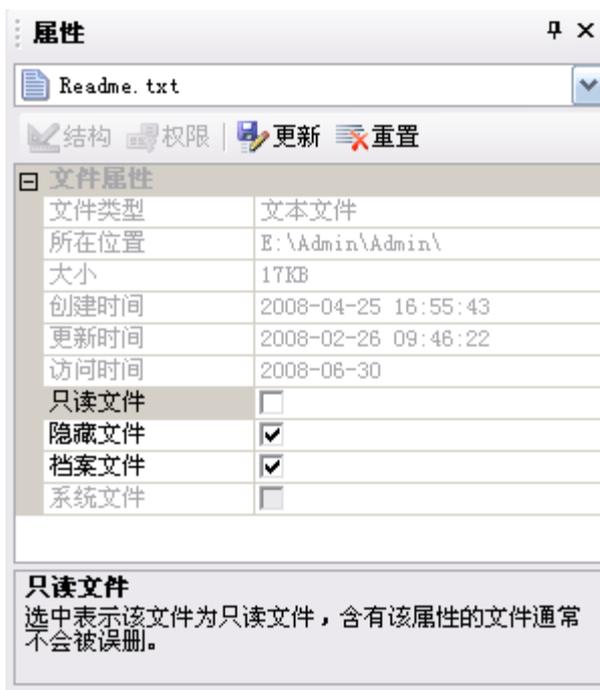


图 1—33 文档编辑窗口对应的属性窗口

属性框里的控制显示当前文件的属性，具体如下：

控制名称	功能简介
文件类型	文件的类型，可为：导航文件、数据文件、控制文件、文本文件之一。
所在位置	文件存放的目录。

控制名称	功能简介
文件大小	文件占用的空间大小（以K字节计算）。
创建时间	文件的创建日期和时间。
修改时间	文件的最后一次修改日期和时间。
访问时间	文件的最后一次访问日期。
只读文件	选中表示该文件为只读文件，含有该属性的文件通常不会被误删。
隐藏文件	选中表示该文件为隐藏文件，含有该属性的文件在默认情况下不显示在文件列表中。
档案文件	选中表示该文件为档案文件，有些系统利用此属性来控制是否备份该文件。
系统文件	选中表示该文件为系统文件，系统文件是系统运行时有用到的，默认情况下不显示在文件列表中，请不要删除含有此标志的文件。

如对文件的属性进行了修改，点击工具栏按钮 ，即可保存相关修改；点击工

具栏按钮 ，重置当前对象的属性，即恢复到上次已存的属性。

第2章 使用数据库对象

2.1 TRS 数据库对象概述

数据库是 TRS 系统中的主要数据对象，它物理地存储了用户加载到系统中的所有数据资料。每个 TRS 系统可管理多达数十亿个数据库。

数据库的规模主要指以下两个方面：能够容纳的最大记录数和数据库文件的最大容量。一个数据库的最大记录数取决于操作系统的位数和 TRS 所采用的文件系统的位数：在 32 位文件系统中，最多可容纳 4 亿多条记录，而在 64 位文件系统中，如果是 32 位操作系统，则最多可容纳 20 多亿条记录，如果是 64 位操作系统，则最多可容纳 40 多亿条记录。制约数据库文件最大容量的唯一因素是文件系统的位数，如果 TRS 采用了 32 位文件系统，则单个文件的大小不能超过 2G 字节（在 Windows 平台上不能超过 4G 字节），但如果采用 64 位文件系统，则单个文件的容量实际上没有限制。**操作系统与文件系统的位数没有必然的联系，许多 32 位操作系统都支持 64 位文件系统**，如 WIN32、Linux、Solaris 等，所以在 32 位操作系统上，TRS 也尽可能地采用 64 位文件系统。

每个记录的最大长度为 256 兆字节，每个字段值的最大长度为 16 兆字节（但以独立，即 ALONE 方式存储的字段值的长度不受此限制。如果需要并允许压缩，则计算压缩后的字段值长度）。但在实际中，如果一个记录（或字段值）太长，会引起操作系统忙于虚拟内存的交换，使机器性能急剧下降，所以系统人为地限制一个记录的最大长度为物理内存兆字节数的 64 的倍数（最小为 1）兆字节，如物理内存为 128 兆字节，则允许的最大记录长度为 2 兆字节。这里的“记录长度”是指集中存储在数据库主文件中的数据块的长度（包含各字段的存储控制信息，以及以集中方式存储的字段值。对于字段值，如果需要并允许压缩，则计算压缩后的字段值长度），或者是各个全文型和二进制型字段中的文本数据的总长度（如果需要并允许压缩，则计算压缩后的数据长度）；“物理内存”是指 TRS 系统配置中的值，而不是实际机器所拥有的硬件内存，调节 TRS 系统配置就可以改变当前系统中一个记录所允许的最大长度，但最大不能超过 256 兆字节。字段值的最大长度不能调节。

数据库对象具有下列属性：

- 名称

名称是 TRS 数据库的对象名，即数据库名，是访问、维护和管理该数据库的入口。数据库名不能与属于同一用户的其它数据库、视图以及各种词典同名。

- **别名列表**

别名是该数据库名的别称，一般是为了方便使用而给出的数据库名的缩写，或其英文名称。数据库别名与数据库名具有同等的地位。一个数据库可以定义多个别名。

- **拥有者**

拥有者表明是哪个用户拥有该数据库。拥有者与名称（别名）唯一确定一个数据库对象。拥有者对该数据库能够进行任何操作，而不受其它数据安全机制的制约。

- **缺省检索字段**

缺省检索字段是指在对该数据库进行查询时，检索表达式没有明确指定检索目标字段的子表达式所默认的目标字段。

- **分词词典**

分词词典是该数据库按词索引时所引用的词汇表。如果不为数据库指定分词词典，则该数据库只按字索引。

- **附加分词词典**

附加分词词典是该数据库按词索引时所引用的扩充词汇表。

- **停用词典**

停用词典是该数据库按词索引时所引用的不允许索引的词汇表。

- **附加停用词典**

附加停用词典是该数据库按词索引时所引用的不允许索引的扩充词汇表。

- **稀疏词典**

稀疏词典是该数据库按词索引时所引用的允许索引的词汇表。如果不为数据库指定稀疏词典，则该数据库索引所有不属于停用词典和附加停用词典中的词汇。

- **公共访问权限**

公共访问权限是指 TRS 系统中所有用户对该数据库所拥有的权限。对数据库可设置的公共访问权限共有五种：

- ◇ 记录检索权(SELECT)
- ◇ 数据更新权(UPDATE)
- ◇ 数据索引权(INDEX)
- ◇ 数据库结构修改权(ALTER)
- ◇ 数据库删除权(DROP)

- **词根索引开关**

词根索引开关决定数据库的短语型和全文型字段数据中的字母单词是否按词根进行索引。如果设置该开关，则索引不区分大小写。

- **大小写敏感开关**

大小写敏感开关决定数据库的字符型、短语型和全文型字段数据中的字母串是否区分大小写进行索引。

- **数字索引开关**

数字索引开关决定数据库的短语型和全文型字段数据中的数字是否进行索引。

- **二元索引开关**

二元组索引开关决定数据库的短语型和全文型字段是否建立二元组索引。

- **禁止压缩开关**

禁止压缩开关决定数据库是否建立压缩索引。

- **使用语言**

使用语言是指存储在数据库中的非英文数据所使用的字符集。TRS 系统支持的字符集有：

- ◇ 简体中文(GB2312、GBK、GB18030)
- ◇ 繁体中文(BIG5)
- ◇ 纯英文(ENGLISH)
- ◇ 统一码(UTF8)

- **数据宿主系统**

有一类应用：TRS 系统只管理和维护索引，而数据存储在其他系统中。数据宿主系统就是用来描述存储数据的系统的信息。描述信息对 TRS 没有实际的意义，如果数据库没有数据宿主系统，则 TRS 系统负责管理数据库的数据。

- **存放路径**

数据存放路径指明数据库数据文件的存储目录。

- **说明信息**

说明信息是对 TRS 数据库对象进行说明的信息。

在定义数据库时，对其每一个属性都必须给出其正确取值。在数据加载之前，对数据库的属性可以进行任意修改。

在加载了数据后，下列属性不再允许修改：

- 数据字符集
- 数据宿主系统
- 数据存放路径

当创建了索引后，下列属性不再允许修改：

- 分词词典
- 附加分词词典
- 停用词典
- 附加停用词典
- 稀疏词典
- 缺省检索字段
- 词根索引开关
- 大小写敏感开关
- 数字索引开关

- 二元组索引开关
- 禁止压缩开关

TRS 还支持一种自动分裂数据库，其属性、结构和权限由所属的自动模式视图决定，用户不能修改，但对记录和索引的维护与普通数据库一样。

2.2 TRS 数据库字段对象概述

数据库字段是 TRS 系统对不同数据类型进行存储管理的主要依据，也是对数据库进行查询的唯一入口，所有的查询操作都是在一定的字段上进行的。每个数据库可定义 1 到 1023 个字段。

数据库字段对象具有下列属性：

- **字段名**
字段名是 TRS 数据库字段的对象名，是访问、维护和管理该数据库字段的入口。字段名不能与同一数据库中的其它字段同名。
- **别名列表**
别名是该字段名的别称，一般是为了方便使用而给出的字段名的缩写，或其英文名称。字段别名与字段名具有同等的地位。一个字段可以定义多个别名。
- **字段类型**
字段类型决定了该字段所能存储的数据类型。TRS 能存储和索引六种类型的数据：
 - ◇ 日期型(DATE)
 - ◇ 数值型(NUMBER)
 - ◇ 字符串型(CHAR)
 - ◇ 短语型(PHRASE)
 - ◇ 全文型(DOCUMENT)
 - ◇ 二进制型(BIT)
- **缺省取值**
缺省取值是指在数据加载时，对于未给出字段值的 DATE、NUMBER 或 CHAR 型字段所对应的缺省取值，对其它类型字段没有意义。
- **格式限定**
该属性对于不同的字段类型有不同的意义。
 - ◇ 对于 CHAR 型字段，字段值格式是指数据加载时，给出的字段值必须遵循的格式。
 - ◇ 对于 DATE 和 NUMBER 型字段，字段值格式是指输出字段值时使用的格式，并且由外部完成字段值的格式化。

◇ 对于其他类型的字段，字段值格式没有意义。

● 取值范围

取值范围是指在加载数据时，DATE、NUMBER 或 CHAR 型字段的合法取值的集合，对其它类型字段没有意义。取值范围的定义格式包括：

◇ 枚举定义：格式为V1,V2,V3,Vn，即枚举该字段上所有可能取值；

◇ 区间定义：格式为V1,...,V2，表示该字段上所有可能取值必须在最小值V1和最大值V2之间；

◇ 混合定义：格式为V1,V2,...,V3,Vn，是以上两种方式的混合。

以上各格式中的逗号和点号都必须是半角符号。CHAR 型字段的取值范围所给出的每个值串首尾不允许空格。DATE 型字段的取值范围所给出的每个日期中只支持绝对（完全）年份，而不支持缺省世纪数的短格式年份，即总是认为缺省世纪数为 0。只有 DATE 型字段的取值范围支持模糊格式，即允许日期格式中缺省一些成份，但对于区间定义定义的范围，其中的 V1 和 V2 “前模糊”的成份必须是相同的，即在“年月日时分秒”的顺序下，V1 和 V2 的第一个确定的日期成份的位置必须是相同的。对于 DATE 型字段还必须注意：其日期的书写格式中不要使用逗号 ‘,’ 和点号 ‘.’ 分隔各成分，以免与“取值范围”本身的表示格式中的逗号和点号发生冲突。如果第一个字符为 ‘@’，则表示其中存储的是取值范围所在文件的文件名。缺省值为空，表示字段的值没有范围限制。

● 字段显示名

字段显示名是用户界面上该字段的对应名字。

● 显示区宽度

显示区宽度是字段值在用户界面上的显示区域的缺省宽度单位数。

● 显示区高度

显示区高度是字段值在用户界面上的显示区域的缺省高度单位数。

● 公共检索开关

公共检索开关决定该字段是否允许 TRS 系统中的所有用户进行查询。

● 禁止索引开关

禁止索引开关决定该字段是否建立索引。BIT 字段始终不建索引。

● 允许多值开关

允许多值开关决定在一个记录中该字段是否允许存在多个值。PHRASE 和 DOCUMENT 字段始终不允许多个字段值。

● 值唯一性开关

值唯一性开关决定在数据库的所有记录中，该字段上（DATE、NUMBER 或 CHAR 型字段）是否允许出现重复的字段值。

● 禁止空值开关

禁止空值开关决定在一个记录中该字段（DATE、NUMBER 或 CHAR 型字段）是

索引语种就是用来指定数据库中字母文字的语种，目前支持 12 种语言：英语、俄语、法语、西班牙语、葡萄牙语、意大利语、德语、荷兰语、瑞典语、挪威语、丹麦语、芬兰语。显然当一个字段中有多种字母文字时，就无法进行正确的词根处理。如果不指明词根索引语种，则系统将根据字母的种类默认一个语种：拉丁字母默认为英语；西里尔字母默认为俄语。

● 混合模式

混合模式只对具有“按词”索引属性的数据库的 PHRASE 和 DOCUMENT 字段有效。字词混合索引可以细分成三部分：只属于“按词”的索引；只属于“按字”的索引；属于“按词”和“按字”的公共索引。混合索引模式是指明其中字词公共索引部分的存储方式，共有四种：

- ◇ 全冗余模式 将公共部分分别在“按词”和“按字”索引中各存一份。优点是检索时直接取各自的索引即可；缺点是索引空间膨胀大。
- ◇ 半冗余模式 将公共部分只存储在“按词”的索引中。优点是“按词”检索时可直接取“按词”的索引即可，且索引空间膨小；缺点是“按字”检索时要取“按字”的索引和“按词”的索引进行合并，多一次“或”运算。
- ◇ 零冗余模式 将公共部分单独存储。优点是索引空间膨小；缺点是“按字”检索时要取“按字”的索引和“按词”的索引进行合并，多一次“或”运算，逻辑上，“按词”检索也多一次“或”运算。
- ◇ 禁止词模式 这种模式实际不是“混合索引”，而是为了实现在其它未指明混合索引模式的字段“按词”索引的情况下，该字段“按字”索引。

一般地，如果一个字片是“按词”和“按字”的公共部分，则它不可能存在只属于“按词”的索引。因此零冗余模式与半冗余模式的实际效果可能是一样的，没有明显的差别。如果一个 PHRASE 或 DOCUMENT 字段未指明任何混合索引模式，则该字段是“按词”，还是“按字”索引，取决于数据库的属性。

● 允许单字检索

允许单字检索开关只有在混合索引模式有效或者数据库按字索引的前提下才有效。如果是全字索引，则“按字”检索能够进行单个字的查询。

● 索引路径

索引路径指明该字段上的索引文件的存储目录。

● 说明信息

说明信息是对 TRS 数据库字段对象进行说明的信息。

在定义数据库字段时，对其每一个属性都必须给出其正确取值。在数据加载之前，对数据库字的属性可以进行任意修改，但只要加载了数据，则有许多属性就不再允许修改。

在加载了数据后，下列属性不再允许修改：

- 数据类型
- 多值允许开关

- 值唯一性开关
- 空值禁止开关
- 最大长度或精度

当创建了索引后，下列属性不再允许修改：

- 索引存放路径
- 索引属性
- 混合索引模式
- 全字索引开关
- 词根索引语种。

2.3 创建 TRS 数据库对象

2.3.1 TRS 数据库结构设计

数据库结构的设计是把非结构化数据建立数据模型的过程，这就要求用户对数据库结构的设计要考虑到如下多方面因素。

- 用户数据特点
- 数据量
- 查询对象
- 各字段属性
- 建库策略等

2.3.2 TRS 数据库建库策略

TRS 系统有两种类型的数据库索引：**B⁺**树索引和倒排(INVERT)索引。这两种类型的索引相辅相成，互为补充。

B⁺树索引的突出特点是很好地解决了键值的插入、溢出、删除和空间回收等问题，从而使**B⁺**树索引可以适应成批插入、文件易变的情况；**B⁺**树索引在操作中可“动态地”进行维护，无须周期性地重新组织文件；还可以采用压缩索引项的办法，提高扇出率，降低树的高度，减少读取次数，加快查找速度。因此**B⁺**树索引是组织大型索引文件的一种最主要的方法。

B⁺树索引实现了索引键值的唯一化，即一个键值在**B⁺**树索引中只有一个唯一的入口，而倒排索引则用来存储索引键值在一个记录和不同记录中出现的情况。倒排索引的建立大大提高了重复键值的查询速度和空间使用率，同时也简化了**B⁺**树索引的维护操作。

倒排索引由一系列倒排索引项组成，每个倒排索引项又可包含多个倒排索引属性，但不同的字段类型，所支持的倒排索引属性也不尽相同：

- DATE、NUMBER和CHAR字段不支持任何倒排索引属性。
- PHRASE字段上所支持的倒排索引属性是固定的，由系统自动定义。
- DOCUMENT字段上的倒排索引属性则可在创建数据库时据实际需要由用户自定义。

对于英文数据库，TRS 系统采用按词索引的固定模式。对于中文数据库，TRS 系统则有四种类型的索引策略：按中文词检索的索引策略、按中文字检索的索引策略、按用户自定义的关键词索引策略，以及二元组索引策略。在实际应用中，采用何种索引策略主要取决于数据的特点、检索的需求和系统的资源等。

不同的索引策略所需的索引词典也不尽相同。下面是索引策略与词典的关系：

策略	稀疏词典	分词词典	附加分词词典	停用词典	附加停用词典
按词	置空	必需	可有可无	可有可无	可有可无
按字	置空	置空	可有可无	可有可无	可有可无
按关键词	必需	忽略	忽略	忽略	忽略
二元组	置空	可有可无	可有可无	可有可无	可有可无

按词索引：TRS 系统内嵌的汉语自动分词系统 TRS 的领先技术，使得 TRS 在国内外同类全文检索系统中独具特色。对大多数中文全文检索应用来说，按词索引是首选的索引策略。使用按词索引策略的方法是：为数据库指定相应的分词词典（全字词典除外）、附加分词词典，或者一部稀疏词典。

按词索引策略的优越性在于：

- 检索的准确率即查准率高。如果采用按词索引策略，则在检索“华人”时，仅包含“中华人民共和国”的记录不会被检索出来。
- 查询的速度快。测试表明，在大型数据库上，按词索引的查询速度要比按字索引的查询速度快5-10倍。
- 空间膨胀率低。由于一个词往往包含了多个字，这样按词索引就大大减少了索引量，再加上停用词典的使用，索引文件的空间膨胀率比按字索引的小50-100%。
- 按词索引是实现智能化概念检索的基础。由于所有辅助知识词典（如主题词典、同义词典、反义词典等）都是基于词，而不是基于字的，所以只有按词索引，才能进行概念检索。

按词索引策略的不足在于：

- 词典维护后需要重新进行索引。
- 由于表意文字的词与词之间没有语法上的分隔符，自动分词难免有误，对检索也会有一定的负面影响。但在相同的语境中，“以错对错”将会自动消除这种影响。

按字索引：如果数据不是现代汉语或数据库规模很小，可采用按字索引策略。按字索引策略的主要优点是：

- 不需要自动分词系统，实现相对简单。
- 没有自动分词的错误，查全率较高。

然而，按字索引策略的不足也是显而易见的，按词索引策略的所有优点均是按字索引策略的致命缺点所以除非需要百分之百的查全率，而不惜牺牲查准率，一般情况下不使用按字索引策略。由于传统意义上的按字索引策略的查询速度太慢，这里的按字索引实际上是一种二元组索引。

使用“**不需要检索单字**”的按字索引策略的方法是：

方法1. 数据库不指定任何分词词典、附加分词词典和稀疏词典。

方法2. 在数据库使用按词索引策略的前提下，将需要按字索引的字段的混合索引模式指定为“禁止词模式”。

使用“**允许检索单字**”的按字索引策略的方法是：

方法1. 为数据库指定分词词典为全字词典。

方法2. 在数据库使用按词索引策略的前提下，将需要按字索引的字段的混合索引模式指定为“禁止词模式”，并开启该字段的全字索引开关。

按关键词索引：按关键词索引实际上是一种抽词索引技术，是在按词索引策略的基础上只抽取其中的某些词进行索引。所有需要索引的关键词构成一部稀疏词典。一般对专用领域的文献可采用这种索引策略。

按关键词索引策略兼有按词索引策略的优点，并且具有更低的空间膨胀率和更快的检索速度。但由于按关键词索引策略只对“关键词”进行了索引，这样就不能像按词索引策略那样可用任意的自由词进行全文检索，而只能对“关键词”进行检索。

二元组索引：二元组索引策略是指在按词索引时对高频词进行二元组合，或者在按字索引时对高频字串进行二元组合。有分词词典或附加分词词典的二元组索引称为按词二元组索引；反之，如果既没有分词词典，也没有附加分词词典，则称为按字二元组索引。由此可见二元组索引策略并不是一种独立的索引策略，它是按词索引策略和按字索引策略的扩展。二元组索引还依赖于二元组词典（高频词词典），不同的数据库语言对应不同字符集的词典，如果数据库语言与系统主字符集不一致，则必须通过一次与数据库语言一致的系统升级来自动创建所需要的词典。

二元组索引策略的最大优势是在大数据库中对含有高频词（字串）的短语的查询时，速度有显著提高。但这种索引策略的缺点也同样明显：一是空间膨胀率较高，尤其是对于

小数据库；二是装库速度要比单独的按词或按字索引策略慢得多。所以对于小规模数据库，在没有二元组索引的情况下，其查询速度也很快，完全没有必要使用这种索引策略。

特别地，按字二元组索引有两个级别：字之间的二元组索引（一级二元索引）和高频字串（即一级二元组）与字之间的二元组索引（二级二元索引，实际上是三元索引，但由于不是所有的字都进行三元索引，所以这里只把它看成是二元索引的一种优化扩展）。第一级按字二元组索引策略就是按字索引策略（因此，并不是所有的字都能单独查询，如果需要能够检索所有的单字，可以给数据库指定系统中的“所有汉字词典”作为分词词典），而这里所说的按字二元组索引就是指第二级按字二元组索引策略。二元组索引策略的选择不是通过词典的选取来决定的，而是通过数据库的二元组索引开关来决定的。

2.3.3 TRIS 数据库创建

创建新的数据库。必须是具有 Resource 权限的用户才可以创建数据库。进入创建数据库对话框有三种方法。

方法一：选择“文件|新建”命令，在“新建”对话框里双击数据库图标；

方法二：对象管理窗口显示数据库列表且焦点也在其上时，按工具栏上的按钮；

方法三：对象管理窗口显示数据库列表且焦点也在其上时，在对象管理框按鼠标右键，出现菜单后选择新建。

通过上述方法之一将会开始执行“创建数据库向导”，该向导共有如下三步：

第一步：输入数据库名、拥有者名、说明信息、公共属性等信息，并可选择一个已存在的数据库作为模板。创建数据库向导第一步的对话框如下：

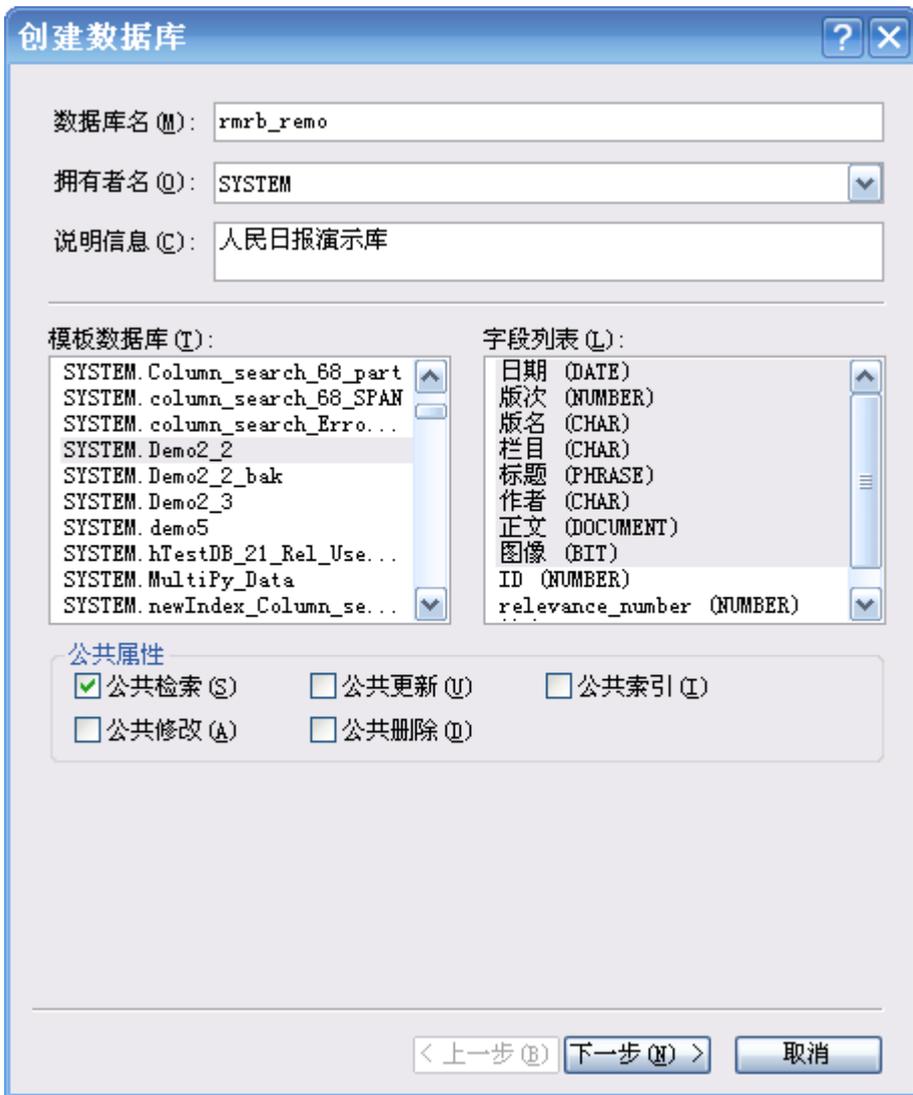


图 2—1 创建数据库第一步

对话框里各控制的说明如下：

控制名称	功能简介
数据库名	新建数据库的名字，不能包含中文、字母、数字或下划线以外的字符，大小写无关。名字长度在1-31个字符之间，不能是TRS保留字，也不能和拥有者的数据库、视图和词典重名。
拥有者名	新建数据库所属的拥有者名，缺省为当前登录用户的用户名。
说明信息	不超过63个字符长度的数据库说明信息。
模板数据库	现有的所有数据库的列表，可选择其中之一作为新数据库的模板。缺省时不选择模板。
字段列表	所选模板数据库的字段列表，可选择一个或多个字段作为新数据库字段的模板。缺省时选择全部字段。

控制名称	功能简介	
公共属性	数据库的公共访问标志集，包含下面所列的几个标志。	
公共属性	公共检索	选中表示所有用户均可对该数据库进行检索。
	公共更新	选中表示所有用户均可对该数据库进行记录维护。
	公共索引	选中表示所有用户均可对该数据库的索引进行维护。
	公共修改	选中表示所有用户均可修改该数据库的结构。
	公共删除	选中表示所有用户均可删除该数据库。

输入以上信息后，按[下一步]按钮进入向导的第二步。按[取消]则退出创建数据库向导。

第二步：设置数据库的别名列表、使用语言、存放路径、数据宿主、索引词典和索引标志等属性。创建数据库向导第二步的对话框如下：

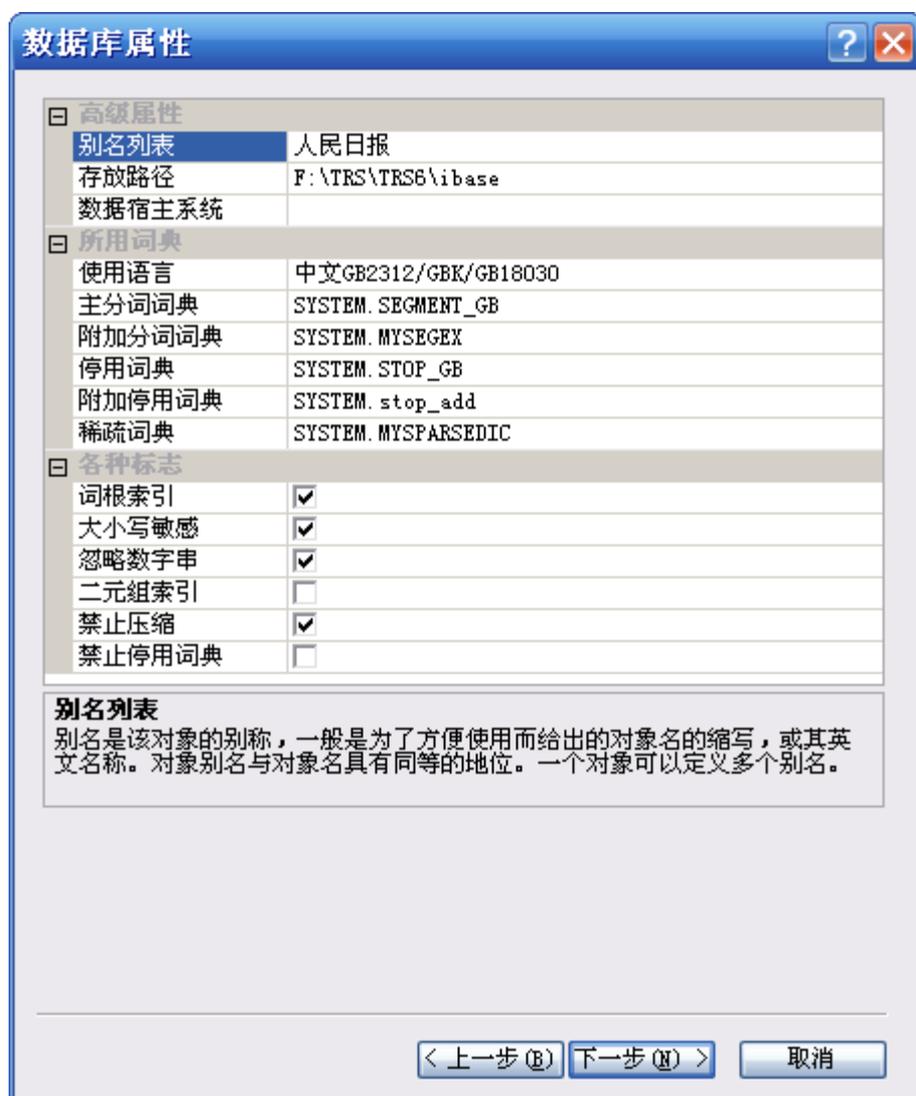


图 2—2 创建数据库第二步

对话框里各控制的说明如下：

控制名称	功能简介
别名列表	数据库的别名列表，多个别名之间以半角分号、逗号或空格分隔，单个别名长度不能超过31个字符，总长度（含分隔符）不能超过63个字符。缺省值为空。
存放路径	数据库存放的路径，必须是系统配置中的数据和索引路径之一，缺省为系统配置中的第一个数据路径。可以在下拉列表中选择。
数据宿主系统	数据库存放的位置。如果不为空，表明实际数据集中存储在一个数据宿主服务器中，并指向存储服务器的有关信息的字符串，如“ORACLE”。缺省值为空。
使用语言	数据库所用的语言，有如下几种取值（缺省为系统配置中的缺省语言环境）：中文GB2312/GBK/GB18030、中文BIG5、英语、统一码UTF8。
主分词词典	该数据库按词索引时所引用的词典。缺省为系统配置中相应语言的缺省主分词词典名。
附加分词词典	该数据库按词索引时所引用的扩充词典。缺省为数据库拥有者的缺省附加分词词典名。
停用词典	该数据库按词索引时所引用的不允许索引的词典。缺省为系统配置中相应语言的缺省停用词典名。
附加停用词典	数据库按词索引时所引用的不允许索引的扩充词典。缺省为数据库拥有者的缺省附加停用词典名。
稀疏词典	该数据库按词索引时所引用的允许索引的词典。缺省为数据库拥有者的缺省稀疏词典名。
词根索引	决定数据库的短语型和全文型字段数据中的字母单词是否按词根进行索引，缺省为系统配置中设定的只建立词根索引标志。
大小写敏感	决定数据库的字符型、短语型和全文型字段数据中的字母串是否区分大小写进行索引，缺省为系统配置中设定的大小写敏感标志。
忽略数字串	选中表示建立索引时忽略数字串，缺省为系统配置中设定的忽略数字串标志。
二元组索引	决定数据库的短语型和全文型字段是否建立二元组索引，缺省为清除。
禁止压缩	决定数据库是否对数据和索引进行压缩。对于CPU速度慢而IO速度较快的系统，不压缩数据与索引对数据库的入库与检索速度会略有提高，但空间膨胀率会显著增大，所以在实际使用时总是压缩索引，以获得较好的时空比，缺省为清除。
禁止停用词典	禁止停用词典开关的作用是允许建立停用词的索引（即创建数据的索引不启用停用词典），检索时停用词是否参与运算由应用进行选择。此功能为TRS全文数据库6.50.4000版本新增。

设置好数据库的别名列表、使用语言、存放路径、数据宿主、索引词典和索引标志等属性以后，按[下一步]按钮进入向导的第三步。按[上一步]可返回向导的第一步，按[取消]则退出创建数据库向导。

第三步：设计数据库的字段结构以及各字段的具体属性等。创建数据库向导第三步的对话框如下：

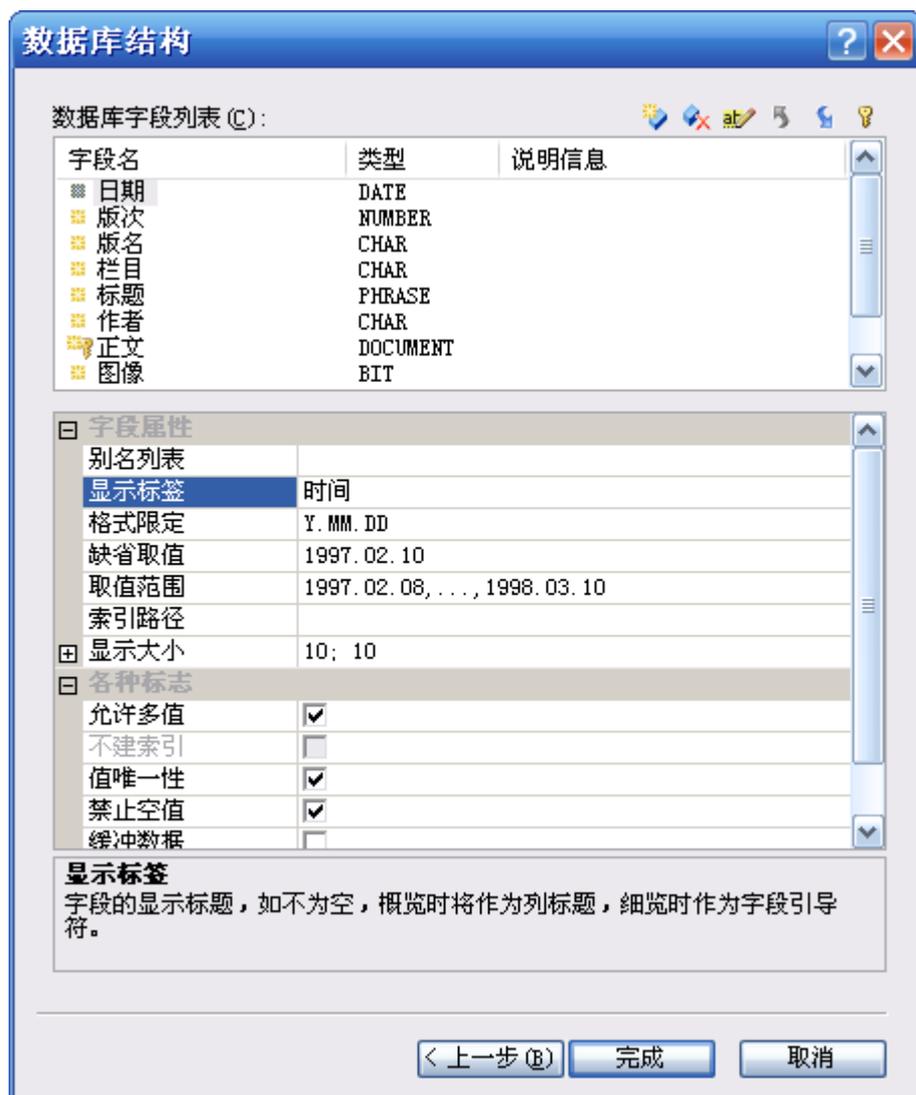


图 2—3 创建数据库第三步

数据库字段列表控制里列出新建数据库的所有字段，第一次进入时显示的是第一步里选择的模板字段。字段列表的每一行分别显示字段的名称、类型和说明，字段名前有的标志的字段为新增字段，有的标志的字段为数据库进行全文检索的缺省全文字段名，缺省为最后一个 DOCUMENT 或者 PHRASE 字段，DOCUMENT 字段优先。用户可在模板字段的基础上进行添加字段、修改字段属性、删除字段、为字段改名、调整字段顺序(仅限于新字段)等操作。数据库字段列表控制的右上角有六个按钮，它们的作用分别如下：

-  添加新字段，快捷键为 Insert
-  删除当前选中字段，快捷键为 Delete



为当前选中字段改名，快捷键为空格，直接在字段名上单击也能起到同样作用



当前选中字段位置上移，仅限于新字段，快捷键为 Shift+ ↑



当前选中字段位置下移，仅限于新字段，快捷键为 Shift+ ↓



把当前选中字段设为缺省字段，快捷键为 Shift+空格

数据库字段列表控制下面是字段属性组框，组框里的控制显示当前数据库字段列表控制里选中的字段的各项属性，具体如下：

控制名称	功能简介
别名列表	数据库字段的别名列表，多个别名之间以半角分号、逗号或空格分隔，单个别名长度不能超过31个字符，总长度（含分隔符）不能超过63个字符。缺省值为空。
显示标签	数据库字段的显示标题，如不为空，概览时将作为列标题，细览时作为字段引导符。
格式限定	数据库字段的输入输出格式。 对于CHAR型字段，其格式规定为： <ul style="list-style-type: none"> • A 表示在相应位置上必须出现一个英文字母； • 9 表示在相应位置上必须出现一个数字； • 空格 表示在相应位置上必须出现一个空格； • X 表示在相应位置上可出现任意字符。 对于DATE和NUMBER型字段，则为字段值的输出限定格式，并且格式由用户确定，如“Y.MM.DD HH:MM:SS”、“L.DP”等。 如果限定格式第一个字符为‘@’，则表示其中存储的是限定格式所在文件的文件名。按右边的...按钮可打开“格式限定”编辑对话框，在此对话框中可以直接进行编辑或者浏览格式限定文件。缺省值为空，表示字段的值没有格式限制。对其它类型字段没有意义。
缺省取值	DATE、NUMBER或CHAR型字段的缺省值，对其它类型字段没有意义。如果第一个字符为‘@’，则表示其中存储的是缺省值所在文件的文件名。按右边的...按钮可打开“缺省取值”编辑对话框，在此对话框中可以直接进行编辑或者浏览缺省取值文件。缺省值为空，表示没有缺省值。
取值范围	DATE、NUMBER或CHAR型字段的取值范围，对其它类型字段没有意义。取值范围的定义格式包括： <ul style="list-style-type: none"> • 枚举定义：格式为V1,V2,V3,Vn，即枚举该字段上所有可能取值； • 区间定义：格式为V1,...,V2，表示该字段上所有可能取值必须在最小值V1和最大值V2之间； • 混合定义：格式为V1,V2,...,V3,Vn，是以上两种方式的混合。 以上各格式中的逗号和点号都必须为半角符号。CHAR型字段的取值范围所给出的每个值串首尾不允许空格。DATE型字段的取值范围所给出的每个日期中只支持绝对（完全）年份，而不支持缺省世纪数的短格式年份，即总是认为缺省世纪数为0。只有DATE型字段的取值范围支持模糊格式，即允许日期格式中缺省一些成份，但对于区间定义定义的范围，其中的V1和V2“前模糊”的成份必须是相同的，即在“年月日时分秒”的顺序下，V1和V2的第一个确定的日期成份的位置必须是相同的。如果第一个字符为‘@’，则表示其中存储的是取值范围所在文件的文件名，按右边的...按钮可打开“取值范围”编辑对话框，在此对话框中可以直接进行编辑或者浏览取值范围文件。缺省值为空，表示字段的值没有范围限制。
索引路径	数据库字段索引的存放路径，必须是系统配置中的数据和索引路径之一。缺省

控制名称	功能简介
	值为空，表示字段索引路径和数据库存放路径一致。
索引属性	<p>按右边的...可打开“索引属性”编辑对话框，在此对话框中可以直接进行编辑或者浏览索引属性文件。</p> <ul style="list-style-type: none"> ◆ 对DOCUMENT字段来说，即倒排索引属性名。多个属性名之间用‘;’分隔，如“SEG;SEN;POS”等。如果第一个字符为‘@’，则表示其中存储的是索引属性所在文件的文件名。对DOCUMENT字段缺省值为“SEG;SEN;POS”，对其它类型的字段缺省值为空。 ◆ 对于CHAR字段来说，索引属性是用来定义冗余索引模式的。CHAR字段的冗余索引是从原字段值中按规则抽取部分值的组合并进行索引，允许从字段值中抽取多个冗余值。抽取规则采用正则表达式语法，每个正则表达式对应一个冗余值，正则表达式之间用换行符分隔，每行一个正则表达式。表达式语法规则请参见附录三。 ◆ 对BIT字段来说，只有它被设置成虚拟逻辑字段时才支持索引属性。虚拟逻辑字段的索引属性存储的是该虚拟逻辑字段所对应的实际字段列表（至少包含一个字段）及其逻辑关系和运算优先模式，如：“字段1, 字段2, *&”。定义中的‘*’是‘字段1’和‘字段2’之间的逻辑关系(AND运算符)，还可以是‘+’（OR运算符）、‘^’（XOR运算符）和‘-’（NOT运算符），逻辑关系缺省时默认为OR运算；‘&’是运算优先模式符号，表示“字段逻辑优先”，即虚拟逻辑在每个检索原子（检索键值）上进行扩展，缺省时表示“表达式逻辑优先”，即虚拟逻辑在虚拟逻辑字段所限定的最大子表达式上进行扩展。字段及逻辑关系之间必须用单字节的逗号‘,’进行分隔，逻辑关系符和运算优先模式也必须是单字节的，并且也必须放在最后。
词根索引语种	对DOCUMENT和PHRASE字段的索引进行词根处理时字段值的语种。可为Alpha（根据字母表确定一个主要的语种）、English（英语）、Russian（俄语）、French（法语）、Spanish（西班牙语）、Portuguese（葡萄牙语）、Italian（意大利语）、German（德语）、Dutch（荷兰语）、Swedish（瑞典语）、Norwegian（挪威语）、Danish（丹麦语）、Finnish（芬兰语）之一。缺省为Alpha。
混合索引模式	对DOCUMENT和PHRASE字段进行字词混合索引时，索引混合存储模式开关，可为None（不进行字词混合索引）、Char（禁止按词索引）、Full（全冗余模式）、Half（半冗余模式）、Zero（零冗余模式）之一。缺省为None。该开关只有当数据库指定了分词词典、扩展分词词典或稀疏词典时才有效。
最大长度或最大精度	数据库字段的长度限定。对CHAR字段为最大的取值长度，为零表示字段所允许的最大长度，超过规定长度的字符将被忽略，此时直接在后面的编辑框中进行编辑，缺省值为零；对NUMBER字段表示数值的小数位精度，此时可以点击右边的...按钮打开“最大精度”对话框，进行具体的精度选择，缺省值为100，表示双精度型数值，不带小数位；对其它类型的字段无意义。
显示大小	数据库字段在用户界面上显示的宽度和高度。
允许多值	选中表示允许多值（PHRASE、DOCUMENT型字段除外），缺省为选中。
不建索引	选中表示不建立索引（BIT型字段除外），对BIT型字段缺省值选中，对其它类型字段缺省值为清除。
值唯一性	选中表示取值不允许重复，仅对DATE、NUMBER和CHAR型字段有效，缺省为清除。
禁止空值	选中表示不允许空值，仅对DATE、NUMBER和CHAR型字段有效，缺省为清除。
缓冲数据	选中表示允许进行检索缓冲，缓冲可以提高检索时有关操作的速度。此开关决定该字段的“数据”是否集中存储，并在“检索缓冲”启动的情况下时，允许进入内存缓冲。虽然所有类型的字段都允许进行缓冲，但DOCUMENT和BIT类型的字段只存储一个标志，表示字段是否有值，而不是存储实际的字段值。缺省为清除。

控制名称	功能简介
公共检索	选中表示所有用户均可对本数据库字段进行检索，缺省为清除。
允许单字检索	对DOCUMENT和PHRASE字段进行字词混合索引时，字索引的方式开关：如果选中，表示所有的表意文字都可被单独查询，否则表示单个表意文字不能查询。该开关只有当数据库指定了分词词典，并且对该字段进行字词混合索引时才有效。缺省为清除。
允许相关排序	选中表示检索命中该字段时，该字段参与相关性排序，缺省为清除。对BIT字段该开关无效。
虚拟逻辑字段	如果虚拟逻辑字段开关被打开，则表示该字段被设置成一个虚拟字段。虚拟字段不能存储数据和索引，而只能在该字段的索引属性中定义一些它所对应的实际字段及其逻辑关系，在检索时作为单一的字段使用。只有BIT字段才能设置为虚拟逻辑字段。此功能为TRS全文数据库6.50.4000版本新增。
缺省虚拟字段	如果缺省虚拟字段开关被打开，则表示在检索时该虚拟逻辑字段的字段名可以用特殊的字符串“\$DUMMY”来替代，以便于检索表达式的书写或选项的指定。此功能为TRS全文数据库6.50.4000版本新增。

设计好数据库的字段结构以及各字段的具体属性后，按[完成]可结束向导，并创建数据库。按[上一步]可返回向导的第二步，按[取消]则退出创建数据库向导。

2.4 修改 TRS 数据库对象

如果满足下列条件之一，则有权修改数据库：

- 超级用户。
- DBA用户。
- 数据库的拥有者。
- 数据库具有公共ALTER属性。
- 与数据库的拥有者同用户组的GDA用户。
- 拥有该数据库的ALTER权的用户。

TRS 数据库建立后，一般不应对该数据库的属性或结构进行修改。当数据库不为空时（数据库中有 1 条或以上记录），不能修改数据库的缺省检索字段、数据字符集、数据宿主系统、数据存放路径以及所有使用词典与索引开关，不可修改数据字段的数据类型、缺省字段值、字段值格式、字段值范围、最大长度或精度、索引属性、索引存放路径和所有索引开关，但可增加新字段或对字段重命名。当数据库不为空而确有必要修改结构时，可先把数据库中数据存储成装库格式，然后按要求建立新的数据库并把数据装入，最后删除原数据库。

数据库的属性包括常规属性、结构和访问权限三个方面，其中常规属性在属性窗口中的属性框里显示，结构和访问权限以对话框显示，用户可对其属性进行修改（当前登录用户必须对该数据库有修改权限）。修改完毕后，按属性窗口上的按钮  则保存对常规属性的修改。按对话框上的[确定]按钮，将保存对结构或者访问权限所作的修改，按[取消]

则放弃所作的修改并关闭对话框。若属性窗口已关闭，可选择菜单上的“文件|显示属性”命令，或按工具栏上的按钮重新打开属性窗口。

2.4.1 修改数据库的常规属性

在对象管理窗口里选中要查看属性的数据库对象，属性窗口将会看到如下图：

常规属性	
内部编号	2634
对象名称	CCTV_XHS
拥有者	SYSTEM
对象类型	数据库
说明信息	中央电视台
创建时间	2008.03.10 11:53:08
修改时间	2010.05.04 01:00:44
字段数	23
记录数	6202264
锁定状态	共享 (SHARE)
锁定用户	
占用空间	点击查看
高级属性	
别名列表	
存放路径	F:\TRS\TRS6\ibase
数据宿主系统	
使用语言	中文GB2312/GBK/GB18030
主分词词典	SYSTEM.SEGMENT_GB
附加分词词典	
停用词典	SYSTEM.STOP_GB
附加停用词典	
稀疏词典	SYSTEM.MYSPARSEDIC
词根索引	<input checked="" type="checkbox"/>
大小写敏感	<input type="checkbox"/>
忽略数字串	<input type="checkbox"/>
二元组索引	<input type="checkbox"/>
禁止压缩	<input type="checkbox"/>
禁止停用词典	<input type="checkbox"/>
公共属性	
公共检索	<input type="checkbox"/>
公共更新	<input checked="" type="checkbox"/>
公共索引	<input type="checkbox"/>
公共修改	<input type="checkbox"/>
公共删除	<input type="checkbox"/>

图 2—4 数据库属性

属性框里各控制的说明如下：

控制名称	功能简介	
常规属性	内部编号	数据库在系统中的内部标识号。
	对象名称	名称是TRS数据库的对象名，即数据库名，是访问、维护和管理该数据库的入口。数据库名不能与属于同一用户的其它数据库、视图以及各种词典同名。
	拥有者	数据库的拥有者。
	对象类型	对象的类型。附有“(自动分裂)”字样的为自动模式视图包含的数据库，这种数据库的属性、结构和权限均不允许修改。
	说明信息	不超过63个字符长度的数据库说明信息。
	创建时间	数据库的创建时间。
	修改时间	数据库的最后一次修改时间。
	字段数	数据库的字段数。
	记录数	数据库的记录数。
	锁定状态	数据库的当前访问状态，有如下几种取值：共享锁(SHARE)、访问互斥锁(ACCESS_EXCL)、修改锁(UPDATE_EXCL)、修改等待锁(UPDATE_WAIT)、拒绝访问锁(ACCESS_DENY)。
	锁定用户	锁定数据库的用户名。
占用空间	点击打开“数据库大小”对话框，统计数据库所占用的磁盘文件空间，包括总占用空间、数据空间、索引空间和各字段索引的大小。	
高级属性	别名列表	数据库的别名列表，多个别名之间以半角分号、逗号或空格分隔，单个别名长度不能超过31个字符，总长度（含分隔符）不能超过63个字符。
	存放路径	数据库存放的路径，必须是系统配置中的数据 and 索引路径之一。可以在下拉列表里选择。
	数据宿主系统	有一类应用：TRS系统只管理和维护索引，而数据存储在其他系统中。数据宿主系统就是用来描述存储数据的系统的信息。描述信息对TRS没有实际的意义，如果数据库没有数据宿主系统，则TRS系统负责管理数据库的数据。不存储数据不针对日期型、数值型和字符型字段，因一些特殊应用和性能要求，这三种字段值始终存储。
	使用语言	数据库所用的语言，有如下几种取值：中文GB2312/GBK/GB18030、中文BIG5、英语、统一码UTF8。
	主分词词典	该数据库按词索引时所引用的词典。如果不为数据库指定分词词典，则该数据库将按字索引。
	附加分词词典	该数据库按词索引时所引用的扩充词典。
	停用词典	该数据库按词索引时所引用的不允许索引的词典。
	附加停用词典	该数据库按词索引时所引用的不允许索引的扩充词典。
	稀疏词典	该数据库按词索引时所引用的允许索引的词典。如果不为数据库指定稀疏词典，则该数据库索引所有不属于停用词典和附加停用词典中的词汇。
	词根索引	决定数据库的短语型和全文型字段数据中的字母单词是否按词根进行索引。如果设置该开关，则索引不区分大小写。
	大小写敏感	决定数据库的字符型、短语型和全文型字段数据中的字母串是否区分大小写进行索引。
	忽略数字串	选中表示建立索引时忽略数字串。
	二元组索引	决定数据库的短语型和全文型字段是否建立二元组索引。
禁止压缩	决定数据库是否对数据和索引进行压缩。对于CPU速度慢而IO速度较快的系统，不压缩数据与索引对数据库的入库与检索速度会略有提高，但空间膨胀率会显著增大，所以在实际使用时总是压缩索引，以获得较好的时空比。	
禁止停用词典	禁止停用词典开关的作用是允许建立停用词的索引（即创建数据的索引不启用停用词典），检索时停用词是否参与运算由应用进行选择。此功能为TRS全文数据库6.50.4000版本新增。	
公共属性	公共检索	选中表示所有用户均可对该数据库进行检索。
	公共更新	选中表示所有用户均可对该数据库进行记录维护。
	公共索引	选中表示所有用户均可对该数据库的索引进行维护。
	公共修改	选中表示所有用户均可修改该数据库的结构。
	公共删除	选中表示所有用户均可删除该数据库。

如对数据库的属性作了修改，按  将保存所作的修改。

2.4.2 修改数据库的结构

若当前属性窗口显示的对象为数据库，直接选择菜单上的“文件|结构”命令，或点击工具栏上的  按钮，将会看到如下对话框：



图 2—5 数据库的结构对话框

数据库字段列表控制里列出数据库的所有字段。字段列表的每一行分别显示字段的名称、类型和说明，字段名前带有  标志的字段为新增字段，有  标志的字段为数据库进行全文检索的缺省全文字段名，缺省为最后一个 DOCUMENT 或者 PHRASE 字段，DOCUMENT 字段优先。用户可在字段列表上进行添加字段、修改字段属性、删除字段、为字段改名、调整字

段顺序(仅限于新字段)等操作。数据库字段列表控制的右上角有六个按钮，它们的作用分别如下：

-  添加新字段，快捷键为 **Insert**
-  删除当前选中字段，快捷键为 **Delete**
-  为当前选中字段改名，快捷键为空格，直接在字段名上单击也能起到同样作用
-  当前选中字段位置上移，仅限于新字段，快捷键为 **Shift+ ↑**
-  当前选中字段位置下移，仅限于新字段，快捷键为 **Shift+ ↓**
-  把当前选中字段设为缺省字段，快捷键为 **Shift+空格**

数据库字段列表控制下面是字段属性组框，组框里的控制显示当前数据库字段列表控制里选中的字段的各项属性，具体如下：

控制名称	功能简介
内部编号	数据库字段在系统中的内部标识号，新字段不显示此属性。
别名列表	数据库字段的别名列表，多个别名之间以半角分号、逗号或空格分隔，单个别名长度不能超过31个字符，总长度（含分隔符）不能超过63个字符。缺省值为空。
显示标签	数据库字段的显示标题，如不为空，概览时将作为列标题，细览时作为字段引导符。
格式限定	数据库字段的输入输出格式。 对于CHAR型字段，其格式规定为： <ul style="list-style-type: none"> • A 表示在相应位置上必须出现一个英文字母； • 9 表示在相应位置上必须出现一个数字； • 空格 表示在相应位置上必须出现一个空格； • X 表示在相应位置上可出现任意字符。 对于DATE和NUMBER型字段，则为字段值的输出限定格式，并且格式由用户确定，如“Y.MM.DD HH:MM:SS”、“L.DP”等。 如果限定格式第一个字符为‘@’，则表示其中存储的是限定格式所在文件的文件名，按右边的...按钮可打开“格式限定”编辑对话框，在此对话框中可以直接进行编辑或者浏览格式限定文件。缺省值为空，表示字段的值没有格式限制。对其它类型字段没有意义。
缺省取值	指在数据库加载时，对于未给出字段值的DATE、NUMBER或CHAR型字段所对应的缺省取值。如果第一个字符为‘@’，则表示其中存储的是缺省值所在文件的文件名。按右边的...按钮可打开“缺省取值”编辑对话框，在此对话框中可以直接进行编辑或者浏览缺省取值文件。
取值范围	DATE、NUMBER或CHAR型字段的取值范围，对其它类型字段没有意义。 取值范围的定义格式包括： <ul style="list-style-type: none"> • 枚举定义：格式为V1,V2,V3,Vn，即枚举该字段上所有可能取值； • 区间定义：格式为V1,...,V2，表示该字段上所有可能取值必须在最小值V1和最大值V2之间； • 混合定义：格式为V1,V2,...,V3,Vn，是以上两种方式的混合。 以上各格式中的逗号和点号都必须是半角符号。CHAR型字段的取值范围所给出的每个值串首尾不允许空格。DATE型字段的取值范围所给出的每个日期中只支持绝对（完全）年份，而不支持缺省世纪数的短格式年份，即总是认为缺省世纪数为0。只有DATE型字段的取值范围支持模糊格式，即允许日期格式中缺省一些成份，但对于区间定义定义的范围，其中的V1和

控制名称	功能简介
	V2“前模糊”的成份必须是相同的，即在“年月日时分秒”的顺序下，V1和V2的第一个确定的日期成份的位置必须是相同的。如果第一个字符为‘@’，则表示其中存储的是取值范围所在文件的文件名。按右边的...按钮可打开“取值范围”编辑对话框，在此对话框中可以直接进行编辑或者浏览取值范围文件。缺省值为空，表示字段的值没有范围限制。
索引路径	数据库字段索引的存放路径，必须是系统配置中的数据和索引路径之一。缺省值为空，表示字段索引路径和数据库存放路径一致。
索引属性	<p>按右边的...按钮可打开“索引属性”编辑对话框，在此对话框中可以直接进行编辑或者浏览索引属性文件。</p> <ul style="list-style-type: none"> ◆ 对DOCUMENT字段来说，即倒排索引属性名。多个属性名之间用‘;’分隔，如“SEG;SEN;POS”等。如果第一个字符为‘@’，则表示其中存储的是索引属性所在文件的文件名。对DOCUMENT字段缺省值为“SEG;SEN;POS”，对其它类型的字段缺省值为空。 ◆ 对于CHAR字段来说，索引属性是用来定义冗余索引模式的。CHAR字段的冗余索引是从原字段值中按规则抽取部分值的组合并进行索引，允许从字段值中抽取多个冗余值。抽取规则采用正则表达式语法，每个正则表达式对应一个冗余值，正则表达式之间用换行符分隔，每行一个正则表达式。表达式语法规则请参见附录三。 ◆ 对BIT字段来说，只有它被设置成虚拟逻辑字段时才支持索引属性。虚拟逻辑字段的索引属性存储的是该虚拟逻辑字段所对应的实际字段列表（至少包含一个字段）及其逻辑关系和运算优先模式，如：“字段1, 字段2, *&”。定义中的‘*’是‘字段1’和‘字段2’之间的逻辑关系（AND运算符），还可以是‘+’（OR运算符）、‘^’（XOR运算符）和‘-’（NOT运算符），逻辑关系缺省时默认为OR运算；‘&’是运算优先模式符号，表示“字段逻辑优先”，即虚拟逻辑在每个检索原子（检索键值）上进行扩展，缺省时表示“表达式逻辑优先”，即虚拟逻辑在虚拟逻辑字段所限定的最大子表达式上进行扩展。字段及逻辑关系之间必须用单字节的逗号‘,’进行分隔，逻辑关系符和运算优先模式也必须是单字节的，并且也必须放在最后。
词根索引语种	对DOCUMENT和PHRASE字段的索引进行词根处理时字段值的语种。可为Alpha（根据字母表确定一个主要的语种）、English（英语）、Russian（俄语）、French（法语）、Spanish（西班牙语）、Portuguese（葡萄牙语）、Italian（意大利语）、German（德语）、Dutch（荷兰语）、Swedish（瑞典语）、Norwegian（挪威语）、Danish（丹麦语）、Finnish（芬兰语）之一。缺省为Alpha。
混合索引模式	对DOCUMENT和PHRASE字段进行字词混合索引时，索引混合存储模式开关,可为None（不进行字词混合索引）、Char（禁止按词索引）、Full（全冗余模式）、Half（半冗余模式）、Zero（零冗余模式）之一。缺省为None。该开关只有当数据库指定了分词词典、扩展分词词典或稀疏词典时才有效。
最大长度或最大精度	数据库字段的长度限定。对CHAR字段为最大的取值长度，为零表示字段所允许的最大长度，超过规定长度的字符将被忽略，此时直接在后面的编辑框中进行编辑；对NUMBER字段表示数值的小数位精度，此时可以点击右边的...按钮打开“最大精度”对话框，进行具体的精度选择；对其它类型的字段无意义。
显示大小	数据库字段的显示宽度和高度。
允许多值	选中表示允许多值（PHRASE、DOCUMENT型字段除外），缺省为清除。
不建索引	选中表示不建立索引（BIT型字段除外）。
值唯一性	选中表示取值不允许重复，仅对DATE、NUMBER和CHAR型字段有效。
禁止空值	选中表示不允许空值，仅对DATE、NUMBER和CHAR型字段有效。

控制名称	功能简介
缓冲数据	选中表示允许进行检索缓冲，缓冲可以提高检索时有关操作的速度。此开关决定该字段的“数据”是否集中存储，并在“检索缓冲”启动的情况下时，允许进入内存缓冲。虽然所有类型的字段都允许进行缓冲，但DOCUMENT和BIT类型的字段只存储一个标志，表示字段是否有值，而不是存储实际的字段值。
公共检索	选中表示所有用户均可对本数据库字段进行检索，缺省为清除。
允许单字检索	对DOCUMENT和PHRASE字段进行字词混合索引时，字索引的方式开关：如果选中，表示所有的表意文字都可被单独查询，否则表示单个表意文字不能查询。该开关只有当数据库指定了分词词典，并且对该字段进行字词混合索引时才有效。
允许相关排序	选中表示检索命中该字段时，该字段参与相关性排序。对BIT字段该开关无效。
虚拟逻辑字段	如果虚拟逻辑字段开关被打开，则表示该字段被设置成一个虚拟字段。虚拟字段不能存储数据和索引，而只能在该字段的索引属性中定义一些它所对应的实际字段及其逻辑关系，在检索时作为单一的字段使用。只有BIT字段才能设置为虚拟逻辑字段。此功能为TRS全文数据库6.50.4000版本新增。
缺省虚拟字段	如果缺省虚拟字段开关被打开，则表示在检索时该虚拟逻辑字段的字段名可以用特殊的字符串“\$DUMMY”来替代，以便于检索表达式的书写或选项的指定。此功能为TRS全文数据库6.50.4000版本新增。

如对数据库的结构作了修改，按[确定]将保存所作的修改并关闭对话框。

2.4.3 修改数据库的访问权限

查看和修改当前选中数据库的权限，当前用户必须是 DBA 用户，或为选中数据库的拥有者，或对选中数据库有 ALTER 权限。数据库的访问权限有检索权限、更新权限、索引权限、修改权限和删除权限五种，字段只有检索权限。

若当前属性窗口显示的对象为数据库，直接选择菜单上的“文件|权限”命令，或按工具栏上的  按钮，将会看到如下对话框：



图 2—6 数据库的访问权限

“访问权限”属性页里各控制的说明如下：

控制名称	功能简介
用户组/用户	当前服务器上所有用户组/用户的列表。用户组/用户名前面的小图标表明了该用户组/用户的类型：表示用户组、表示全域管理员用户、表示组管理员用户、表示资源用户、表示登录用户。
字段	对象名所包含字段的列表，“(对象权限)”相当于对数据库的操作权限。
访问权限	当前选中的用户组/用户对指定对象/字段的访问权限，包括以下所列几项。 直接单击组里的复选框可修改所有处于选择状态的用户组/用户对所有选择的对象/字段的访问权限，不需按“确定”或“应用”即可立刻生效。
检索权限	选中表示当前用户组/用户对数据库和指定字段有检索权限。
更新权限	选中表示当前用户组/用户对数据库有记录更新权限。
索引权限	选中表示当前用户组/用户对数据库有索引维护权限。
修改权限	选中表示当前用户组/用户对数据库有结构修改权限。
删除权限	选中表示当前用户组/用户对数据库有删除权限。

按[确定]或[取消]将并关闭对话框。

2.5 数据库重命名

数据库重命名要求操作用户必须拥有 DBA 权限或为当前数据库的拥有者。先选中要

改名的数据库，然后选择“文件”菜单上的“重命名”，或按工具栏上的按钮，或直接按空格键，当前数据库名字的位置上将出现一编辑框；在编辑框中输入新的名字，然后按 ENTER 键，所选数据库将被重新命名为所输入的名字；如要取消重命名操作，可按 ESC 键。此外，还可以在属性窗口中直接修改，点击数据库，属性窗口将显示其相应的属性，在“对象名称”一栏可直接对数据库名进行修改，修改完毕点击工具栏上的按钮即可保存新名字。

2.6 删除 TRS 数据库对象

删除数据库。如果满足下列条件之一，则有权删除数据库：

- 超级用户。
- DBA用户。
- 数据库的拥有者。
- 数据库具有公共DROP属性。
- 与数据库的拥有者同用户组的GDA用户。
- 拥有该数据库的DROP权的用户。

使用 TRS 管理员工具删除 TRS 数据库对象，首先要打开对象管理的数据库标签，通过四种方法可以完成对数据库对象的删除。

方法一：选择要删除的数据库，按下鼠标的右键并选择“删除”或“剪切”；

方法二：选择要删除的数据库，选择菜单“编辑|删除”或“编辑|剪切”；

方法三：选择要删除的数据库，按下“Delete”键或“Ctrl+X”；

方法四：选择要删除的数据库，选择工具栏中的或按钮。

系统会出现如下对话框：



图 2—7 删除对象确认对话框

选择“是(Y)”删除数据库结构（同时删除所有数据）；

选择“取消”放弃删除操作。

“删除”和“剪切”的区别在于剪切操作把所选的数据库复制到剪贴板里，而删除则不进行复制。利用“剪切”和“粘贴”功能的配合可快速复制数据库的结构。

2.7 数据库/视图检索

检索是 TRS 里最常用的一种操作。使用记录维护窗口打开数据库或视图可方便地进行检索。由于视图是逻辑的数据库，检索方法与数据库相似，在这里一并介绍。词典是一种特殊的数据库，其检索和浏览方法有别于用户自定义数据库，在后面的章节里另外介绍。

用户可使用浏览器通过 Web 发布系统进行检索和浏览，有关内容请参阅相应资料。

2.7.1 一般检索

有几种方法可以对数据库/视图进行检索。最简单的方法是在对象管理窗口直接双击所要检索的数据库/视图进行无条件检索；另外可以先在对象管理窗口里选择目标对象（可选择多个，此时各目标对象必须包含至少一个相同名字和类型的字段），然后按回车，或者在工具栏的检索表达式输入框里输入检索表达式再按回车或者按按钮；比较复杂的检索和排序可以使用对象管理窗口里的“文件|检索”命令或记录维护窗口里的“检索|高级检索/排序”命令或者点击工具栏上的按钮进行高级检索和排序(详见下一小节的介绍)。

检索表达式表示你要查找的目标记录，为空表示查找所有记录。

例如：输入“中国*北京”表示在默认字段查询中国和北京同时出现的记录。输入“中

华%”表示在默认字段查询出现“中华……”的记录，这里%为通配符。输入“日期>1997.01.01 *作者=王%”表示查询日期为1997年1月1日以后（不含）并且作者是姓王的记录。

检索表达式可以包括合法的算术运算符、比较运算符、逻辑运算符、重复逻辑运算符、属性运算符、限定运算符、加权运算符、统计函数、检索函数、词典函数、通配符等，具体操作可参见附录中检索表达式构造规则。

2.7.2 高级检索与排序

在对象管理窗口里选择要检索的目标数据库/视图（可选择多个，此时各目标对象必须包含至少一个相同名字和类型的字段），然后选择“文件|检索”命令；或在记录维护窗口选择“检索|高级检索/排序”命令；或者点击工具栏上按钮，将出现以下对话框：

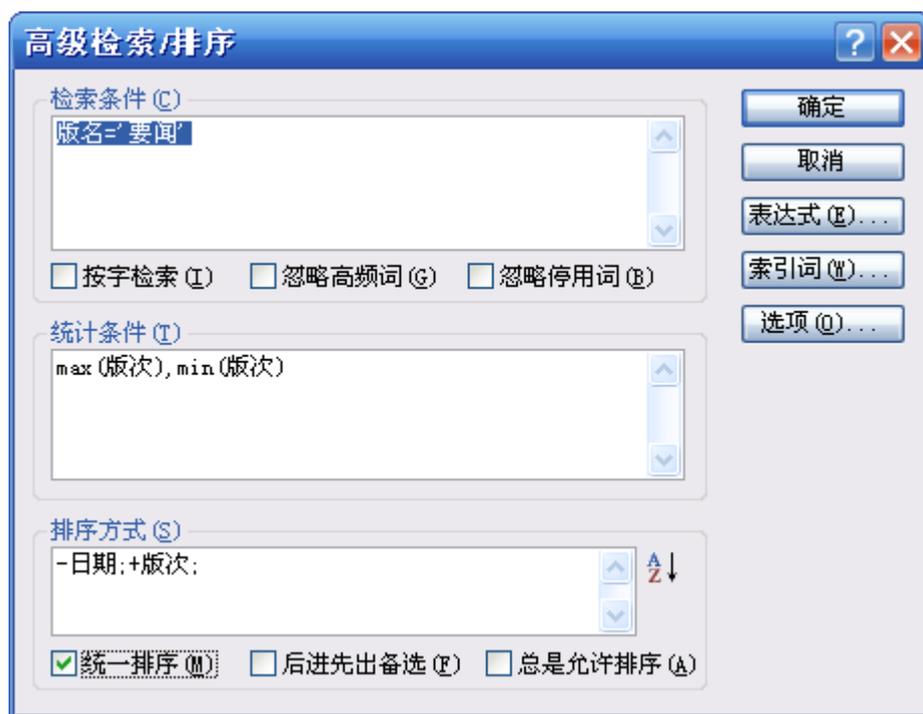


图 2—8 高级检索/排序对话框

对话框中各控制的说明如下：

控制名称	功能简介
检索条件	检索条件表达式。缺省情况下为当前检索结果的检索条件。

按字检索	选中表示表意文字按字进行检索。
忽略停用词典	选中表示检索时忽略停用词典。
统计条件	统计结果表达式，多个表达式之间以逗号分隔，可为空。所设统计条件应用于以后的各次检索中。要查看统计结果，可选择记录维护窗口的“文件 检索信息”命令。
排序方式	结果记录的排序方式，可直接输入，也可以点击旁边的  按钮打开“选择排序方式”对话框进行排序字段和选项的设置。
统一排序	对多个对象进行检索时，选中表示对检索结果进行混合排序，否则将按各数据库/视图分别排序。对单个对象检索没有作用。
总是允许排序	在命中记录数不超过系统配置的最高限值的情况下，“总是”允许排序检索结果。
表达式	打开“表达式生成器”对话框，帮助用户生成检索表达式。
索引词	打开“查询索引词”对话框，对数据库或者视图的索引词进行查询。
选项	打开“检索选项”对话框，设置检索的选项，所设选项将应用于以后的所有检索中。

设置好检索条件、统计条件、排序字段、检索选项等参数以后，按[确定]，TRS 管理员会以所设参数对当前检索目标对象进行检索，并刷新记录维护窗口。若此操作费时较长，可在过程中按 **Ctrl+Break** 中断操作，此时检索排序取消。按[取消]则放弃检索，关闭对话框。

2.7.3 表达式生成器

帮助用户生成检索表达式或者统计条件，对初学者尤为有用。在“高级检索/排序”对话框里，按[表达式]按钮，或者“复制记录”对话框和“新建导航文件”对话框里，按  按钮，或者在导航节点的“检索条件”或“统计条件”属性中，点击  按钮，都会弹出此对话框。

若当前焦点停留的编辑框为“检索条件”编辑框，表达式生成器如下图所示：



图 2—9 生成检索条件的表达式生成器

若当前焦点停留的编辑框为“统计条件”编辑框，表达式生成器如下图所示：

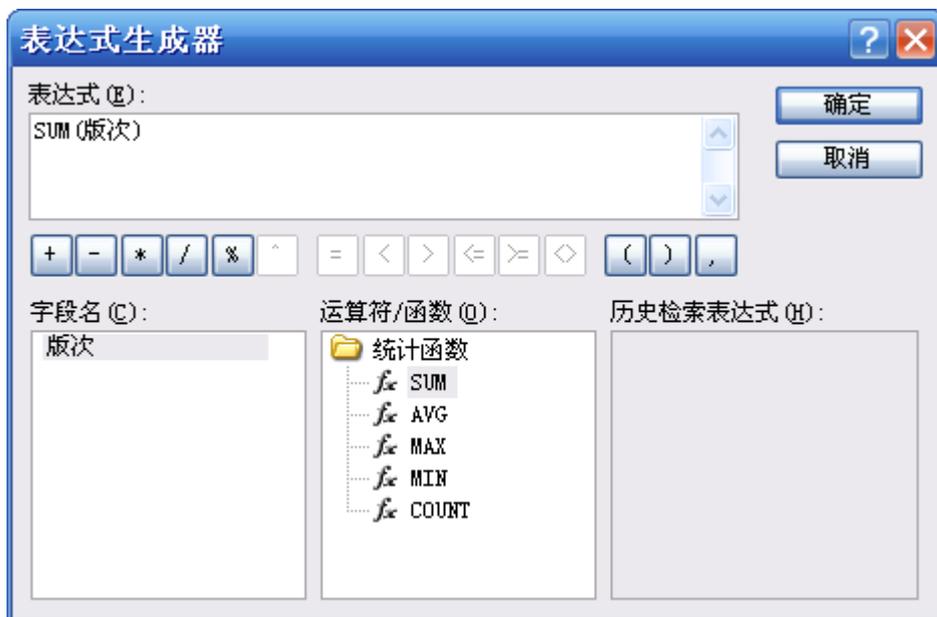


图 2—10 生成统计条件的表达式生成器

对话框中各控制的说明如下：

控制名称	功能简介
表达式	生成的检索表达式或者统计表达式。
运算符按钮	从  到  为常用的运算符按钮，按下按钮会把相应运算符复制到表达式编辑框里。
字段名	检索目标的所有共同字段（字段名和类型都一样）列表。双击可把字段名复制到表达式编辑框里。
运算符/函数	RS的运算符和函数列表。运算符分为逻辑运算符、重复逻辑运算符、属性运算符、限定运算符，函数分为检索函数和词典函数；当生成统计条件时函数为统计函数。双击文件夹可显示该项包含的所有运算符/函数列表，再双击某一项则可把该运算符/函数复制到表达式编辑框里。
历史检索表达式	历史检索表达式列表，列出最近用到的20个检索表达式，双击可把检索历史项复制到表达式编辑框里。当生成统计条件时此列表框无用。

生成合适的检索表达式或统计条件后，按[确定]，将关闭“表达式生成器”对话框并把检索表达式或统计条件复制到相应对话框的检索条件编辑框或统计条件编辑框里。按[取消]则关闭对话框。

2.7.4 选择排序方式对话框

选择排序方式对话框用于帮助用户生成排序方式，对初学者尤为有用。设置排序方式和选项，在“高级/检索排序”、“分类统计”、“复制记录”等对话框里，按  按钮，或者在导航节点的排序方式属性项中，点击  按钮，将弹出如下对话框：

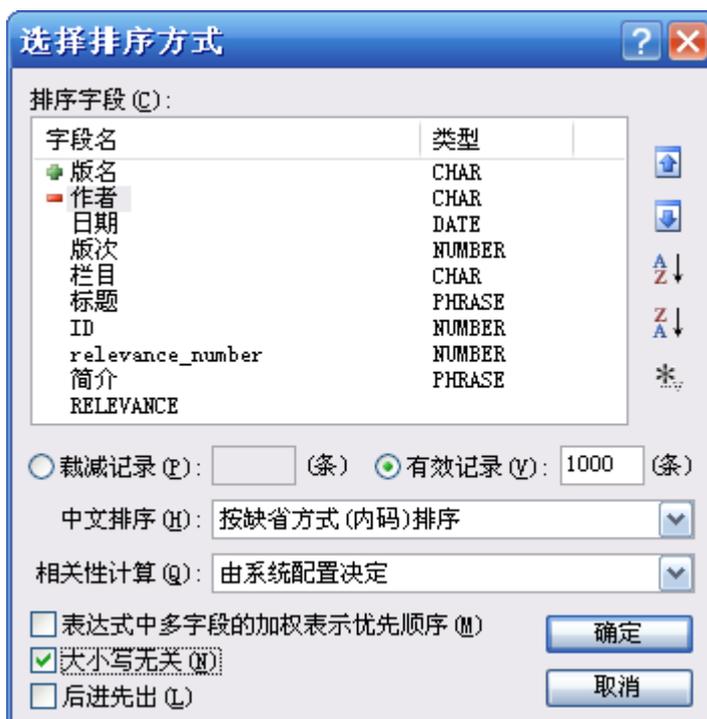


图 2—11 选择排序方式对话框

对话框中各控制的说明如下：

控制名称	功能简介
排序字段	<p>可以作为排序字段的字段列表，包括检索目标的所有共同非正文/二进制字段（字段名和类型都一样）和RELEVANCE（相关排序）。选择字段名，然后点击旁边的按钮 可把该字段的排序方式设为升序（前面显示 图标，RELEVANCE除外）；点 设为降序（前面显示 图标）；对于NUMBER字段点 按钮可把该字段设为按相关性排序时将相关度调整为“计算所得相关度值 * 该字段的值”（前面显示 图标）；对DATE字段点 按钮则弹出“分级排序”对话框，分别设置小时、天、周（顺序无关，内部将自动调整），表示以当前日期为起点，将记录中的日期字段值分级进行降序排序（前面显示图标）。用 和 按钮可调整已选字段的顺序。若字段前无任何图标，则表示此字段与排序方式无关。</p>
载减记录	<p>存储之前完成排序，并只存储排在前面的部分记录，同时自动设置“非精确命中点”模式。0表示由系统配置决定，缺省是不进行载减。如果值等于65535，则表示忽略系统配置，关闭排序载减功能。</p>
有效记录	<p>即部分排序，只保证前一部分记录是有序的。0表示由系统配置决定，缺省是全部排序。如果值等于65535，则表示则表示忽略系统配置，关闭部分排序功</p>

	能。
中文排序	中文的排序方式，可选择按内码、拼音、多音字拼音或笔画数排序。
相关性计算	不同的相关性的计算模式和IDF模式组合下会有不同的排序结果。分为7种情况：由系统配置决定；忽略词的位置和使用基于单个数据库范围的IDF值；忽略词的位置和词IDF值不参与相关性计算；考虑词的位置和使用基于单个数据库范围的IDF值；考虑词的位置和词IDF值不参与相关性计算；用一个记录中命中词的单位向量长度（每个命中词只计算一次）作为记录的相关；用一个记录中命中词的单位向量长度，以及命中词的词频和作为记录的相关度。
表达式中多字段的加权表示优先顺序	相关性排序时表达式中多字段的加权表示优先顺序。
大小写无关	选中表示忽略大小写。
后进先出	选中表示按记录的物理记录号的降序排列，不能和其它排序方式混用。

2.7.5 查询索引词

“查询索引词”对话框用于查询数据库或者视图里的索引词，索引词可作为整个检索表达式或检索表达式的一部分。在记录维护窗口选择“检索[查询索引词]”命令，或者在“高级检索/排序”对话框里，按[索引词]按钮，将弹出如下对话框：



图 2—12 查询索引词对话框

对话框中各控制的说明如下：

控制名称	功能简介
查询条件	查询表达式。可为“目标词”或“函数名(目标词, ...)”形式，其中“目标词”既可是个确定的词汇或字段值，也可是一个符合TRS规范的模糊词汇或模糊字段值。按右边的  按钮进行索引词查询，并把结果显示在结果列表里。若此操作费时较长，可在过程中按Ctrl+Break中断操作，此时查询取消。
显示个数	显示结果的个数，缺省为不设置。
当前页面	当前显示的页数。在编辑框中输入想要跳转的页数以后点击回车键，将会跳转到指定的页数。
结果列表	查询结果的索引词列表，每页显示个数可在“选项[集成环境]”里设定，缺省为100个。每个索引词显示其词语、命中数和词频数，如词语前方带  表示按词索引，带  表示按字索引，带  表示字词混合索引。要看其它页的索引词信息，可用  和  按钮进行前后翻页。点击“词语”、“命中数”和“词频数”标题可进行排序，三角形尖角向上表示升序，尖角向下表示降序。
对象名	待查询的目标数据库名或者视图名。改变查询对象将重新检索索引词并刷新检索列表。
字段名	待查询的字段名，改变查询字段将重新检索索引词并刷新检索列表。
复制	把结果列表里当前选中的检索词复制到“高级检索/排序”对话框的检索条件编辑框里。
关闭	关闭“查询索引词”对话框。
另存为	打开“词语导出”对话框，把当前查询结果的一部分或全部保存到一个文件里。

查询完毕后，按[复制]或直接双击结果列表的索引词项将把指定索引词复制到“高级检索/索引”对话框的检索条件编辑框里。按[关闭]则关闭对话框。

如果在查询索引词里按下[另存为]按钮，将出现如下对话框：



图 2—13 词语导出对话框

对话框中各控制的说明如下：

控制名称	功能简介
------	------

控制名称	功能简介
目标文件名	保存的目标文件名，可按右边的  按钮来选择目标文件名。
词语范围	需要保存的词语的范围，缺省为所有词语。
当前记录	只保存当前记录。
所有记录	保存检索结果里的所有记录。
指定范围	保存指定范围的记录，可用“起始记录号-终止记录号”的方式表示一定范围的记录，或仅用单个数字表示一条记录，多个分段之间以“;”分隔，各分段不能重叠。还可用“>词频数”或“>=词频数”的方式指定保存索引词的范围。
保存格式	结果保存的格式，对索引词典和索引词查询无效，缺省为TRS装库数据格式。
记录文件	以记录数据格式保存检索结果，一般只需浏览检索结果时选择此格式。
装库文件	以TRS装库数据格式保存检索结果，所保存的文件可以用来重新进行装库。
ISO2709	以ISO2709国际标准格式保存检索结果。
XML	以XML格式保存检索结果。
保存词频数	在保存索引词查询结果时，指定是否保存各索引词的词频数。在词典维护时无此项控制。

输入目标文件名，设置词语范围、保存格式等选项以后，按[确定]，TRS 管理员将开始保存查询结果到目标文件里。按[取消]则放弃保存操作并关闭对话框。由于保存查询结果是一个耗时过程，因此保存过程中有“保存进度”对话框显示目前已保存的记录数和保存完成的百分比。如要终止保存过程，可按[终止]按钮。

2.7.6 修改检索选项

用户可通过修改检索的高级选项来设置检索的环境。在“高级检索/排序”对话框里，按[选项]按钮，或者在导航节点的“缺省字段及选项”属性项中，点击  按钮，将弹出如下对话框：



图 2—14 检索选项对话框

对话框中各控制的说明如下：

控制名称	功能简介	
缺省目标字段	字段列表	检索条件表达式中缺省字段的子检索表达式的目标字段名集列表，字段名之间以半角分号分隔。为空表示在数据库数据词典中指定的字段上进行检索；为‘*’表示在所有DOCUMENT字段和所有PHRASE字段上同时进行检索。点击  按钮，可弹出选择缺省字段对话框，对缺省字段进行选择。
	逻辑关系	检索条件表达式中缺省字段的子检索表达式的目标字段名集各字段间的逻辑关系，可为：与、或、非、异或四种方式之一。
	优先于表达式中的逻辑	缺省字段检索的字段逻辑优先于检索表达式中的逻辑。

自动扩展检索	主题词典	利用知识辅助词典自动进行知识扩展检索时的主题词典。
	同义词典	利用知识辅助词典自动进行知识扩展检索时的同义词典。
	反义词典	利用知识辅助词典自动进行知识扩展检索时的反义词典。
	同义词	选中表示自动进行“同义词”扩展检索。
	反义词	选中表示自动进行“反义词”扩展检索。
	等同词	选中表示自动进行“等同词”扩展检索。
	替代词	选中表示自动进行“替代词”扩展检索。
	缩略词	选中表示自动进行“缩略词”扩展检索。
	被缩略词	选中表示自动进行“被缩略词”扩展检索。
	相关词	选中表示自动进行“相关词”扩展检索。
	上位词	选中表示自动进行“上位词”扩展检索。
	下位词	选中表示自动进行“下位词”扩展检索。
	所有上位词	选中表示自动进行所有“上位词”扩展检索。
	所有下位词	选中表示自动进行所有“下位词”扩展检索。
	同根词	选中表示自动进行“同根词”扩展检索。
	外文等同词	选中表示自动进行“外文等同词”主题控制扩展检索。
全角/半角	选中表示自动进行“半角/全角字符”扩展检索。	
简体/繁体	选中表示自动进行“中文简体/繁体”扩展检索。	
估算和加权	估算分段单位	使用估算检索模式，其值为数据的分段单位(记录数)。
	估算终止条件	使用估算检索模式，其值为检索终止条件(记录数)。
	标题加权系数	命中点在标题(第0段/第0句)位置上的加权系数(倍数)。
	知识扩展权重	知识扩展检索(主题/同义/反义)词的权重系数(百分比)。
	命中分支权重	命中分支数的权重系数。
命中点存储方式	不同的命中点存储模式，不但影响检索速度，还会影响相关性排序结果。	
检索超时时间	检索的超时时间(秒数，不能大于系统的的超时时间设置)。如果值为0，表示使用系统的的超时时间设置。	
集群裁减设置	集群对后台检索结果的裁减设置，仅在所登录的服务器为集群时才有有效，按右边的  按钮可打开对话框进行设置。0表示不作限制；1~9表示集群至少需要等待相应数量的后台返回结果；10~99表示集群至少需要等待相应比例的后台返回结果；100及以上表示至少需要返回的结果记录数。	
中断时输出部分结果	选中表示当检索过程被中断时也输出已经得到的部分检索结果，否则按出错处理。	
忽略不存在的索引属性	对于不存在的索引属性，简单地忽略，不按出错处理。	

设置好缺省目标字段和自动扩展检索选项以后，按[确定]保存这些选项并关闭对话框，所设选项将用于以后所有的检索操作中。如不想改变原有选项，按[取消]键关闭对话框。

2.7.7 检索附加参数

设置导航的附加参数，供用户保存一些附加导航信息以及供将来扩展时所用。选择“检索|附加参数”命令后，将弹出如下对话框：

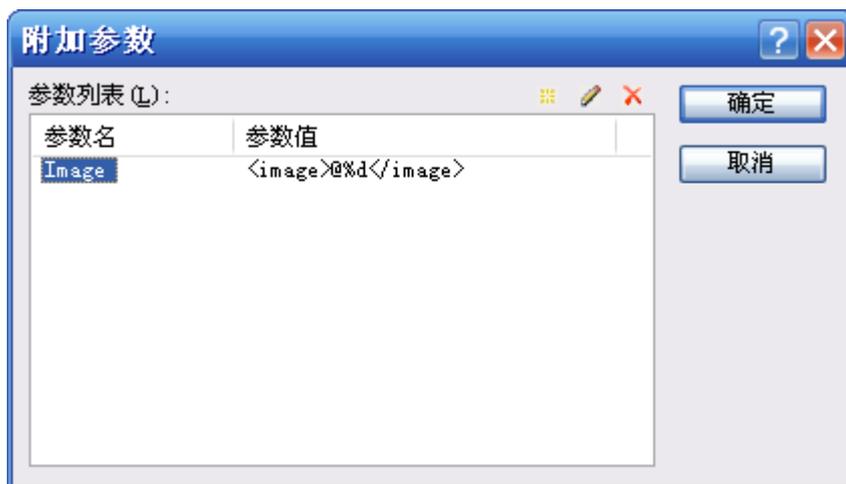


图 2—15 检索附加参数对话框

对话框中各控制的含义如下：

控制名称	功能简介
参数列表	显示当前设置的附加参数，这些参数将保存在导航文件里。右上角的  、  和  按钮分别用于添加、修改和删除参数。

设置好参数列表以后，按[确定]将把参数列表保存在导航文件的修改记载里，保存修改结果时会把这些设置写入导航文件，这样下一次打开此对象时将自动恢复上一次的设置。按[取消]放弃所有修改并关闭对话框。

2.7.8 二次检索和排序

在记录维护窗口里，可以在上一次检索的基础上进行二次检索，以缩小查找的范围或对上次检索结果进行重新排序。

在记录维护窗口的选择一段文字，然后选择“检索|按选定内容检索”命令或按下工具栏中的  按钮。可显示所有当前字段包含当前选择内容的记录。选择“检索|内容排除检索”命令或按下工具栏中的  按钮。可显示所有当前字段不包含当前选择内容的记录。（注意：如果被选字段是多值字段，而且您选定的检索内容也是多值字段的话，您需要将检索表达式对话框中显示的表达式中的“;”替换为“*”或“and”。例如：作者多值字段：“王一;李二”，检索表达式：“作者='王一;李二'”，修改检索表达式为：“作者='王一' * 作者='李二'”。）

在浏览检索结果时单击字段标题可对检索结果直接排序。选择菜单“检索|降序排列”命令，可按当前列内容的降序排列记录。选择菜单“检索|升序排列”命令，可按当前列内

容的升序排列记录。若排序操作费时较长，可在过程中按 **Ctrl+Break** 中断操作，此时排序取消。

以上命令仅对概览窗口（排序时不能为 **DOCUMENT/BIT** 字段）、细览窗口（当前不是处于源文件浏览状态）和表单编辑窗口有效，且检索结果是针对当前检索结果的二次检索。如想在当前导航结果的基础上进行检索，请先选择“检索|取消检索/排序”命令，或直接在工具栏上的检索表达式组合框里输入检索条件并按回车。如想直接对检索的条件和排序方式等进行修改，可选择“检索|高级检索”命令。

2.8 浏览和处理检索结果

浏览检索结果在记录维护窗口里实现。该窗口可有四种方式：只有概览、只有细览、上下分隔、表单编辑，一般缺省为上下分隔方式（上方为概览窗口，下方为细览窗口，边界可以上下调整）。左侧上部的导航窗口显示当前对象对应的导航树，左侧下部的属性窗口显示当前对象的检索属性和上方导航树的属性。

对于查询检索结果可通过菜单“记录”浏览到各条记录，也可通过浏览维护窗口的状态栏中的小按钮定位到各条记录，如下图：



图 2—16 记录维护窗口的状态栏

状态栏最左边部分用于显示菜单命令、工具栏按钮、操作结果等提示信息。中间为记录定位部分，显示当前记录号，并可快速定位到某条记录。后边的两个状态框分别显示键盘的数字键和大小写状态。最右边的  和  按钮则分别用于显示对象管理窗口和打开命令语言处理工具。

记录定位部分中间的编辑框显示当前记录号，最右端显示当前检索结果的总记录数。如要直接定位到某条记录，请在该编辑框内输入目标记录的序号，然后按回车键。记录定位部分还有几个小图标按钮，它们的作用分别是：

-  首记录 定位到第一条记录
-  上一记录 定位到上一条记录
-  下一记录 定位到下一条记录
-  尾记录 定位到最后一条记录
-  添加记录 添加新记录到数据库
-  修改记录 修改当前记录的内容
-  删除记录 从数据库中删除记录

Ctrl+Home	移到整个表格第一行
Ctrl+End	移到整个表格最后一行
PageUp	移到上一页第一行
PageDown	移到下一页最后一行

表格列 修改单元格的显示标题和宽度。选择“查看|表格列”命令后，TRS 管理员弹出如下对话框：



图 2—18 表格列对话框

对话框中各控制的说明如下：

控制名称	功能简介
字段列表	当前显示的所有概览字段。
显示标签	当前选定字段的显示标题，缺省为该字段的字段名。
显示宽度	当前选定字段的显示宽度，缺省情况下为字段规定的显示宽度（以字符为单位）。
缺省宽度	选中表示将该字段的显示宽度设为缺省值。

修改完各概览字段的显示标题和宽度后，按[确定]，所作修改将影响概览窗口的显示。按[取消]则放弃所作修改并关闭对话框。注意：所作修改并不影响下一次打开该记录维护窗口，如要修改字段标题和宽度的缺省设置，请修改字段的显示标题和显示宽度属性。

2.8.2 细览窗口

细览窗口以浏览器的方式显示用户指定的细览字段值，可直接显示记录所带的 jpg 或者 gif 类型图片，其他类型的附件在默认状态下以连接形式显示，双击带下划线的文字以打开或保存该文件或连接。但处于直接浏览记录附件状态（即按钮按下）时，点击记录时将直接显示该记录所带的附件，或者弹出对话框询问用户是否打开或保存所带的附件，选择[打开]，将在细览窗口中直接打开此附件，选择[保存]，则可将此附件保存到指定的目录中。显示的字段列表可以通过选择“检索|显示字段”命令或设置导航节点的显示方式来改变。

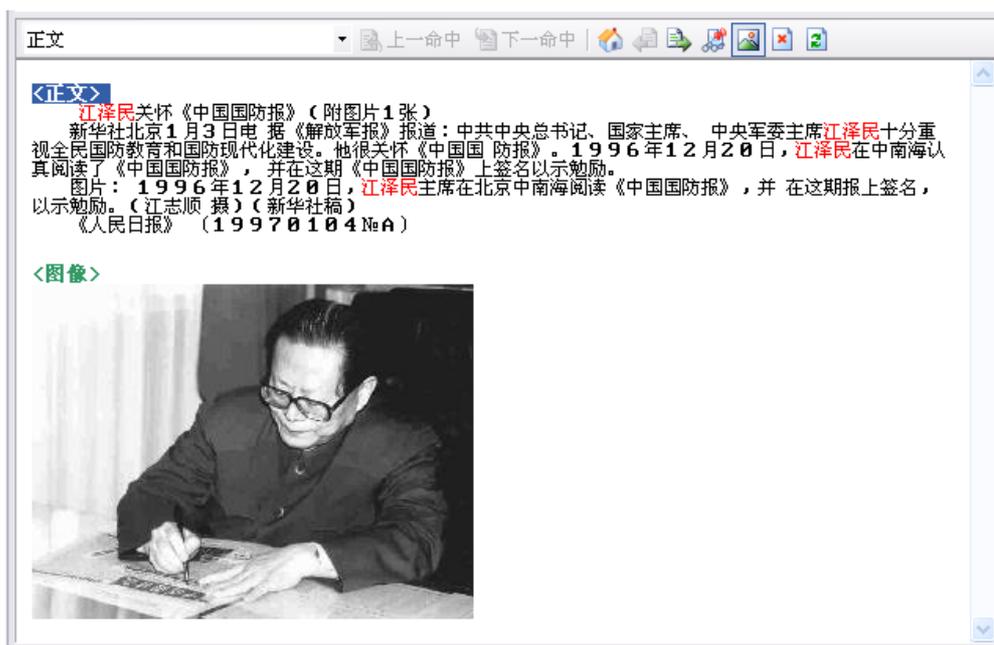


图 2—19 细览窗口

上方工具栏最前端是一个字段列表框,其下拉框中列有当前细览字段的列表,选择相应的字段会直接定位到该字段处。工具栏上其他各按钮的具体功能如下:

 上一命中	定位到上一命中点处。对当前对象进行检索后,细览窗口中短语型(PHRASE)和全文型(DOCUMENT)字段中的命中点会以特殊的颜色显示,便于查看。点击此命令后,将定位到当前位置前的上一个命中点处,命中点以蓝色反显。
 下一命中	定位到下一命中点处。对当前对象进行检索后,细览窗口中短语型(PHRASE)和全文型(DOCUMENT)字段中的命中点会以特殊的颜色显示,便于查看。点击此命令后,将定位到当前位置后的下一个命中点处,命中点以蓝色反显。
	返回显示细览字段和附件列表的主页。在此版本中,细览窗口直接用浏览器的方式对细览字段进行显示,记录包含的图片被直接显示,记录包含的其他类型的附件以连接的方式显示。如果当前显示的是记录所带附件,点击此命
	打开当前记录的上一个附件。若当前处于附件直接浏览状态(即  按钮按下),且当前所浏览的附件不是当前记录的头一个附件时,选择此命令可以
	打开当前记录的下一个附件。若当前处于附件直接浏览状态(即  按钮按下时),且当前所浏览的附件不是当前记录的最后一个附件时,选择此命令

	直接浏览记录的附件。此按钮处于按下状态时，点击记录时将直接显示该记录所带的附件，或者弹出对话框询问用户是否打开或保存所带的附件，选择[打开]，将在细览窗口中直接打开此附件，选择[保存]，则可将此附件保存
	直接显示记录所带的 jpg 或者 gif 类型图片。此按钮处于按下状态时，点击记录将直接显示该记录所带的 jpg 或者 gif 类型图片，若是其他类型的图片还是显示其连接；若此按钮未按下，点击记录时则只显示连接而不显示实际
	停止显示当前附件。在显示当前记录所带附件时，若附件过大或过多，显示时间较长，可点击此命令停止显示当前附件。
	重新装入并显示当前附件

细览窗口的快捷菜单如下所示：

按选定内容检索	显示所有包含当前选择内容的记录
内容排除检索	显示所有不包含当前选择内容的记录
复制	复制所选内容到剪贴板上

在细览窗口中，可以用以下编辑键移动选择文本：

←	插入点左移一个字符
Ctrl+←	插入点左移一个单词
→	插入点右移一个字符
Ctrl+→	插入点右移一个单词
↑	插入点上移一行
↓	插入点下移一行
Home	插入点移到行首
End	插入点移到行尾
PageUp	屏幕上滚一屏
PageDown	屏幕下滚一屏
Shift+移动键	选择文本
Ctrl+Home	滚动到文本的头一个字符
Ctrl+End	滚动到文本的最后一个字符
Ctrl+PageUp	滚动到屏幕最顶端
Ctrl+PageDown	滚动到屏幕最底端

2.8.3 表单窗口

表单窗口以表单方式显示用户指定的字段值，每一个字段对应一个编辑框，可直接在编辑框里修改字段值。对于 DOCUMENT/BIT 型字段可点击右边按钮  或者  添加文件，点击按钮  删除文件。显示的字段列表可以通过选择“检索|显示字段”命令或设置导航节点的显示方式来改变。

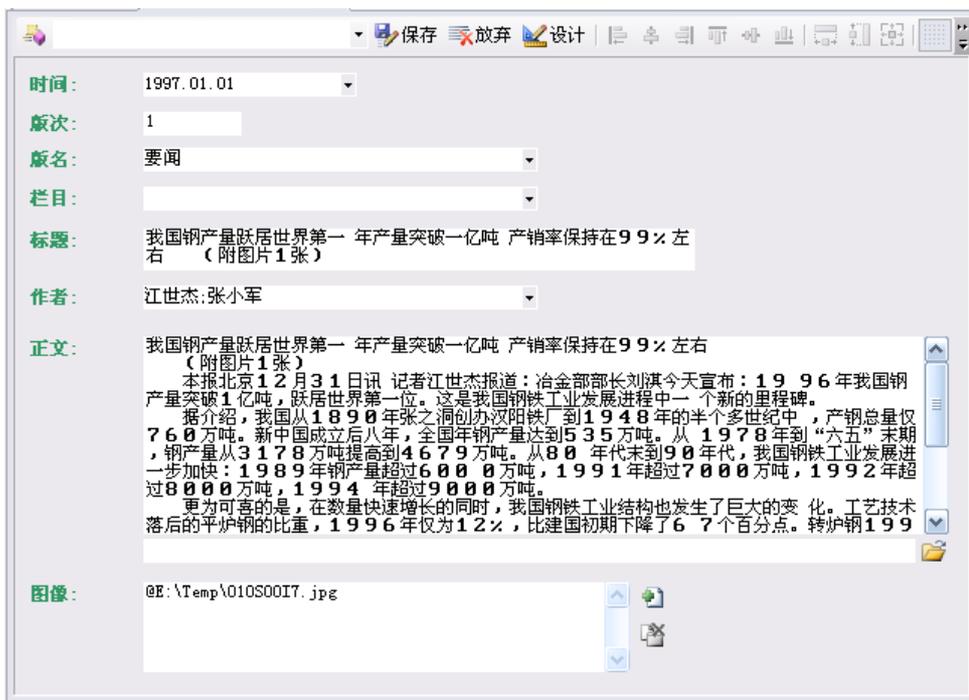


图 2—20 表单窗口

同时,日期、字符型字段可以进行可能值的选择:



图 2—21 日期型字段的选择

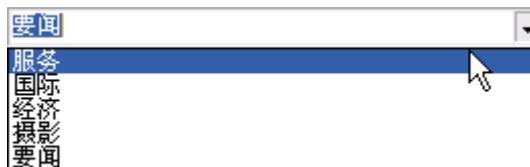


图 2—22 字符型字段可能取值的选择

上方工具栏最前端是当前记录的标志图标,不同的标志代表不同的记录状态,如下所示:

	当前记录为新纪录
	当前记录已设密码
	当前操作者的密级不够查看当前记录
	当前操作者不属于当前记录的开放用户组

	当前记录已被开锁
	当前记录已被修改
	读取当前记录失败
	当前记录已在别处被删除

本版本表单窗口新增了编辑框位置和大小设计。点击  按钮，所有编辑框变灰，表单窗口处于控件设计状态。点击某一个编辑框，可对其大小进行改变或者随意拖动其位置。修改完以后点击  按钮即可保存这些设置，下次再打开此对象的表单窗口时，将自动恢复上一次对各编辑框的设置。如下图：

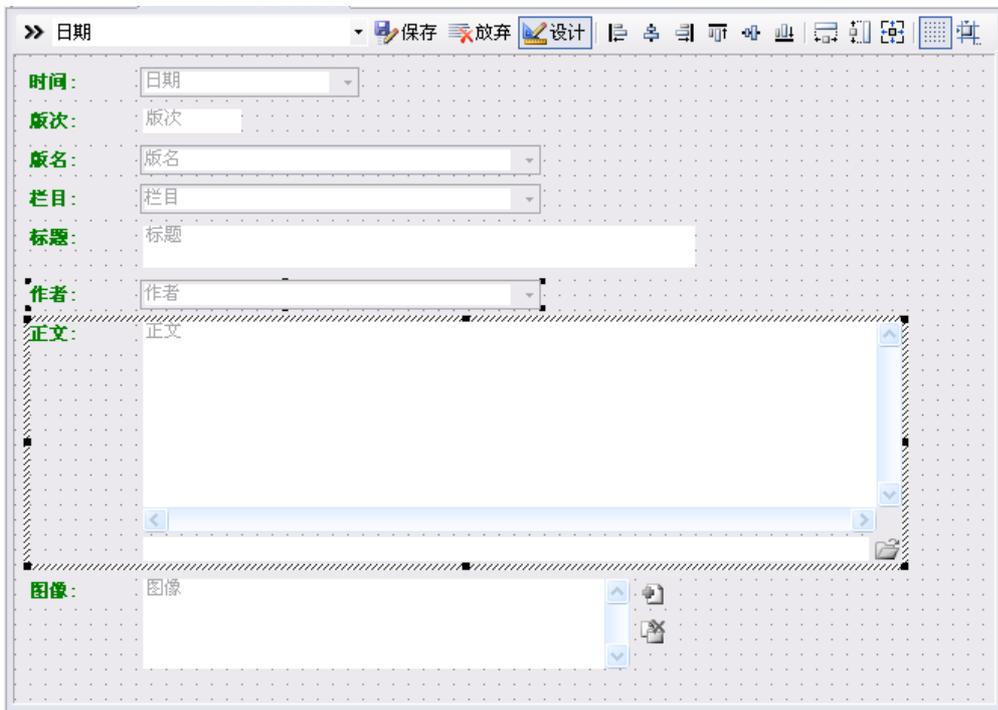


图 2—23 控件设计状态下的表单窗口

标志后面是一个字段列表框，其下拉框中列有当前可修改的各字段的列表，选择某一字段，输入焦点将定位到相应的字段编辑框处。工具栏上其它按钮的具体功能如下：

	保存当前所做的修改并返回原先的浏览状态
	放弃当前所做修改并返回原先的浏览状态

 设计	对当前窗口中各编辑框的大小及位置进行修改
	所选控件靠左对齐
	所选控件水平居中对齐
	所选控件靠右对齐
	所选控件靠上对齐
	所选控件垂直居中对齐
	所选控件靠下对齐
	使所选控件的宽度相同
	使所选控件的高度相同
	使所选控件的大小相同
	以网格方式显示界面
	弹出“布局设置”对话框，对网格的间距和标签的宽度等进行设置

表单窗口的快捷菜单如下所示（仅针对窗口里当前拥有输入焦点的编辑框）：

按选定内容检索	显示所有包含当前选择内容的记录
内容排除检索	显示所有不包含当前选择内容的记录
剪切	删除所选内容并把它放到剪贴板上
复制	复制所选内容到剪贴板上
粘贴	插入剪贴板内容
删除	删除所选内容
全部选定	选定所有内容

在表单窗口中，可以用以下编辑键切换当前编辑框和编辑文本：

←	插入点左移一个字符
Ctrl+←	插入点左移一个单词
→	插入点右移一个字符
Ctrl+→	插入点右移一个单词
↑	插入点上移一行
↓	插入点下移一行
Home	插入点移到行首
End	插入点移到行尾

PageUp	屏幕上滚一屏
PageDown	屏幕下滚一屏
Shift+移动键	选择文本
Enter	换行
Delete	删除插入点后的字符或所选内容
Ctrl+Delete	删除插入点后的单词
Back	删除插入点前的字符
Ctrl+Back	删除插入点前的单词
Ctrl+Home	滚动到文本的头一个字符
Ctrl+End	滚动到文本的最后一个字符
Ctrl+PageUp	滚动到屏幕最顶端
Ctrl+PageDown	滚动到屏幕最底端
Ctrl+Shift+Tab	插入一个制表符
Tab	把输入焦点转移到下一编辑框
Shift+Tab	把输入焦点转移到上一编辑框

2.8.4 查看检索结果属性

检索结果信息和统计结果早打开记录维护窗口时将直接显示在属性窗口中



图 2—24 检索结果属性

对话框中各控制的说明如下：

控制名称	功能简介
对象列表	进行统计的对象，可对所有对象进行综合统计或分别对每个对象进行单独统计。
字段数	检索结果的字段数，相当于所有检索目标数据库的共同字段（字段名和类型都相同的字段）的总数。
总记录数	检索结果的总命中记录数。
有效记录数	实际有效的、能够读取的结果记录数。
命中点数	检索结果的总命中点数，只有在检索时没有指定“无命中点”选项才有效。
检索消耗时间	检索请求所消耗的时间（毫秒数）。如果为0，则表示服务器（可能是版本较低）没有进行时间统计。
估算部分结果	选中表示结果集是估算检索的部分结果，否则表示结果集是精确检索的完整结果。
统计结果	统计子表达式的值。属性项左边显示一个统计子表达式，属性项右边显示其统计结果值。若统计条件只为count(*)时，因其结果实际上就是命中记录数，所以此种情况下统计结果不作显示。

2.8.5 检索结果导出

当前检索结果可以多种方式保存到文本文件里。选择“文件|另存为”命令后，TRIS 管理员弹出如下对话框：



图 2—25 检索结果导出对话框

对话框中各控制的说明如下：

控制名称	功能简介
目标文件名	保存的目标文件名，可按右边的  按钮来选择目标文件名。
保存字段	可选字段 可以选为保存字段的字段列表，可按  按钮把当前选中字段选到已选字段里。
	已选字段 已经被选为保存字段的字段列表，缺省情况下为所有的概览/细览字段。可按  按钮把当前选中字段清除出已选字段，或用  和  按钮调整保存字段的顺序。
记录范围	当前记录 只保存当前记录。
	所有记录 保存检索结果里的所有记录。为缺省选项。
	指定范围 保存指定范围的记录，可用“起始记录号-终止记录号”的方式表示一定范围的记录，或仅用单个数字表示一条记录，多个分段之间以“,”分隔。
	每批读取 设置每批读取的记录数，最大值是5120，不设置即和系统的默认值一致。

	每文件限制	设置每个文件存放记录的条数，清空或为零表示不限制，最大值为FFFFFFD。
保存格式	保存格式	结果保存的格式，对索引词典无效，缺省为TRS装库数据格式。
	记录文件	以记录数据格式保存检索结果，一般只需浏览检索结果时选择此格式。
	装库文件	以TRS装库数据格式保存检索结果，所存文件可以用来重新进行装库。
	ISO2709文件	以ISO2709国际标准格式保存检索结果。
	XML文件	以XML格式保存检索结果。
	传输装库文件	以传输格式保存检索结果。传输格式是数据加载时用于传输一批记录的内部格式文件。这种格式的一个最大优点是能够将文本与非文本数据打包在一个文件里，避免大量外挂文件可能引起的性能问题。
	保存记录ID	保存记录时保存记录的ID。
	压缩字段值	存储字段值时进行压缩，仅可与传输格式组合使用。
	添加记录前缀	记录及字段标记前添加的引导前缀，只对装库文件格式有效。
文件选项	保存附件方式	保存记录里包含的源文件的模式。
	附件存放目录	指定记录所含附件存放的目录。
	每组容纳文件	指记录附件存放的目录中，每个目录容纳多少个附件，最大值为32768。若为0或者清空表示不限制，即所有附件存入目标文件所在的目录中；若指定了数值比如为10，则每个目录中存储10个附件，附件目录自动按序号生成。
	最大文件长度	设置保存的目标文件的最大长度，清空或为零时取缺省值2G。若当前所要保存的记录大小超过此数值，将自动分为N个大小为此数值的目标文件进行保存，如Demo2_1.tris、Demo2_2.tris...
	超时重试时间	读取记录超时后等待重试的时间，以秒为单位。

输入目标文件名，设置保存字段、记录范围、保存格式等选项以后，按[确定]，TRS 管理员将开始保存检索结果到目标文件里。按[取消]则放弃保存操作并关闭对话框。由于保存检索结果是一个耗时过程，因此保存过程中有“保存进度”对话框显示目前已保存的记录数和保存完成的百分比。如要终止保存过程，可按[终止]按钮。

2.8.6 发送检索结果

选择“文件|发送”命令，TRS 管理员将首先打开“另存为”对话框，把检索结果的指定部分保存到一个临时文件里（文件名由 TRS 管理员决定）。然后调用相应的邮件处理软件（如 Microsoft Outlook Express 等），用户可以利用该软件将检索结果临时文件发送到其他用户手里。最后，TRS 管理员自动删除检索结果临时文件。

2.8.7 打印检索结果

选择“文件|打印”命令，TRS 管理员将首先打开“另存为”对话框，把检索结果的指定部分保存到一个临时文件里（文件名由 TRS 管理员决定），然后弹出如下对话框：

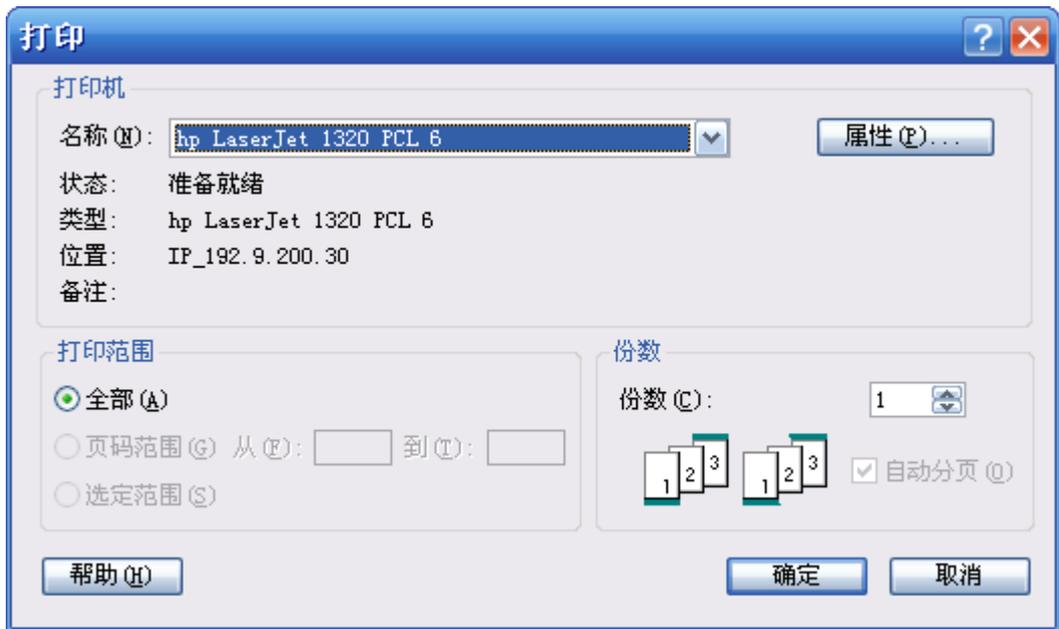


图 2—26 打印选项对话框

在对话框中选择目标打印机，输入打印份数后，按[确定]键后，TRS 管理员将开始打印检索结果临时文件，按[取消]则放弃打印操作。由于打印文件是一个耗时过程，因此打印过程中会显示“打印进度”对话框，显示当前已打印的页数和打印完成的百分比，按[终止]按钮将中断打印过程。打印完毕后，TRS 管理员将自动删除检索结果临时文件。

打印纸张及打印的页面参数等由“页面设置”对话框设定。需要打印前，先选择“文件|页面设置”命令，TRS 管理员弹出如下对话框：

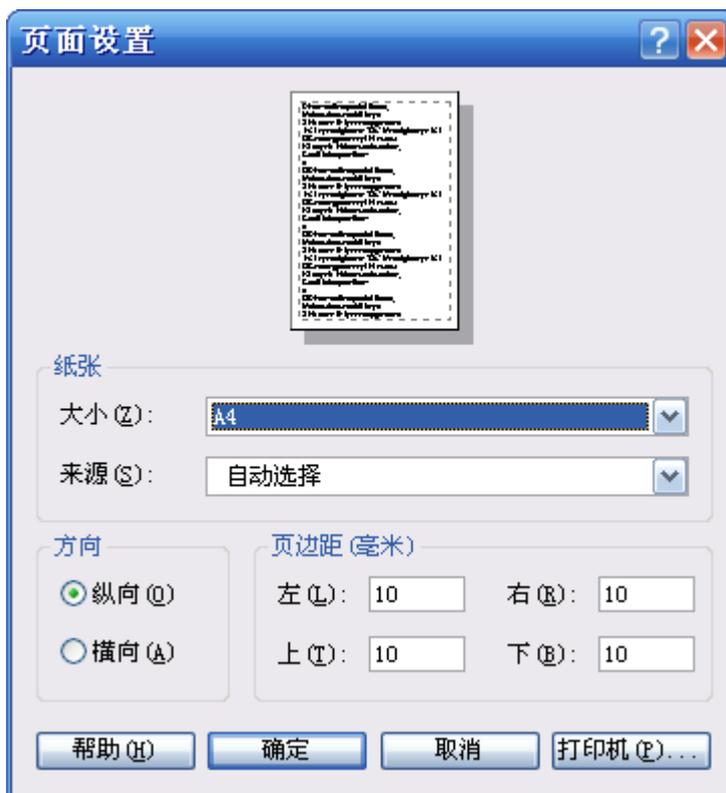


图 2—27 页面设置对话框

对话框中各控制的说明如下：

控制名称	功能简介
纸张	指定打印纸张的大小和来源。
大小	指定要使用的纸张或信封大小。
来源	指定纸张在打印机上的位置。不同打印机有不同的纸张来源，如上层纸盒、信封送纸器和手动送纸器等。
方向	显示整个文档在整个页面上的位置。可为横向或纵向，缺省为纵向。
页边距	设置页面的打印区域，此处设置的边距不在打印范围内。缺省情况下，页面的上下左右四个方向各留下10mm的距离。

每改变一个选项，对话框上部的预览画面都会相应地作出调整。按[确定]，所作设置将作为打印时的页面设置；按[取消]将放弃所作的设置并关闭对话框；按[打印机]将选择打印的目标打印机。

2.8.8 分类统计结果

选择“检索|分类统计”命令，将出现以下对话框：

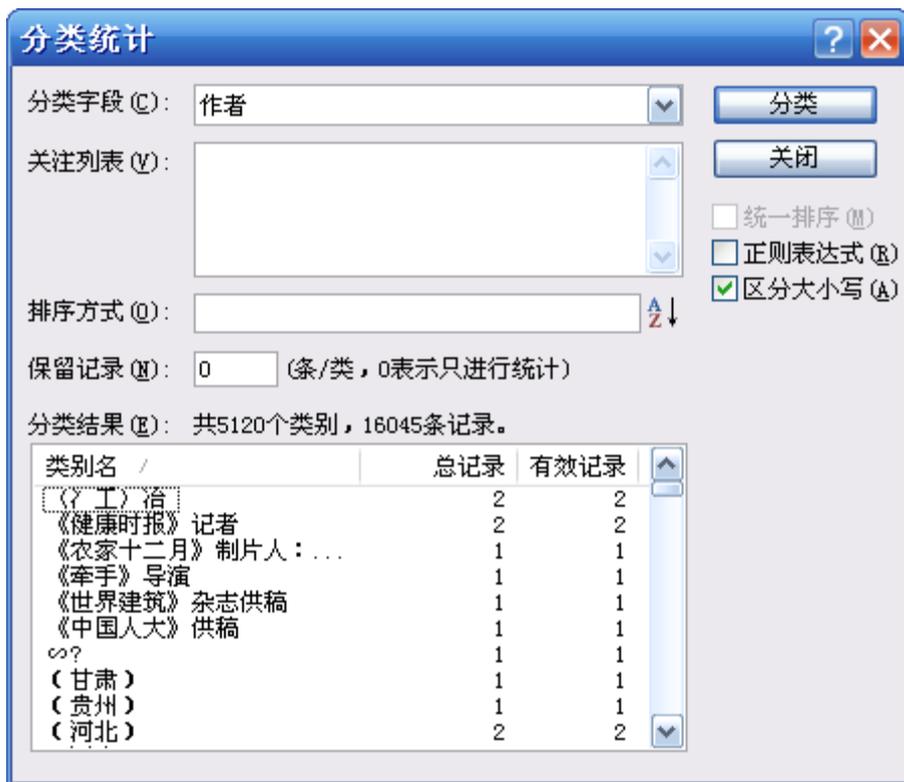


图 2—28 分类统计对话框

对话框中各控制的说明如下:

控制名称	功能简介
分类字段	指定按值分类的字段, 可以是日期、数值或字符类型字段。
关注列表	指定分类时只需关注的字段值, 也就是说只对这些字段值所对应的结果记录进行分类统计和存储, 每个值为一行 (注意不是用其它符号作为分隔符)。如果为空, 则缺省为所指定的字段的所有字段值。但类别的总数 (识别的字段值个数) 取决于系统配置所设定的最大数。
排序方式	结果记录的排序方式, 为空表示不改变原来的排序方式。点击 按钮弹出“选择排序方式”对话框, 帮助用户生成排序方式表达式。
统一排序	选中表示在检索结果所涉及的所有数据库上进行混合排序, 否则将在每个视图或物理数据库中进行独立排序。
正则表达式	选中表示关注列表里给出的是一个正则表达式, 用以筛选符合条件的字段值。
区分大小写	选中表示字段值区分大小写。
保留个数	每个类中保留结果记录的个数: 0表示对结果记录只进行类别统计, 而不改变结果记录的存储; 65535表示对结果记录进行类别统计, 并按类别存储和保留所有的结果记录; 其它值表示对结果记录进行类别统计, 并按类别存储, 每个所保留的结果记录数不超过这里所指定的值。
结果列表	分类统计的结果列表, 每行显示其类别名、总记录数和有效记录数。单击类别名标题可改变其排列顺序, 缺省为升序排列。
分类	执行分类统计并把结果显示在列表里。若此操作费时较长, 可在过程中按

	Ctrl+Break中断操作，此时分类取消。
关闭	关闭“分类统计结果”对话框。

设置好分类字段、关注列表、排序方式、保留个数等参数以后，按[分类]，TRIS 管理员会以所设参数对当前检索结果进行分类统计，并把结果显示在列表里。按[关闭]则关闭对话框，如果分类操作对检索结果作了改变将刷新浏览窗口，直至重新进行检索或排序操作。

2.9 建立和使用导航

记录维护窗口对应的导航窗口显示导航树。导航树的每一个节点记录了检索条件、排序方式、统计表达式以及检索选项等信息，选中一个节点将根据这些信息重新进行检索，并按节点设置的显示方式显示检索结果。此外，导航信息还包括窗口标题、显示字段、记录范围等信息，这些信息统一地存放在指定导航目录下的一个文件里，由 TRIS 管理员进行统一管理。

本版本的导航窗口新增了检索条件具有蕴含关系的多级导航功能，即在继承根节点的检索条件下根据二级导航节点再次进行检索，如下图中，即是在：“读者来信”的检索结果中按照“日期”进行了第二次检索。以此类推，还可进行更多级的检索。但需要注意的是，此多级导航功能只有在使用“自动生成”功能的情况下才能生成，其他情况下生成的多级导航中，根节点和子节点代表的检索条件没有蕴含关系。

导航树的维护可直接在该窗口内进行，也可以在外部编辑一个导航文件，并把它载入到当前窗口里。同样，当前窗口的导航树也可以导出到一个外部的文件中。

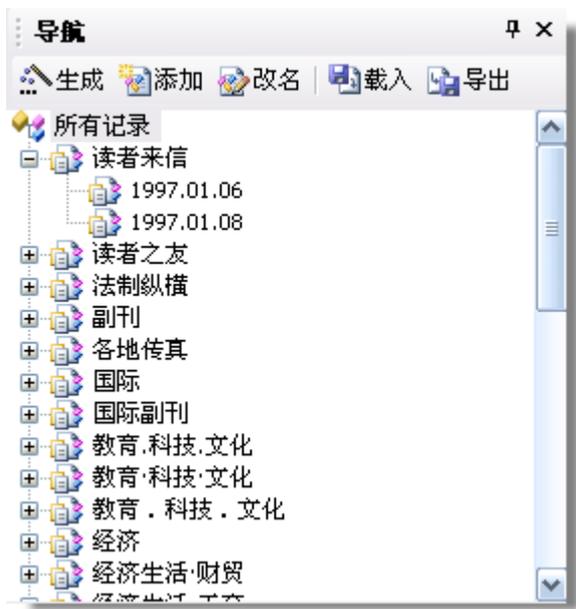
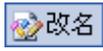


图 2-29 记录维护窗口对应的导航窗口

通过工具栏各个按钮的点击，实现对导航树的维护，工具栏中的每个按钮的功能如下：

-  **生成** 自动生成导航树
-  **添加** 添加新导航节点
-  **改名** 重命名当前导航节点
-  **载入** 从外部文件里导入导航信息
-  **导出** 把导航信息存储到外部文件里

导航树的直接维护通过快捷菜单来完成，选择相应节点并单击鼠标右键即可调出快捷菜单。在导航窗口里，如当前选择节点为根节点，快捷菜单如下：

- 自动生成 自动生成导航树
- 添加导航 添加新导航节点
- 重命名 重命名当前导航节点
- 粘贴 插入剪贴板内容
- 载入 从外部文件里导入导航信息
- 导出 把导航信息存储到外部文件里
- 查看源文件 打开查看当前用户列表树对应的导航文件查看源文件
- 清除导航 删除当前对象对应的导航文件

属性 修改当前导航节点的属性，若属性窗口已关闭，重新打开属性窗口

如当前选择节点不是根节点，快捷菜单如下：

自动生成	自动生成导航树
添加导航	添加新导航节点
重命名	重命名当前导航节点
剪切	删除当前导航节点并把它放到剪贴板上
复制	复制当前导航节点到剪贴板上
粘贴	插入剪贴板内容
删除	删除当前导航节点
上移	当前导航节点位置上移
下移	当前导航节点位置下移
属性	修改当前导航节点的属性，若属性窗口已关闭，重新打开属性窗口

若在导航窗口的空白处点击右键，快捷菜单如下：

自动生成	自动生成导航树
添加导航	添加新导航节点
刷新	刷新当前窗口

导航窗口支持节点的拖动，可以通过鼠标把一个节点及其所有子节点移动/复制到其它位置。操作方法是：按住所要移动/复制的导航节点，并把它拖动到目标节点下，然后放开鼠标左键，然后根据需要选择以下的快捷菜单命令之一：

移动到子节点	把源节点移动到目标节点下
移动到邻接节点	把源节点移动到目标节点的下一节点
复制到子节点	把源节点复制到目标节点下
复制到邻接节点	把源节点复制到目标节点的下一节点
取消	取消节点的拖放操作

在导航窗口中，还可以用以下快捷键编辑导航树：

Insert	“导航 添加”命令，添加新导航节点
Delete	“导航 删除”命令，删除当前导航节点
Space	“导航 重命名”命令，重命名当前导航节点
Shift+↑	“导航 上移”命令，当前导航节点位置上移
Shift+↓	“导航 下移”命令，当前导航节点位置下移

Shift+Enter “导航|属性”命令，修改当前导航节点的属性，若属性窗口已关闭，重新打开

属性窗口

2.9.1 建立导航文件

TRS 管理员提供了一个向导用于生成一个导航文件模板，用户可在此模板的基础上进行修改，以生成自己所需的导航文件。选择“文件|新建”命令，在“新建”对话框里双击导航文件图标，TRS 管理员打开如下“创建导航文件”对话框：



图 2—30 创建导航文件对话框

对话框里各控制的说明如下：

控制名称	功能简介
源对象名	创建导航的对象名，可为数据库或视图。

控制名称	功能简介	
窗口标题	导航文件对应的记录维护窗口的标题，如不指定则自动取对象名作为标题。	
概览字段	浏览时以概览方式显示的字段列表，字段名间以“;”分隔。可按编辑框右边的  按钮来选择字段列表。如概览和细览字段都没有指定则浏览时自动取所有DATE、NUMBER、CHAR、PHRASE型字段作为概览字段。	
细览字段	浏览时以细览方式显示的字段列表，字段名间以“;”分隔。可按编辑框右边的  按钮来选择字段列表。如概览和细览字段都没有指定则浏览时自动取所有DOCUMENT和BIT型字段作为细览字段。	
记录范围	浏览时显示的逻辑记录范围，可用“起始记录号-终止记录号”的方式表示一定范围的记录，或仅用单个数字表示一条记录，多个分段之间以“;”分隔，各分段不能重叠。可按编辑框右边的  按钮来设置记录范围。如不指定则显示检索结果集里的所有记录。	
导航列表	作为导航的导航标签及对应的检索条件。按  按钮弹出“自动生成导航”对话框，帮助用户生成导航列表，此种情况下只能生成一级导航；按  按钮删除当前导航标签。	
显示方式	缺省方式	根据字段的类型决定显示的方式，为缺省设置。
	只有概览	只显示窗口的概览部分，隐藏细览部分。
	只有细览	只显示窗口的细览部分，隐藏概览部分。
	上下分隔	在浏览窗口的上部显示概览，窗口下部显示细览。
	表单编辑	每一个字段对应一个编辑框，可直接在编辑框里修改字段值。
显示导航	选中表示显示导航窗口，否则为隐藏。缺省为显示。	
显示属性	选中表示显示属性窗口，否则为隐藏。缺省为显示。	

设定好各项的内容以后，按[确定]，TRS 管理员将生成导航文件模板并打开文件编辑窗口。按[取消]则放弃生成模板。如上例生成的导航文件如下：

```
' SYSTEM.Demo1
:WindowTitle= 中国专利
:SummarySet= 申请人;公告号;分类号;申请号;申请日;颁证日
:DetailSet= 地址;专利代理机构;摘要
:RecBounds= 1-1000,1005
:DisplayMode= 3
:ShowNavigate= TRUE
0+ SYSTEM.Demo1
  1- 设计人: 设计人=
    2 申请人: 申请人=
```

其中，以“'”开头的为注释行（如第一行），以“:”开头的为参数定义行，最后一个参数定义行以下为导航树定义部分。导航树定义的每一行均以数字开头，数字表示该行所处的层次（0表示根节点），其后跟节点状态表示符（“+”表示打开的文件夹，“-”表示关闭的文件夹，没有表示为页节点），状态表示符以后到“:”前为节点标题，“:”之后为检索条件。检索条件之后还可以带以#字符开头的若干参数（如下表），参数大小写无关，且

彼此顺序无关，但每个参数在一个节点里只能出现一次。

参数名	参数含义
Sort	导航项对应的结果记录的排序方式。多个排序字段间以半角分号、逗号或空格分隔，字段名前加前缀 '+' 表示按升序排列，加前缀 '-' 表示按降序排列，指定 "RELEVANCE" 表示按相关性排序，或用 LIFO 表示按记录的物理记录号的降序排列，如 "+日期; -作者; RELEVANCE"。为空表示不排序。
Stat	导航项设置的统计表达式。
Default	检索条件表达式中缺省字段的子检索表达式的目标字段名集列表，字段名之间以半角分号分隔。为空表示在数据库的数据词典中指定的字段上进行检索；为 '*' 表示在所有 DOCUMENT 字段上进行检索。
Options	检索选项掩码。
Thesaurus	利用知识辅助词典自动进行知识扩展检索时的主题词典。
Synonym	利用知识辅助词典自动进行知识扩展检索时的同义词典。
Antonym	利用知识辅助词典自动进行知识扩展检索时的反义词典。

用户可以根据自己的需要定义相应的导航树，然后保存到系统指定的导航目录里，也可以在记录维护窗口里联机编辑导航树。

2.9.2 查看源文件

打开查看当前导航树对应的导航文件。

导航文件中存储了导航树中各导航节点的属性，用户可以查看或根据自己的需要修改导航树的各属性。

2.9.3 清除导航

删除当前对象对应的导航文件。

因导航文件记录了当前对象对应的导航节点、显示标题、编辑框位置及大小等信息，点击此命令后，对应的导航节点将被删除，上述各值也将会归为默认值。

2.9.4 指定显示的概览字段和细览字段

当某个数据库/视图字段太多时，用户可根据需要指定显示的概览字段和细览字段，所作的修改将应用于概览和细览窗口中。

选择“检索|显示字段”命令，将出现以下对话框：



图 2—31 显示字段对话框

对话框中各控制的说明如下：

控制名称	功能简介
字段列表	未选择的字段列表，该列表里的字段将不在记录维护窗口里显示。按上边的  按钮可把当前选中的字段加入概览字段列表中，按下边的  按钮可把当前选中的字段加入细览字段列表中。
概览字段	显示的概览字段列表，该列表里的字段将按指定顺序显示在概览窗口中。按上边的  按钮把当前选中字段清除出概览字段列表，用概览字段列表框右边的  和  按钮可调整概览字段的顺序。
细览字段	显示的细览字段列表，该列表里的字段将按指定顺序显示在细览窗口中。按下边的  按钮把当前选中字段清除出细览字段列表，用细览字段列表框右边的  和  按钮可调整细览字段的顺序。
全部清除	清除所有概览/细览字段，重新进行设置。
自动重置	自动将全文型(DOCUMENT)字段和二进制型(BIT)字段设为细览字段，其他类型的字段设为概览字段。

设置好概览/细览字段列表以后，按[确定]将改变概览/细览窗口的显示字段，并把所有设置保存在导航文件的修改记载里，保存修改结果时会把这些设置写入导航文件，这样下一次打开此对象时将自动恢复上一次的设置。按[取消]放弃所有设置并关闭对话框。

2.9.5 指定显示的记录范围

当检索结果记录数太多时，可设置只显示一定范围的记录，以方便查看，所作的修改将应用于概览和细览窗口中。

选择“检索|记录范围”命令，将出现以下对话框：



图 2—32 记录范围对话框

对话框中各控制的说明如下：

控制名称	功能简介
记录范围	显示记录的范围。列表里的每一项为一个分段，每个分段可以是单个记录的逻辑记录号，也可以用“起始记录号-终止记录号”的形式表示一定范围内的记录，多个分段之间不能重叠。可用  按钮添加新的分段，也可用  按钮删除当前选中的分段，或者用  和  按钮调整记录分段的顺序。
全部清除	清除所有记录范围的列表，重新进行设置。

设置好记录范围以后，按[确定]将改变当前显示记录的范围，在此范围之外的记录将不会被显示出来。记录范围的设置将保存在导航文件的修改记载里，保存修改结果时会把这些设置写入导航文件，这样下一次打开此对象时将自动恢复上一次的设置。按[取消]放弃所有设置并关闭对话框。

2.9.6 改变记录维护窗口的标题

选择“检索|窗口标题”命令后，将弹出对话框要求用户记录维护窗口的标题。输入完毕后按[确定]，当前浏览窗口的标题将被变为所输入的新字符串，并把标题保存在导航文件的修改记载里，保存修改结果时会把这些设置写入导航文件，这样下一次打开此窗口时将自动恢复上一次设置的标题。

2.9.7 设置控制选项

设置读取结果和记录维护的控制选项。选择“检索|控制选项”命令，将看到如下对话框：



图 2—33 控制选项对话框

对话框中各控制的说明如下：

控制名称		功能简介
读取记录	多值分隔	数据记录文件中多值字段DATE、NUMBER、CHAR、BIT的各个字段值之间的分隔符，缺省为单字节“;”。
	自动折行	只对DOCUMENT和PHRASE字段有效，其中“所有文本不分行”选项一般用于细览字段以便于控制其显示格式，“段落文本不分行”一般用于细览字段以便于自动换行。
	提取格式化文本	只对DOCUMENT字段有效，表示只存储了单个格式化文档字段值的字段，尽量提取文本并给出命中点。
	输出原始的相关度值	选中表示对URL(统一资源定位)输出其对应的文件(如HTML等)，而不是输出URL本身。缺省值为清除。
	对URL输出对应文件	输出记录的原始（绝对）相关度值，而不是缺省方式的相对相关度值。
控制文件		设置用于插入/更新记录的装库控制文件。按 按钮可选择装库控制文件，按 按钮可以对话框的方式设置具体的装库参数。

2.9.8 建立和维护导航树

在记录维护窗口左边为导航窗口，用户可通过导航树方便、直观地浏览信息。

添加新导航项 在导航窗口的快捷菜单里选择“添加”命令，或按 Insert 键，在导航树中将增加一个新节点，节点名默认为“新节点 n”，可按 改名 按钮进行改名，下方属性窗

口会显示相应的各项配置，可在其中对新节点的各项属性进行设置，设置完后点击  更新按钮即可保存。



图 2—34 导航节点的属性

如当前导航节点为组节点，则新节点将添加到当前组节点的最后；如当前导航节点为叶节点，则新节点将添加到当前节点的下面。导航节点的属性将被保存到文件里，并对下一次打开生效。

对话框中各控制的说明如下：

控制名称	功能简介
导航书签	导航项显示的标题。长度在1-31个字符之间，不能包含半角冒号。
统计条件	导航项设置的统计表达式。
检索条件	导航项设置的检索表达式。点击...按钮打开“表达式生成器”对话框，帮助用户生成检索表达式。
排序方式	导航项对应的结果记录的排序方式。点击...按钮打开“选择排序方式”对话框，帮助用户生成排序方式表达式。
缺省字段及选项	打开“检索选项”对话框，设置导航项的检索选项，包括自动缺省目标字段、自动扩展检索等。
不存命中点	选中表示在检索历史中不保存命中点信息，这样可以加快读取和显示结果的速度。
按字检索	选中表示表意文字按字进行检索。

自动生成导航 根据某一字段自动生成导航项。在导航窗口的快捷菜单里选择“自动生成”，或者点击工具栏按钮，TRS 管理员将会弹出如下对话框：



图 2—35 自动生成导航对话框

对话框中各控制的说明如下：

控制名称	功能简介
查询条件	查询表达式。可为“目标词”或“函数名(目标词, ...)”形式，其中“目标词”既可能是一个确定的词汇或字段值，也可是一个符合TRS规范的模糊词汇或模糊字段值。按右边的  按钮进行索引词查询，并把结果显示在结果列表里。若此操作费时较长，可在过程中按Ctrl+Break中断操作。
显示个数	显示结果的个数，缺省为“选项 集成环境”里设置的每页显示词语数。
结果列表	查询结果的词语列表，每个索引词显示其词语、命中数、词频数和它相对于主题词的深度。该结果集里的所有词语将作为自动生成的导航树的节点。按右上角的  和  按钮可调整当前选择词语位置；按  按钮为当前词语改名，即当前词语作为导航节点时以新的名字表示；按  按钮把当前词语从结果集里删除，即不把它作为导航节点。点击“词语”和“命中数”标题可进行

控制名称	功能简介
	排序，三角形尖角向上表示升序，尖角向下表示降序。
查询对象	待查询的目标对象名。改变查询对象将重新检索索引词并刷新检索列表。
查询字段	待查询的字段名，不包括DOCUMENT型、BIT型和PHRASE型字段。改变查询字段将重新检索索引词并刷新检索列表。
继承父节点	选中表示生成的所有子节点自动继承当前根节点的检索条件。
二级导航	进行再次检索时依据的字段。再次检索即在当前根节点检索条件得到的检索结果中根据二级导航节点再次进行检索。只有当“继承父节点”选中时此控制项才能进行选择。
分类条件	二次检索时只需关注的字段值，比如上图所示，即，二次检索时只生成检索条件为“版名=要闻”的导航节点。输入时每个值为一行（注意不是用其它符号作为分隔符）。如果为空，则缺省为依据字段字段的所有字段值。
正则表达式	选中表示通过正则表达式从所有字段值中选取符合条件的字段值作为目标类别，并容许抽取某些匹配项进行组合而形成新的类别。正则表达式的具体规则见《TRSDatabase.doc》4.2节。
区分大小写	选中表示检索的字段值区分大小写。

把查询结果调整到符合自己的要求以后，按[确定]将关闭对话框，并在导航窗口的当前节点下自动生成导航树（每个词语对应一个导航节点）。按[取消]将关闭对话框而不做任何操作。

添加导航节点 在导航窗口的快捷菜单里选择“添加”命令，或按 Insert 键，在导航树中将增加一个新节点，节点名默认为“新节点 n”，可按按钮进行改名，下方属性窗口会显示相应的各项配置，可在其中对新节点的各个属性进行设置，设置完以后点击按钮即可保存。如当前导航节点为组节点，则新节点将添加到当前组节点的最后；如当前导航节点为叶节点，则新节点将添加到当前节点的下面。导航节点的属性将被保存到文件里，并对下一次打开生效。

删除当前导航节点 在导航窗口的快捷菜单里选择“删除命令”，或按 Delete 键，TRS 管理员将会提示是否删除当前导航项及其所有子导航项。按[是]表示删除当前导航项及其所有子导航项（根节点不能删除），按[否]表示只删除当前导航项的所有子导航项，按[取消]则放弃删除操作。

当前导航节点位置上移 在导航窗口的快捷菜单里选择“上移”命令，或按 Shift+ ↑ 键，TRS 管理员将把当前节点上移一个位置（对根节点无效）。

当前导航节点位置下移 在导航窗口的快捷菜单里选择“下移”命令，或按 Shift+ ↓ 键，TRS 管理员将把当前节点下移一个位置（对根节点无效）。

重命名当前导航节点 在导航窗口的快捷菜单里选择“重命名”命令，或按空格键，TRS 管理员将在当前节点的位置显示一个编辑框。输入新导航书签以后按回车，当前导航节点的书签将被修改为新值。另外，也可在下方属性窗口中直接对当前导航节点的名称进行

修改，修改后点击  按钮保存即可。

查看源文件 打开查看当前导航树对应的导航文件。导航文件中存储了导航树中各导航节点的属性，用户可以查看或根据自己的需要修改导航树的各属性。

修改当前导航节点的属性 焦点定位于下方的属性窗口上，以便于对当前节点的属性进行修改，若属性窗口已关闭，重新打开属性窗。修改导航节点的书签以及各项配置以后，按  按钮，TRS 管理员将更新原导航节点的各项属性。

如果当前所选节点为根节点时，还可使用以下两个命令来载入/导出导航文件：

从外部文件里导入导航信息 在导航树根节点的快捷菜单里选择“载入”命令，或点击工具栏上的  按钮，TRS 管理员将弹出“选择导航文件”对话框提示用户选择要载入的导航文件名，选择完毕后按“确定”，TRS 管理员将读入该文件，并刷新当前浏览窗口的导航信息和导航树。

把导航信息存储到外部文件里 在导航树根节点的快捷菜单里选择“导出”命令，或者点击工具栏上的  按钮，TRS 管理员将弹出“选择导航文件”对话框提示用户选择要导出的导航文件名，输入完毕后按“确定”，TRS 管理员将把当前的导航信息和导航树写入到指定的文件里。

2.10 记录维护

对记录的维护直接在检索的结果集上进行，但当前检索对象只能是单个数据库或自动模式视图，不能是普通视图或多个数据库，而且只有对该数据库/自动模式视图具有更新权限的用户才可以进行添加记录、删除记录、更新记录和修改记录安全设置操作。

2.10.1 添加记录

选择“记录|添加记录”命令，或按下状态栏上的  按钮后，浏览窗口将会切换到表单编辑状态，在最后添加一条空的记录，并把当前记录指示符移动到这条空记录上。设定各字段值后按“保存修改”按钮保存所作的更新并返回原先的浏览状态，按“放弃修改”则取消所作的修改并返回原先的浏览状态。可一次添加多条记录，选择“文件|保存”命令，或按工具栏上的  按钮，新记录将被添加到数据库中。

2.10.2 批量删除记录

如要快速删除当前选择的记录而不刷新检索结果，可按状态栏上的  按钮，这些记录会在保存修改结果的时候从数据库里删除。否则，选择“记录|批量删除记录”命令，将看到如下对话框：



图 2—36 批量删除记录对话框

对话框中各控制的说明如下：

控制名称	功能简介	
记录范围	当前记录	只删除当前记录，为缺省设置。
	所有记录	删除检索结果里的所有记录。
	指定范围	删除指定范围的记录，可用“起始记录号-终止记录号”的方式表示一定范围的记录，或仅用单个数字表示一条记录，多个分段之间以“,”分隔，各分段不能重叠。
	指定条件	删除指定条件的记录，该条件和当前检索条件的与就是待删除的记录的范 围。
快速逻辑删除	选中表示只在所删记录上作个标记，而并不真正从数据库中删除它，这些删除的记录只有在优化数据库的时候才会被真正地清除。快速逻辑删除可加快删除的操作速度，为缺省设置。	

设置好删除范围后，按[确定]，TRS 管理员将开始删除指定范围的记录，删除完毕后刷新检索结果。删除过程中会弹出对话框显示删除的进度。若此操作费时较长，可在过程中按 Ctrl+Break 中断操作，此时删除取消。按[取消]则放弃删除操作并关闭对话框。如要快速删除当前选择的记录而不刷新检索结果，可按状态栏上的  按钮，这些记录会在保存修改结果的时候从数据库里删除。

2.10.3 批量更新记录

更新单条记录可以直接在表单编辑窗口中进行，概览/细览模式下按状态栏上的按钮，会进入表单编辑模式，可对记录进行修改后按“保存修改”按钮，将保存更新并返回原显示状态。也可一次编辑多条记录，选择“文件|保存”命令，或按工具栏上的按钮，所作修改将被保存到数据库/自动模式视图中。批量更新记录可选择“记录|批量更新记录”命令，将看到如下对话框：



图 2—37 指定字段值方式批量更新记录对话框

选择“指定文件”方式，则变为下图形式：

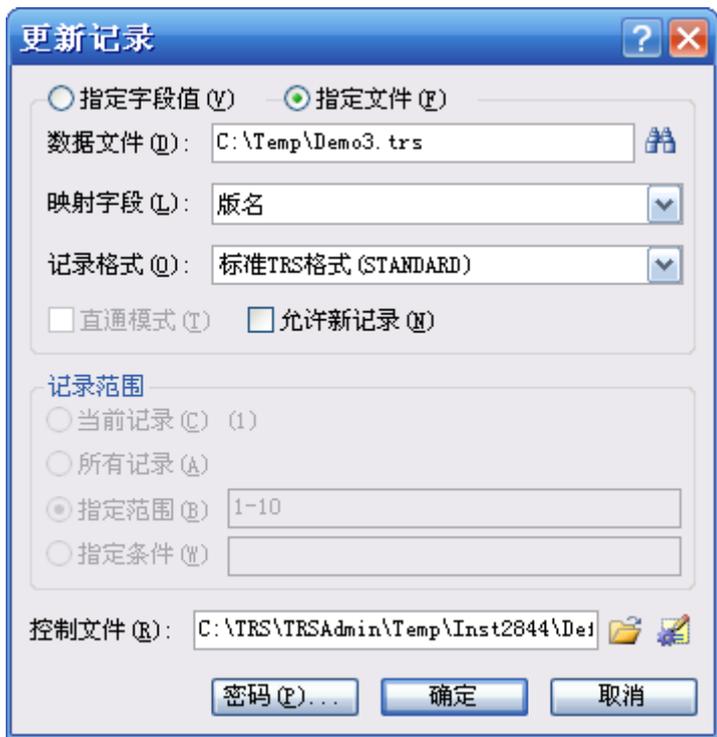


图 2—38 指定文件方式批量更新记录对话框

对话框中各控制的说明如下：

控制名称	功能简介	
更新字段值	更新记录的字段及取值，每一行表示一个需要更新的字段及对应的取值。 按  按钮会弹出“添加字段值”对象框，选择需要更新的字段及对应的字段值，更新的字段名不能重复，即已加入列表框的字段不能再被选择；也可以按  按钮修改当前选中的字段值；或按  按钮删除当前选中的字段值。	
记录范围	当前记录	只更新当前记录，为缺省设置。
	所有记录	更新检索结果里的所有记录。
	指定范围	更新指定范围的记录，可用“起始记录号-终止记录号”的方式表示一定范围的记录，或仅用单个数字表示一条记录，多个分段之间以“,”分隔，各分段不能重叠。
	指定条件	更新指定条件的记录，该条件和当前检索条件的与就是待更新的记录的范围。
指定文件方	数据文件	即用来批量修改记录的记录格式文件。按按钮可选择数据文件的路径。
	映射字段	指定要批量修改的字段名，或者DOCID。比如选择字段“标题”，在数据文件中又有若干<标题>=value的记录，那么这些记录中的数据即会用来修改原数据库中“标题”为value的记录。
	记录格式	数据记录文件的格式，缺省为标准TRS格式。

式	直通模式	选中表示以“直通模式”加载数据。如果数据加载所涉及所有原始数据文件、批处理文件、控制文件，服务器的启动帐号都有读取权限，则可使用直通模式进行数据加载。由于直通模式下服务器直接读取原始数据文件，从而减少了中间环节，所以效率会更高。
	允许新记录	选中表示允许源记录格式文件中出现的新记录，即记录格式文件中的某些记录在原数据库中不存在。
控制文件		设置用于更新记录的装库控制文件。按  按钮可选择装库控制文件，按  按钮可以对话框的方式设置具体的装库参数。
密码		打开“记录密码”对话框，设置记录的密码、安全级别和公开用户组名等安全设置。

输入更新表达式，设置好更新范围和记录密码后，按[确定]，TRS 管理员将开始更新指定范围的记录，更新完毕后刷新检索结果。更新过程中会弹出对话框显示更新的进度。若此操作费时较长，可在过程中按 **Ctrl+Break** 中断操作，后面还未更新的记录将不再更新，前面已更新的记录保持更新。按[取消]则放弃更新操作并关闭对话框。概览/细览模式下按状态栏上的  按钮，会进入表单编辑模式，可对记录进行修改后按“保存修改”按钮，将保存更新并返回原显示状态。

2.10.4 记录密码

每一条记录都可以设置密码、安全级别和开放组名，只有具备足够权限的用户才能浏览该记录。选择“记录|当前记录密码”命令，或在“批量更新记录”对话框里按下“密码”按钮后，将弹出对话框：



图 2—39 记录密码对话框

对话框中各控制的说明如下：

控制名称	功能简介
设置密码	设置记录的访问密码。为保密起见，所输入的字符以“*”显示。密码有效长度是1-8个字符。此控制项初始化为8个“*”，在第一次设置该记录密码时，表示

	密码为空：在修改记录的密码时，表示保持原密码。若此控制项为空，也表示没有密码。当所输入密码首尾出现空格时，Admin会自动将其去除，但密码中间可以包含空格。
确认密码	确认记录的访问密码，必须与设置密码一致。为保密起见，所输入的字符以“*”显示。
安全级别	赋予记录的安全级别，访问这些记录时操作者的安全级别必须不小于此安全级别。为零表示访问此记录无须经过安全级别检查。缺省值为空，表示不进行设置。
开放组名	记录对外开放的用户组，该用户组的所有用户无须经过权限检查就可访问这些记录。为“(不作限制)”表示不限制开放组，即对所有组开放。缺省为空串，表示不修改上次的设置。

输入记录的新密码、安全级别，选择好开放组名后，按[确定]，这些设置将被保存起来。在保存修改或更新记录的时候，这些设置将被更新到数据库/自动模式视图里。以后需要读取记录时，如设置了记录的密码，TRS 管理员将弹出对话框要求输入该记录的正确密码，如所输密码不正确，将无法浏览该记录。

2.11 维护 TRS 数据库对象

2.11.1 数据加载

装入新的数据到数据库/词典中。这里只讲述加载数据的方法，有关数据加载格式规范请参阅附录 B。

- 创建装库数据文件模板

可用 TRS 管理员的“创建装库数据文件”对话框生成一个装库数据文件的模板。“文件|新建”命令，在“新建”对话框里双击数据文件图标，TRS 管理员打开如下“创建数据文件”对话框：

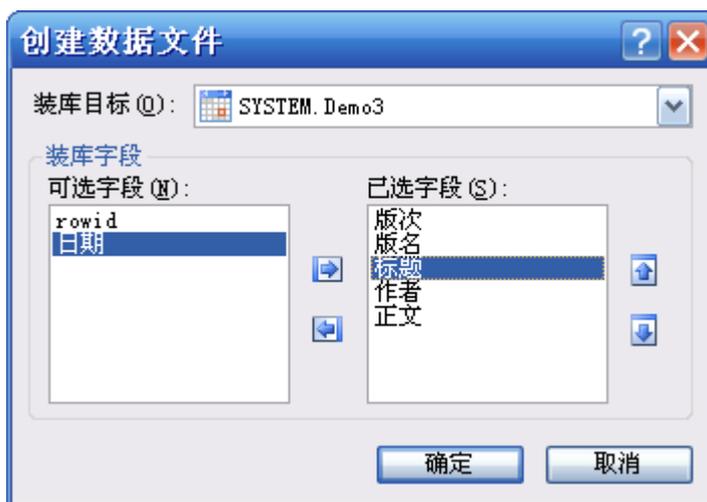


图 2—40 创建数据文件对话框

对话框里各控制的说明如下：

控制名称	功能简介
装库目标	装库文件对应的目标数据库/自动模式视图名。
装库字段	选定作为装库字段的字段列表。
可选字段	可以作为装库字段的字段列表，初始为当前对象中的所有字段。可用  按钮加入已选字段列表。
已选字段	已选择作为装库字段的字段列表，可用  按钮把当前选中字段清除出已选字段列表，或用  和  按钮调整装库字段的顺序。

设定好各项的内容以后，按[确定]，TRS 管理员将生成装库数据文件模板并打开文件编辑窗口。按[取消]则放弃生成模板。如上例生成的装库数据文件如下：

```
<REC>
<申请日>=
<颁证日>=
<申请人>=
<地址>=
<设计人>=
<专利代理机构>=
<代理人>=
<摘要>=
```

用户可以根据需要对此模板进行编辑（填入相应的字段值以及增加记录等），然后把它保存到文件里。

- 创建批装库文件模板

同样，可用 TRS 管理员的“创建批装库文件”对话框生成一个批装库文件的模板。“文

件|新建”命令，在“新建”对话框里双击批装库文件图标，TRS 管理员打开如下“创建批装库文件”对话框：



图 2—41

对话框里各控制的说明如下：

控制名称	功能简介
数据文件	待生成批装库文件的文件列表。按 按钮可添加装库数据文件或批装库文件，按 按钮可添加某个目录下的所有装库文件，按 按钮可删除选定的文件/目录，用 和 按钮可调整文件/目录的顺序。单击列表中的某一项可修改该文件/目录名。

设定好各项的内容以后，按[确定]，TRS 管理员将生成批装库文件模板并打开文件编辑窗口。按[取消]则放弃生成模板。如上例生成的装库数据文件如下：

```
D:\Programs\LoadFile\Patent.tris
D:\Programs\LoadFile\BJWB.tris
D:\Programs\Demo1.tris
D:\Programs\Demo2.tris
D:\Programs\Demo3.tris
D:\RMRB\Rmrb.tris
```

用户可以根据需要对此模板进行编辑（添加、删除、修改装库数据文件），然后把它保存到文件里。

- 把数据加载到数据库/词典里

准备好装库数据文件以后，就可把数据装入到数据库中。选择“工具|装入记录”命令后，TRS 管理员弹出如下对话框：

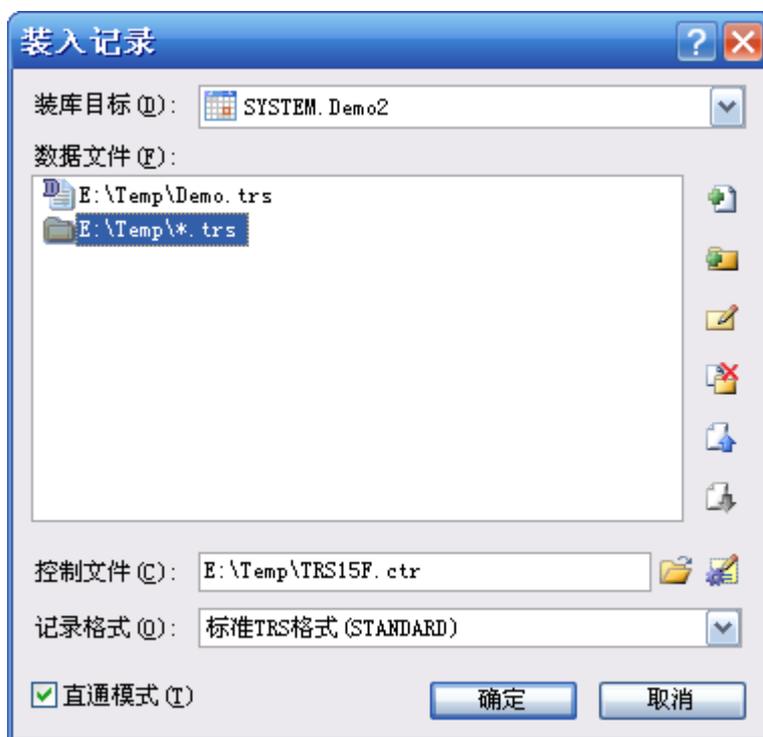


图 2—42 装入记录对话框

对话框中各控制的说明如下：

控制名称	功能简介
装库目标	待装入记录的数据库名。当前用户必须对该数据库有更新权限。
数据文件	待装入的文件列表，可一次装入多个数据文件或批装库文件，或指定一个目录下的所有装库文件。按 按钮可添加装库数据文件或批装库文件，按 按钮可添加某个目录下的所有装库文件，按 按钮对选定的文件/目录进行编辑修改，按 按钮可删除选定的文件/目录，用 和 按钮可调整文件/目录的顺序。单击列表中的某一项可修改该文件/目录名。
控制文件	装库控制文件，用于设置装库参数，为空则取系统缺省值。按 按钮可选择装库控制文件，按 按钮可以对话框的方式设置具体的装库参数。
记录格式	数据记录文件的格式，缺省为标准TRS格式。
直通模式	选中表示以“直通模式”加载数据。如果数据加载所涉及所有原始数据文件、批处理文件、控制文件，服务器的启动帐号都有读取权限，则可使用直通模式进行数据加载。由于直通模式下服务器直接读取原始数据文件，从而减少了中间环节，所以效率会更高。

设定好装库目标、数据文件和控制文件后，按[确定]按钮则开始装入记录，按[取消]

按钮则关闭对话框。由于装库是一个耗时的过程，因此，装库过程中有“装库进度”对话框显示当前正在装入的文件、已装记录数和本批加载完成的百分数。如想终止装库过程，可按下对话框里的[终止]按钮。装库结束后，TRS 管理员会弹出“结果报告”对话框，显示装库的起始时间、终止时间、花费时间、成功记录数、失败记录数、致命错误等结果信息。当装入有失败记录时，如要查看错误报告文件，可选中[打开报告查看]复选框，然后按[生成报告]，即可打开查看错误报告，也可设置为只生成报告而不打开报告查看，节约了时间。

- 设置装库控制参数

在“装入记录”对话框里按按钮，将弹出“设置控制参数”对话框，如下图所示：

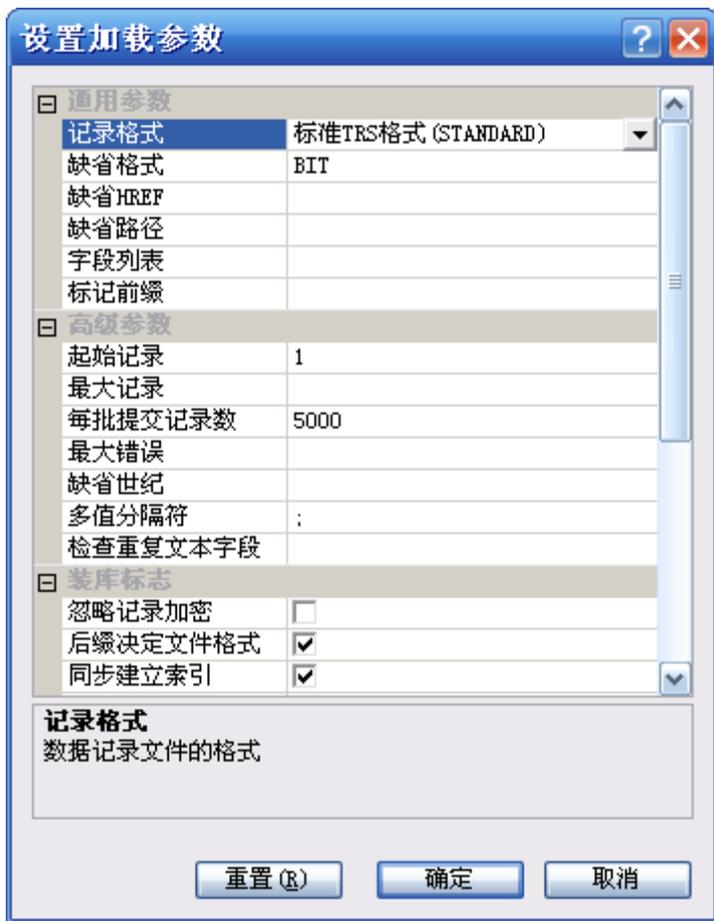


图 2—43 设置加载参数对话框

对话框里各控制的说明如下：

控制名称		功能简介
通用参数	记录格式	数据记录文件的格式，缺省为标准TRS格式。
	缺省格式	数据记录文件中BIT字段值文件的缺省数据格式名，缺省为BIT。
	缺省HREF	指明DOCUMENT字段引用的HTML格式的字段值文件中BASE元素的缺省HREF属性值，只对DOCUMENT字段的HTML字段值有效，显式BASE控制符优先于该变量。此变量没有缺省值。
	缺省路径	数据记录文件中字段值所引用的数据文件的缺省路径，缺省为当前数据文件所在目录。
	字段列表	需要加载数据的字段名列表，缺省为全部字段。可按编辑框右边的...按钮来选择字段列表。
	标记前缀	TRS标记引导符前缀，缺省为空。
高级参数	起始记录	在数据记录文件中数据加载的起始记录号，缺省为1。
	最大记录	需要加载的最大记录数，缺省为全部记录。
	每批提交记录数	加载时分批提交的最大记录数，缺省为5000。
	最大错误	加载时允许错误记录的最大个数，缺省为没有限制。
	多值分隔	数据记录文件中多值字段DATE、NUMBER、CHAR、BIT的各个字段值之间的分隔符，缺省为单字节“;”。
	缺省世纪数	数据记录文件中DATE字段值的年份的缺省世纪数，缺省为当前世纪。
检查重复文本字段	是否允许在数据记录文件中同一个PHRASE或DOCUMENT字段在同一个记录中重复出现的开关。如果数据记录文件的格式是ISO2709或XML，则缺省值为JOIN1，否则为FALSE。	
装库标志	忽略记录加密	是否忽略数据记录文件中的记录加密描述符的开关，缺省为清除。
	同步建立索引	在加载数据的同时建立索引的开关，缺省为设置。
	后缀决定文件格式	当数据记录文件中所引用的字段值文件未指定数据格式控制符时，用文件名的后缀来确定其数据格式的开关，缺省为设置。
	保持文本数据空白	是否保持PHRASE和DOCUMENT字段值文本中首尾空白行的开关，缺省为设置。
	分类存储BIT字段	是否分类存储BIT字段值的开关，缺省为设置。
	分类存储DOCUMENT字段	是否分类存储DOCUMENT字段值的开关，缺省为设置。
	不存储Office文档的文本	是否丢弃由OEM过滤器所生成的文本数据的开关，缺省为设置。
	不存储字段值附加文件	在数据记录文件中，DOCUMENT字段值的附加文件是否存储在数据库中的开关，缺省为清除。
	允许自动匹配入库字段	是否允许自动匹配入库字段，缺省为清除。
以ISO2709映射入库字段	是否以ISO2709格式的目录标记（以及数据标识符）映射入库字段，缺省为清除。	
XML控制参数	分段标记	XML数据作为一个字段值入库时，过滤器进行自动分段的标记。
	分句标记	XML数据作为一个字段值入库时，过滤器进行自动分句的标记。此功能为TRS全文数据库6.50.4000版本新增。
	换行标记	XML数据作为一个字段值入库时，过滤器进行自动换行的标记。
	命中点标记	XML数据作为一个字段值入库后，检索命中点的指示标记。
	CDATA标记	XML数据作为一个字段值入库时，过滤器是否处理标记中的内容，缺省值为FALSE。
	覆盖XML数据中已有的过滤器处理指令	XML数据作为一个字段值入库时，是否需要覆盖数据中已有的关于过滤器的处理指令，缺省为清除。

以上控制变量并不要求全部设置，只须修改需要设置的参数，其它项会自动取缺省值，按[重置]也可把所有选项设回缺省值。设定好各项的内容以后，按[确定]，TRS 管理员将把当前所设选项保存在缺省的装库控制文件里，并作为当前装库所使用的控制文件。按[取消]则放弃装库选项的设置。

● 生成装库控制文件

如果要重复使用某些设置，可以自己创建一个控制文件。选择“文件|新建”菜单，然后选择“控制文件”类型，TRS 管理员将弹出一个与设置控制参数一样的对话框。设定好各个参数以后，按[确定]，TRS 管理员将新建一个控制文件编辑窗口。把该文件保存起来以后，需要再装库时就可使用该文件作为装库控制文件了。如上例，生成的控制文件如下：

```
RECORD_FILE_FORMAT= STANDARD
BIT_FORMAT_DEFAULT= BIT
HTML_BASE_DEFAULT=
DATA_PATH_DEFAULT=
SECTION_TAG_PREFIX=
DATE_CENTURY_DEFAULT=
MULTI_VALUE_SEPARATOR= ;
START_RECORD_FROM= 1
PROCESS_RECORD_NUMBER= 10000
MAX_COMMIT_NUMBER= 5000
MAX_ERROR_NUMBER= 100
IGNORE_RECORD_CRYPT= FALSE
FILE_SUFFIX_PRIOR= TRUE
SYNC_CREATE_INDEX= TRUE
CLASS_BIT_STORE= TRUE
CLASS_DOCUMENT_STORE= TRUE
KEEP_TEXT_ACTUAL= TRUE
CHECK_TEXT_ITERATIVE=
WAIVE_OEMFILTER_TEXT= TRUE
WAIVE_DOCUMENT_REFERER= FALSE
COLUMN_AUTO_MATCH= FALSE
NAMIZE_ISO2709_FIELD= FALSE
COLUMN_NAME_SEQUENCE= 日期;版次;版名;栏目;标题;作者;正文;图像;文件名;标记
XML_SEGMENT_MARK=
XML_NEWLINE_MARK=
XML_HITSHOW_MARK=
XML_FILTER_CDATA= false
XML_OVERWRITE_PI= FALSE
```

2.11.2 结构导入/导出

通过管理员可把对象导出到文件中，或从文件里恢复对象(当前用户必须具备 Resource 权限)。使用此功能，可以在重新安装 TRS 后恢复原有的对象(包括对象信息和结构，但不包括数据库的记录)，或把对象从一个服务器复制到另一个服务器(两台服务器的平台可以不一样)。

在对象管理窗口里选择要导出的对象，然后选择“文件|导入/导出”命令，TRS 管理员弹出如下对话框：

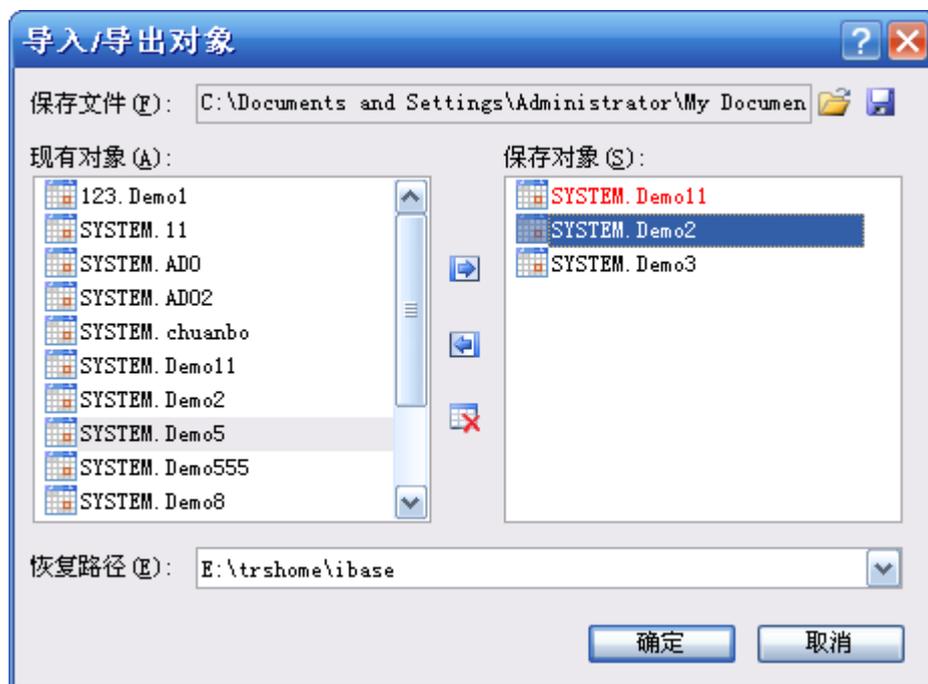


图 2—44 导入导出对话框

对话框中各控制的说明如下：

控制名称	功能简介
保存文件	导入对象的源文件名或对象导出的目标文件名，可按右边的 按钮选择导入对象的源文件名，或按 按钮选择导出对象的目标文件名。
现有对象	当前服务器上的对象列表，可按右边的 按钮把所选的对象添加到保存对象列表框里。
保存对象	当前指定文件里包含的对象列表（以红色字体显示以示与服务器上现有对象的区别），可按左边的 按钮把所选的对象添加到现有对象列表框里（在导入文件里仍保留该对象），或按 按钮把所选对象从导入文件里删除。
恢复路径	导入数据库对象时，指定恢复的目标路径。必须是系统配置中的数据和索引路径之一，缺省为系统配置中的第一个数据路径。可以在下拉列表里选择。

设定好要导入/导出的对象，以及保存文件和恢复路径后，按[确定]按钮，TRS 管理员会先从指定文件中把相应对象恢复到当前服务器中，然后把加入到保存对象列表框里的对象的属性/结构信息存储在指定文件里（除被明确删除的对象以外，保存文件中原有的对象信息继续保留）。按[取消]按钮则关闭对话框。

此功能不仅对数据库对象有效，也同样适用于视图、词典、用户组、用户和数据格式，

对三者只有 SYSTEM 用户能恢复对象。以下章节不再重复。

注意：只有当导入源文件中包含的对象类型与当前“现有对象”列表框中的对象类型一致时，源文件中包含的对象列表才会显示于“保存对象”列表框中。

2.11.3 数据备份

通过管理员可把指定的数据库/自动模式视图备份到磁带或文件中，每次操作可以备份多个数据库（自动模式视图则只能选一个）。操作用户必须为 DBA 用户、拥有者所在组的 GDA 用户或对象的拥有者。（注：备份操作不备份权限信息。）

选择“工具|备份数据”命令后，TRS 管理员弹出如下对话框：



图 2—45 备份数据对话框

对话框中各控制的说明如下：

控制名称	功能简介
备份对象	待备份的数据库/自动模式视图列表，如是数据库可一次选择多个，自动模式视图则只能选一个。
备份设备	备份目标的设备名，缺省值为系统的备份路径。当登录本地服务器时，可点击  按钮选择本地路径，点击  按钮选择备份目录名。

选择[数据库]或[视图]单选项，然后在[备份对象]列表框中选择需要备份的对象，在[备份设备]编辑框中输入备份的设备名或文件名，然后按[确定]，TRS 管理员将开始进行备份操作，按[取消]按钮则关闭对话框。备份数据是一个耗时的过程，过程中有“备份进度”

对话框显示当前正在备份的对象、当前备份的进度和总进度的百分数。如想终止备份过程，可按下对话框里的[终止]按钮。备份数据库结束后，TRS 管理员会弹出“结果报告”对话框，显示备份的起始时间、终止时间、花费时间、备份对象数、致命错误等结果信息。

2.11.4 数据恢复

从指定的设备恢复数据库/自动模式视图，每次操作可以恢复多个数据库（自动模式视图只能是一个）。操作用户必须为 DBA 用户、GDA 用户或 Resource 用户。（注：数据库/自动模式视图恢复后，需要对它进行授权，否则该库只有 DBA 用户、拥有者所在组的 GDA 用户或对象的拥有者才能访问。）

选择“工具\恢复数据”命令后，TRS 管理员弹出如下对话框：



图 2—46 恢复数据

对话框中各控制的说明如下：

控制名称	功能简介
备份设备	备份包的设备名，可以是磁带名或文件名。输入备份设备后，按  按钮可在[数据库] 列表中列出该备份设备里包含的所有数据库或者自动模式视图。当登录本地服务器时，可点击  按钮选择本地路径，点击  按钮选择备份目录名。
恢复对象	待恢复的对象列表。可选择一个或多个对象进行恢复。

恢复路径	恢复的目标路径，必须是系统配置包含的数据库或索引路径，为空表示系统配置中数据路径的第一个。
------	---

在[备份设备]里输入备份设备名，在[数据库]列表框中选择需要恢复的数据库，在[恢复路径]里选择目标路径，然后按[确定]，TRS 管理员将开始进行恢复操作，按[取消]按钮则关闭对话框。由于恢复数据库是一个耗时的过程，因此，恢复过程中有“恢复进度”对话框显示当前正在恢复的数据库、当前恢复数据库的进度和总进度的百分数。如想终止恢复过程，可按下对话框里的[终止]按钮。恢复数据库结束后，TRS 管理员会弹出“结果报告”对话框，显示恢复的起始时间、终止时间、花费时间、恢复数据库数、致命错误等结果信息。

2.11.5 数据库的复制

数据库的复制是指结构的复制，如果要同时复制数据，可先复制结构然后参看 2.9.4 复制记录。数据库的复制分两步。

第一步：把当前所选数据库复制到剪贴板。

先选择需要复制的数据库，再选择“编辑|复制”命令或按下  按钮把所选数据库复制到剪贴板（也可按 Ctrl+C 键）

第二步：粘贴剪贴板里的数据库。

选择“编辑|粘贴”命令或按下  按钮（也可按 Ctrl+V），即可粘贴剪贴板里的数据库，使用此功能不仅可以在当前服务器上创建数据库的副本，还可以在不同服务器之间复制数据库。如果当前服务器上已有同名的数据库，则 TRS 管理员会提示用户输入新的数据库名。

此功能不仅对数据库对象有效，也同样适用于视图、词典、用户组、用户和数据格式，对三者只有 SYSTEM 用户能粘贴对象。以下章节不再重复。

2.11.6 复制记录

把符合指定条件的记录从一个数据库/自动模式视图复制到另一个数据库/自动模式视图中。

选择“工具|复制记录”命令后，TRS 管理员弹出如下对话框：



图 2—47 复制记录对话框

对话框中各控制的说明如下：

控制名称	功能简介
复制目标	复制记录的目标数据库/自动模式视图名。
数据来源	复制记录的源数据库/自动模式视图名。
字段列表	待复制的字段列表。每次改变复制目标或数据来源时，字段列表框里会自动列出两者共同的字段（字段名和字段类型均一致），用户可以选择需要复制的字段（如不选择表示复制全部字段）。点击“字段名”和“类型”标题可进行排序，三角形尖角向上表示升序，尖角向下表示降序。
复制条件	复制记录的条件，数据来源里只有符合此条件的记录才会被复制到复制目标中。点击  按钮弹出“表达式生成器”对话框，帮助用户生成检索表达式。
排序方式	复制记录时所用的排序方式，可直接输入也可在下拉列表里选择。点击  按钮弹出“选择排序方式”对话框，帮助用户生成排序方式表达式。
同步索引	选中表示在复制记录的过程中同时建立索引，否则用户必须在复制完毕以后手动建立索引。缺省为选中。

设定好复制目标、数据来源、字段列表、复制条件和排序方式等参数后，按[确定]按钮则开始复制记录，按[取消]按钮则关闭对话框。由于复制记录是一个耗时的过程，因此，复制过程中有“复制进度”对话框显示当前已复制的记录数和本批复制完成的百分数。如想终止复制过程，可按下对话框里的[终止]按钮。复制记录结束后，TRS 管理员会弹出“结

果报告”对话框，显示复制的起始时间、终止时间、花费时间、复制记录数、致命错误等结果信息。

2.11.7 数据优化

优化选择的数据库/自动模式视图。即对数据文件和索引文件进行优化重组，以达到提高检索速度和提高存储空间利用率的目的，一般在数据装载结束或进行数据库维护操作之后进行。操作用户必须为 DBA 用户、或拥有者所在组的 GDA 用户、所或为所选对象的拥有者，或对指定对象有 Index 权限。

选择“工具|优化数据”命令后，TRS 管理员弹出如下对话框：



图 2—48 优化数据对话框

对话框中各控制的说明如下：

控制名称	功能简介
优化对象	待优化的数据库/自动模式视图名。
字段列表	待优化的数据库字段列表，可选择一个或多个字段，不选表示对所有字段进行优化。点击“字段名”或“类型”标题可进行排序，三角形尖角向上表示升序，尖角向下表示降序。
工作空间	优化时使用的临时目录，为空表示缺省目录。当登录本地服务器时，可点击  按钮选择本地路径。

在[优化对象]组合框中选中要进行优化的数据库/自动模式视图，在[字段列表]列表框

中选择需要优化的字段（不选表示对所有字段索引以及数据文件进行优化；如指定了字段则只优化指定字段的索引），在[工作空间]编辑框里输入优化的临时目录，然后按[确定]开始进行优化，按[取消]按钮则关闭对话框。由于优化数据是一个耗时的过程，因此，优化过程中有“优化进度”对话框显示当前正在优化的字段名及当前优化进度的百分数。如想终止优化过程，可按下对话框里的[终止]按钮。优化数据库结束后，TRS 管理员会弹出“结果报告”对话框，显示优化的起始时间、终止时间、花费时间、优化字段数、致命错误等结果信息。

2.11.8 索引维护

索引维护是指对数据库/自动模式视图的索引进行创建、删除或验证操作。操作用户必须为 DBA 用户、或所有者所在组的 GDA 用户、或对象的拥有者，或对所选对象有 Index 权限。

选择“工具|索引维护”命令后，TRS 管理员弹出如下对话框：



图 2—49 索引维护对话框

对话框中各控制的说明如下：

控制名称	功能简介
维护对象	待进行索引维护的数据库/自动模式视图名。
字段列表	待进行索引维护的字段列表（设置了“不建索引”标志的字段除外），可选择一个或多个字段，点击“字段名”或“类型”标题可进行排序，三角形尖角向上表示升序，尖角向下表示降序。 按  按钮可对所选字段创建索引，对数据库需要输入创建索引的终止记录

	<p>号（为零表示为所有记录创建索引），对自动模式视图只能全部创建。</p> <p> 按钮可删除所选字段上的索引，对数据库需要输入删除索引的起始记录号（为零表示删除所有记录的索引），对自动模式视图只能全部删除。</p> <p> 按钮可验证所选字段上的索引的正确性，如果验证表明某个字段的索引不正确，说明该字段的索引文件可能已被损坏，可以删除该字段的索引，然后重建，以保持索引的正确性。</p> <p>有  标志的字段表示索引不完整，前面标有  标志的字段表示已验证过索引的正确性。</p>
临界记录号	当前选中字段上建立索引的临界记录号，该记录号之前（包含该记录号）的记录已经建立索引，该记录号之后的记录则还没有建立索引。
最大记录号	当前选中字段上已经加载数据的最大记录号。
检查	对当前所有数据库和自动模式视图进行字段索引的不完整性检查，检查完后上方的下拉框列表中只显示字段索引不完整的数据库或自动模式视图，此时 [检查] 按钮变为 [重置] 按钮，点击重置按钮后重新显示所有的数据库和自动模式视图。

创建索引、删除索引和验证索引均为耗时操作。创建/删除索引过程中会显示进度对话框显示操作成功的总记录数和当前正在提交的索引进度的百分数。验证索引过程中会显示“验证进度”对话框，以显示正在验证索引的字段名和验证进度的百分数，验证完成后会弹出“结果报告”对话框，显示验证的起始时间、终止时间、花费时间、验证字段数、致命错误等结果信息。操作过程中按下进度对话框的里的 [终止] 按钮按中断操作。索引维护操作完成后，按 [关闭] 按钮关闭对话框。

2.11.9 数据库解锁

由于某种原因，某些数据库可能被加锁（非共享模式），为这些状态异常的数据库解锁（设为共享模式），使数据库可被访问就是解锁。操作者如果满足下列条件之一，则有权解锁数据库（数据库的异常一般是由数据库操作的意外中断引起）：

1. 超级用户。
2. DBA 用户。
3. 数据库所有者所在用户组的 GDA 用户。
4. 数据库锁的拥有者，并且数据库具有公共 ALTER 属性或者操作者拥有该数据库的 ALTER 权。

但如果一个数据库被“拒绝访问锁”锁定，则任何用户都不能解锁，只能由系统内部自动解除，或者重新启动系统。

选择“工具|解锁数据库”命令后，TRS 管理员弹出如下对话框：



图 2—50 数据库解锁对话框

对话框里的“数据库”列表框里列出了当前已加锁的数据库及其状态。选择需要解锁的数据库，然后按[确定]，TRs 管理员会把所选择的数据库恢复到共享状态，并在状态条显示成功信息。按[取消]则放弃解锁操作并关闭对话框。

2.11.10 自动分库

自动分库是指将物理数据库转化为能自动分裂数据库的视图。可对生成的分裂视图进行装入记录和复制记录操作，结果将自动反映到其对应的物理数据库里。

从对象管理窗口中数据库页的快捷菜单中选择“转成自动模式视图”命令，TRs 管理员弹出如下对话框：



图 2—51 自动分库对话框

对话框里的“数据库”列表框里列出了当前所有可以转化为自动模式视图的数据库。“分裂条件”用来描述数据库分裂的条件，目前支持的条件是记录数，其表示方法为： $\$RECORD=n$ ， n 是一个 10000 到 20000000 之间的任意数，当视图中最后一个数据库的记录数不小于 n 时将分裂出一个新的数据库。

选择需要转化的数据库（可一次选择多个），输入分裂条件，然后按[确定]，TRS 管理员会把所选择的数据库化为能自动分裂数据库的视图，并在状态条显示成功信息。按[取消]则放弃操作并关闭对话框。

第3章 使用视图对象

3.1 TRS 视图概述

视图是 TRS 系统中的一种辅助数据对象。TRS 系统支持对数据库记录和字段的进行直接的授权机制来实现数据的安全保密性控制，但有时显得有些烦琐。通过视图来限制某些用户对数据库字段和记录的查询，就可以用一种简便的方式间接地完成对数据库记录和字段查询权的控制。视图实际上是一个逻辑数据表，它是逻辑地从一个或多个数据库中抽取一个或多个字段和满足指定条件的记录。视图并不物理地存储数据，只是描述了组成该视图的数据所遵循的逻辑。每个 TRS 系统可管理多达数十亿个视图，每个视图中可包含 1 到 128 个数据库。

视图对象具有下列属性：

- **名称**

名称是 TRS 视图的对象名，即视图名，是访问、维护和管理该视图的入口。视图名不能与属于同一用户的其它视图、数据库以及各种词典同名。

- **别名列表**

别名是该视图名的别称，一般是为了方便使用而给出的视图名的缩写，或其英文名称。视图别名与视图名具有同等的地位。一个视图可以定义多个别名。

- **拥有者**

拥有者表明是哪个用户拥有该视图。拥有者与名称（别名）唯一确定一个视图对象。拥有者对该视图能够进行任何操作，而不受其它数据安全机制的制约。

- **缺省检索字段**

缺省检索字段是指在对该视图进行查询时，检索表达式没有明确指定检索目标字段的子表达式所默认的目标字段。

- **公共访问权限**

公共访问权限是指 TRS 系统中所有用户对该视图所拥有的权限。对视图可设置

的公共访问权限共有三种：

- 记录检索权(SELECT)
- 视图结构修改权(ALTER)
- 视图删除权(DROP)

● 所含数据库

所含数据库是指该视图中所包含的数据库列表。视图中不能包含视图。

● 记录选取条件

记录选取条件实际上是一个对该视图所包含的每个数据库进行检索的检索表达式。记录选取条件中的检索目标字段（包括缺省的目标字段）动态地对应于该视图所包含的每个数据库中的字段，而不是该视图的字段。**如果从多个数据库中选取数据的条件不一致，则可在各个数据库中建立“空字段”（不存储任何数据的字段）进行相互映射，以保证选取条件中的字段在各个数据库中都存在。**

● 说明信息

说明信息是对 TRS 视图对象进行说明的信息。

一般情况下，视图只能用于对数据的查询，而不支持对数据的“增、删、改”等操作，但 TRS 系统支持一种特殊的视图：数据库自动分裂模式视图，简称为自动模式视图。这种视图不但支持对数据记录的维护，而且随着数据记录的增加，能够自动分裂数据库，以维持单个数据库的规模。自动模式视图除了普通视图的三种访问权限外，还具有数据索引权限和数据更新权限的管理。

3.2 TRS 视图字段对象概述

视图字段是对视图进行查询的唯一入口，所有的查询操作都是在一定的字段上进行的。每个视图最多可定义 1023 个字段，每个字段可映射 1 到 64 个数据库中的某个字段。

视图字段对象具有下列属性：

● 字段名

字段名是 TRS 视图字段的对象名，即是访问、维护和管理该视图字段的入口。字段名不能与同一视图中的其它字段同名。

● 别名列表

别名是该字段名的别称，一般是为了方便使用而给出的字段名的缩写，或其英文

名称。字段别名与字段名具有同等的地位。一个字段可以定义多个别名。

- **字段类型**

字段类型表明了该字段所映射的数据库字段所存储的数据类型。对应于数据库字段，视图字段也有六种类型：

- 日期型(DATE)
- 数值型(NUMBER)
- 字符串型(CHAR)
- 短语型(PHRASE)
- 全文型(DOCUMENT)
- 二进制型(BIT)

- **字段显示名**

字段显示名是在用户界面上该字段的对应名字。

- **显示区宽度**

显示区宽度是字段值在用户界面上显示区域的缺省宽度单位数。

- **显示区高度**

显示区高度是字段值在用户界面上显示区域的缺省高度单位数。

- **公共检索开关**

公共查询开关决定是否允许 TRS 系统中的所有用户在该字段上进行查询。

- **源数据库字段**

源数据库字段决定了该字段对应于哪些数据库中的哪个字段。

- **说明信息**

说明信息是对 TRS 视图字段对象进行说明的信息

3.3 创建新的视图

必须是具有 Resource 权限的用户才可以创建视图。进入创建视图对话框有三种方法。

方法一：选择“文件|新建”命令，在“新建”对话框里双击视图图标；

方法二：当对象管理窗口显示视图列表切焦点也在其上时，按工具栏上的按钮；

方法三：当对象管理窗口显示视图列表切焦点也在其上时，在对象管理框按鼠标右键，出现菜单后选择新建。

通过上述方法之一将会开始执行“创建视图向导”，该向导共有如下三步：

第一步：输入视图名、拥有者名、说明信息、公共属性等信息，并可选择一个已存在的视图作为模板；对话框如下：



图 3—1 创建视图第一步

对话框里各控制的说明如下：

控制名称	功能简介
视图名称	新建视图的名字，不能包含中文、字母、数字或下划线以外的字符，大小写无关。名字长度在1-31个字符之间，不能是TRS保留字，也不能和拥

		有者的数据库、视图和词典重名。
拥有者名		新建视图所属的拥有者名，缺省为当前登录用户的用户名。
说明信息		不超过63个字符长度的视图说明信息。
模板视图		现有的所有视图的列表，可选择其中之一作为新视图的模板。缺省时不选择模板。
字段列表		所选模板视图的字段列表，可选择一个或多个字段作为新视图字段的模板。缺省时选择全部字段。
公共属性	公共检索	选中表示所有用户均可对该视图进行检索。
	公共修改	选中表示所有用户均可修改该视图的结构。
	公共删除	选中表示所有用户均可删除该视图。

输入视图名、拥有者名、说明信息、公共属性，选择好模板视图和字段列表以后，按[下一步]按钮进入向导的第二步。按[取消]则退出创建视图向导。

第二步：设置视图的别名列表、创建条件和源数据库等属性；

创建视图向导第二步的对话框如下：



图 3—2 创建视图第二步

对话框里各控制的说明如下：

控制名称	功能简介
别名列表	视图的别名列表，多个别名之间以半角分号、逗号或空格分隔，长度不能超过63个字符。缺省值为空。
记录选取条件	视图的创建条件。缺省值为空。长度不能超过1023个字符。
源数据库集	视图对应的源数据库集。点击...按钮弹出“选择源数据库”对话框，对视图对应的源数据库集进行选择或更改。 注意：只有当所包含的数据库字段成为视图的源数据库字段以后，当前新增加的数据库才会保存入视图的源数据库列表中，否则在提交更新后，服务器会自动将其从源数据库列表中剔除。

设置好视图的别名列表、创建条件和源数据库等属性以后，按[下一步]按钮进入向导的第三步。按[上一步]可返回向导的第一步，按[取消]则退出创建视图向导。

第三步：设计视图库的字段结构以及各字段的具体属性等。

创建视图向导第三步的对话框如下：

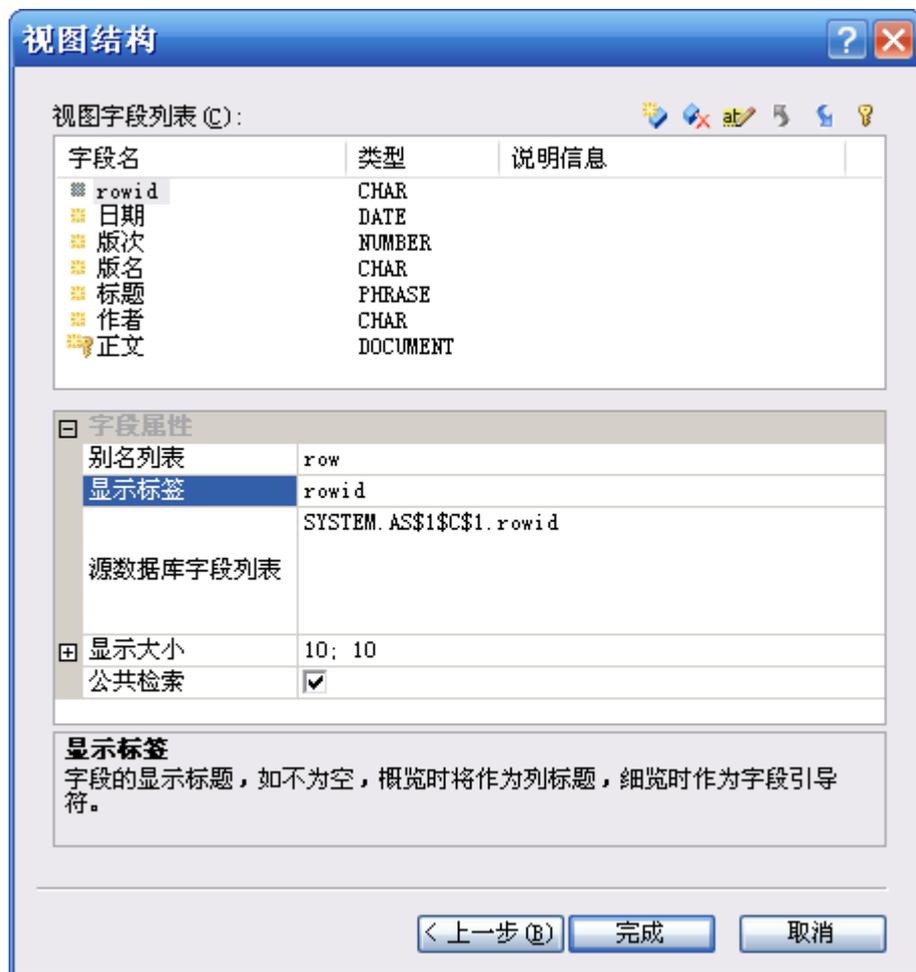


图 3—3 创建视图第三步

视图字段列表控制里列出新建视图的所有字段，第一次进入时显示的是第一步里选择的模板字段，若没有选择模板，则自动将所选的各源数据库中名称类型相同的字段列出，字段属性取第一个字段的相应属性。字段列表的每一行分别显示字段的名称、类型和说明，字段名前有  标志的字段为新增字段，有  标志的字段为视图进行全文检索的缺省全文字段名，不指定的话自动设为最后一个 DOCUMENT 字段（没有 DOCUMENT 字段的话则选择最后一个 PHRASE 字段）。用户可在模板字段的基础上进行添加字段、修改字段属性、删除字段、为字段改名、调整字段顺序(仅限于新字段)等操作。视图字段列表控制的右上角有六个按钮，它们的作用分别如下：



添加新字段，快捷键为 Insert，点击此按钮后将弹出“选择源数据库字段”对话框，选定源数据库字段以后，点击[确定]将自动生成一个相应的字段，类型和源数据库字段类型一样



删除当前选中字段，快捷键为 Delete



为当前选中字段改名，快捷键为空格，直接在字段名上单击也能起到同样作用



当前选中字段位置上移，仅限于新字段，快捷键为 Shift+↑



当前选中字段位置下移，仅限于新字段，快捷键为 Shift+↓



把当前选中字段设为缺省字段，快捷键为 Shift+空格

视图字段列表控制下面是字段属性组框，组框里的控制显示当前视图字段列表控制里选中的字段的各项属性，具体如下：

控制名称	功能简介
别名列表	视图字段的别名列表，多个别名之间以半角分号、逗号或空格分隔。缺省值为空。
显示标签	视图字段的显示标题，如不为空，概览时将作为列标题，细览时作为字段引导符。
源数据库字段列表	视图字段对应的源数据库字段列表，每个数据库字段名的表示方式为：“所有者.数据库名.字段名”，源数据库字段数必须大于零。当视图字段对应多个源数据库字段时，这些源数据库字段的类型必须一致。按按钮弹出“选择源数据库字段”对话框，对视图对应的源数据库字段进行选择或更改。
显示大小	视图字段的显示宽度和高度。
公共检索	选中表示所有用户均可对本视图字段进行检索。

设计好视图的字段结构以及各字段的具体属性后，按[完成]可结束向导，并创建视图。按[上一步]可返回向导的第二步，按[取消]则退出创建视图向导。

3.4 修改 TRS 视图对象

修改视图。如果满足下列条件之一，则有权修改视图：

- 超级用户。
- DBA用户。
- 视图的拥有者。
- 视图具有公共ALTER属性。
- 与视图的拥有者同用户组的GDA用户。
- 拥有该视图的ALTER权的用户。

视图的属性包括常规属性、结构和访问权限三个方面，其中常规属性在属性窗口中的属性框里显示，结构和访问权限以对话框显示，用户可对其属性进行修改（当前登录用户必须对该视图有修改权限）。修改完毕后，按属性窗口上的按钮更新则保存对常规属性

的修改。按对话框上的[确定]按钮，将保存对结构或者访问权限所作的修改，按[取消]则放弃所作的修改并关闭对话框。若属性窗口已关闭，可选择菜单上的“文件|显示属性”命令，

或按工具栏上的按钮重新打开属性窗口。

3.4.1 修改视图的常规属性

在对象管理窗口里选中要查看属性的视图对象，属性窗口将会看到如下图：

常规属性	
内部编号	118
对象名称	rmbxxxxxx
所有者	SYSTEM
对象类型	视图
说明信息	
创建时间	2008.05.14 09:41:22
修改时间	2009.08.05 11:55:45
字段数	10
数据库数	8
高级属性	
别名列表	人民日报
记录选取条件	中国,北京
源数据库集	SYSTEM.rmb1948; SYSTEM.rmb1949; SYSTEM.rmb1977;
公共属性	
公共检索	<input type="checkbox"/>
公共修改	<input checked="" type="checkbox"/>
公共删除	<input type="checkbox"/>

图 3—4 视图属性

属性框里各控制的说明如下：

	控制名称	功能简介
常规属性	内部编号	视图在系统中的内部标识号。
	对象名称	TRS视图的对象名，即视图名，是访问、维护和管理该视图的入口。视图名不能与属于同一用户的其它视图、数据库以及各种词典同名。
	所有者	视图的拥有者。
	对象类型	对象的类型。附有“(自动分裂)”字样的为自动模式视图。
	说明信息	不超过63个字符长度的视图说明信息。
	创建时间	视图的创建时间。
	修改时间	视图的最后一次修改时间。
	字段数	视图的字段数。
	数据库数	视图对应的源数据库数。

高级属性	别名列表	视图的别名列表，多个别名之间以半角分号、逗号或空格分隔，长度不能超过63个字符。
	记录选取条件	实际上是一个对该视图所包含的每个数据库进行检索的检索表达式。记录选取条件中的检索目标字段（包括缺省的目标字段）动态地对应于该视图所包含的每个数据库中的字段，而不是该视图的字段。如果从多个数据库中选取数据的条件不一致，则可在各个数据库中建立“空字段”（不存储任何数据的字段）进行相互映射，以保证选取条件中的字段在各个数据库中都存在。其长度不能超过1023个字符。
	源数据库集	视图对应的源数据库集。此处为只读选项，添加或删除源数据库在“视图结构”对话框中的添加新字段操作时进行。 注意：只有当所包含的数据库字段成为视图的源数据库字段以后，新增加的数据库才会保存入视图的源数据库列表中，否则在提交更新后，服务器会自动将其从源数据库列表中剔除。
公共属性	公共检索	选中表示所有用户均可对该视图进行检索。
	公共更新	选中表示所有用户均可对该视图进行记录维护（仅限自动模式视图）。
	公共索引	选中表示所有用户均可对该视图的索引进行维护（仅限自动模式视图）。
	公共删除	选中表示所有用户均可删除该视图。

如对视图的属性作了修改，按  将保存所作的修改。

3.4.2 修改视图结构

若当前属性窗口显示的对象为视图，直接选择菜单上的“文件|结构”命令，或点击工具栏上的  按钮，将会看到如下对话框：

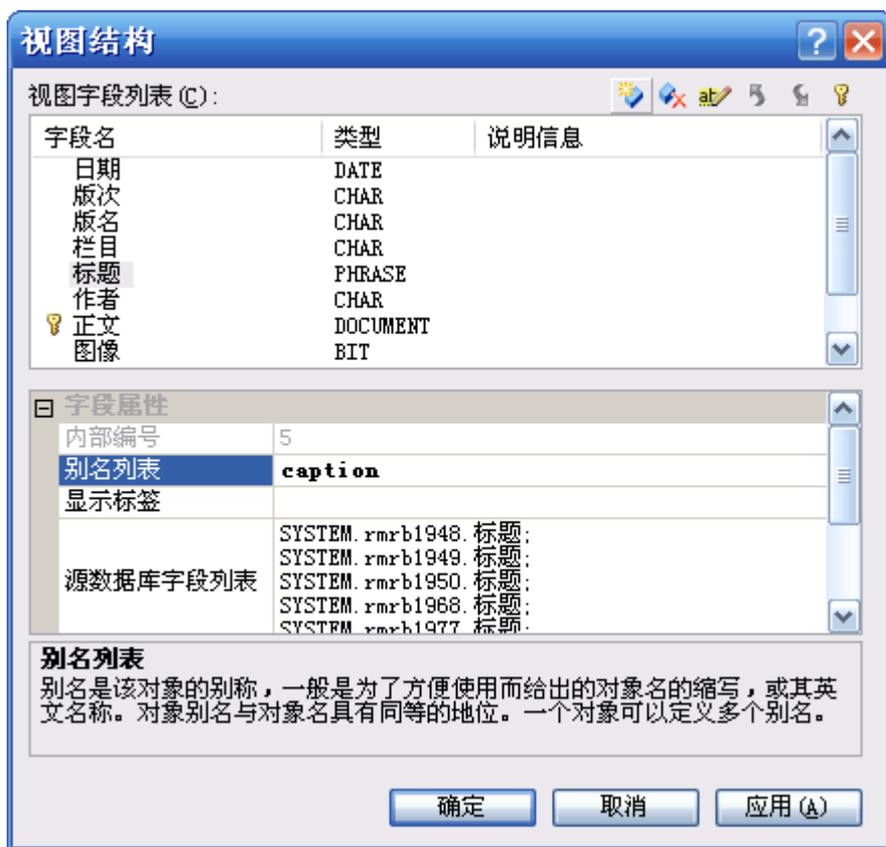


图 3—5 视图结构对话框

视图字段列表控制里列出视图的所有字段。字段列表的每一行分别显示字段的名称、类型和说明，字段名前带有 标志的字段为新增字段，带有 标志的字段为视图进行全文检索的缺省全文字段名，不指定的话自动设为最后一个 DOCUMENT 字段（没有 DOCUMENT 字段的话则选择最后一个 PHRASE 字段）。用户可在字段列表上进行添加字段、修改字段属性、删除字段、为字段改名、调整字段顺序（仅限于新字段）等操作。视图字段列表控制的右上角有六个按钮，它们的作用分别如下：

- 添加新字段，快捷键为 Insert
- 删除当前选中字段，快捷键为 Delete
- 为当前选中字段改名，快捷键为空格，直接在字段名上单击也能起到同样作用
- 当前选中字段位置上移，仅限于新字段，快捷键为 Shift+↑
- 当前选中字段位置下移，仅限于新字段，快捷键为 Shift+↓
- 把当前选中字段设为缺省字段，快捷键为 Shift+空格

视图字段列表控制下面是字段属性组框，组框里的控制显示当前视图字段列表控制里

选中的字段的各项属性，具体如下：

控制名称	功能简介
内部编号	视图字段在系统中的内部标识号，新字段不显示此属性。
别名列表	视图字段的别名列表，多个别名之间以半角分号、逗号或空格分隔，单个别名长度不能超过31个字符，总长度（含分隔符）不能超过63个字符。
显示标签	视图字段的显示标题，如不为空，概览时将作为列标题，细览时作为字段引导符。
源数据库字段列表	视图字段对应的源数据库字段列表，每个数据库字段名的表示方式为：“所有者.数据库名.字段名”，源数据库字段数必须大于零。当视图字段对应多个源数据库字段时，这些源数据库字段的类型必须一致。按  按钮弹出“选择源数据库字段”对话框，对源数据库字段列表进行选择或更改。
显示大小	视图字段的显示宽度和高度。
缓冲数据	选中表示允许进行检索缓冲，缓冲可以提高检索时有关操作的速度。只有自动模式视图的DATE、NUMBER或CHAR型字段的值才允许进入检索缓冲区，并且只有当“检索缓冲”已经启动时，该开关才有效。此属性只有自动模式视图拥有。
允许相关性排序	选中表示检索命中该字段时，该字段参与相关性排序，缺省为清除。对BIT字段该开关无效。此属性只有自动模式视图拥有。
公共检索	选中表示所有用户均可对本视图字段进行检索。

如对视图的字段结构作了修改，按[确定]将保存所作的修改并关闭对话框。

3.4.3 访问权限

若当前属性窗口显示的对象为视图，直接选择菜单上的“文件|权限”命令，或按工具栏上的按钮，将会看到如下对话框：



图 3—6 视图的访问权限

“访问权限”属性页里各控制的说明如下：

控制名称	功能简介
用户组/用户	当前服务器上所有用户组/用户的列表。用户组/用户名前面的小图标表明了该用户组/用户的类型：表示用户组、表示全域管理员用户、表示组管理员用户、表示资源用户、表示登录用户。
字段	对象名及其所有字段的列表，“(对象权限)”相当于对视图的操作权限。
访问权限	当前选中的用户组/用户对指定对象/字段的访问权限，包括以下所列几项。 直接单击组里的复选框可修改所有处于选择状态的用户组/用户对所有选择的对象/字段的访问权限，不需按“确定”或“应用”即可立刻生效。
检索权限	选中表示当前用户组/用户对视图和指定字段有检索权限。
更新权限	选中表示当前用户组/用户对视图有记录更新权限（仅限自动模式视图）。
索引权限	选中表示当前用户组/用户对视图有索引维护权限（仅限自动模式视图）。
修改权限	选中表示当前用户组/用户对视图有结构修改权限。
删除权限	选中表示当前用户组/用户对视图有删除权限。

按[确定]或[取消]将并关闭对话框。

3.5 视图重命名

视图重命名要求操作用户必须拥有 DBA 权限或为当前视图的拥有者。

先选中要改名的视图，然后选择“文件”菜单上的“重命名”，或按工具栏上的按钮，或直接按空格键，当前视图名字的位置上将出现一编辑框；在编辑框中输入新的名字，然后按 ENTER 键，所选视图将被重新命名为所输入的名字；如要取消重命名操作，可按 ESC 键。此外，还可以在属性窗口中直接修改，点击视图，属性窗口将显示其相应的属性，在“对象名称”一栏可直接对视图名进行修改，修改完毕点击工具栏上的按钮即可保存新名字。

3.6 删除 TRS 视图对象

删除视图。如果满足下列条件之一，则有权删除视图：

- 超级用户。
- DBA用户。
- 视图的拥有者。
- 视图具有公共DROP属性。
- 与视图的拥有者同用户组的GDA用户。
- 拥有该视图的DROP权的用户。

使用 TRS 管理员工具删除 TRS 视图对象，首先要打开对象管理的视图标签，通过四种方法可以完成对视图对象的删除。

方法一：选择要删除的视图，按下鼠标的右键并选择“删除”或“剪切”。

方法二：选择要删除的视图，选择菜单“编辑|删除”或“编辑|剪切”。

方法三：选择要删除的视图，按下“Delete”键或“Ctrl+X”。

方法四：选择要删除的视图，选择工具栏中或按钮。

系统会出现如下对话框：



图 3—7 删除视图确认对话框

选择“是(Y)”删除视图；选择“否(N)”放弃删除操作。

“删除”和“剪切”的区别在于剪切操作把所选的视图复制到剪贴板里，而删除则不进行复制。利用“剪切”和“粘贴”功能的配合可快速复制视图的结构。

3.7 加入数据库

此命令仅对自动模式视图有效，用于往自动模式视图里添加新的数据库。注意新增数据库的结构必须和自动模式视图一致。

从对象管理窗口中视图页的快捷菜单中选择“加入数据库”命令，TRS 管理员弹出如下对话框：



图 3—8 加入数据库对话框

对话框里的“目标自动模式视图”列表框里列出了当前所有自动模式视图，从“加入数据库名”列表中选择需要加入的数据库（可一次选择多个），然后按[确定]，TRS 管理员会把所选择的数据库添加到指定自动模式视图里，并在状态条显示成功信息。按[取消]则放弃操作并关闭对话框。

第4章 使用词典对象

4.1 TRS 词典概述

TRS 的全文检索不单是一种快速的字符串匹配系统，要获得良好的检索效果，必须使用和参考词典。词典对象是 TRS 系统中非常重要的基础资源，是按词索引和检索技术的根本依据。

词典对象具有下列属性：

- 名称

名称是 TRS 词典的对象名，即词典名，是访问、维护和管理该词典的入口。词典名不能与属于同一用户的其它视图、数据库以及各种词典同名。

- 别名列表

别名是该词典名的别称，一般是为了方便使用而给出的词典名的缩写，或其英文名称。词典别名与词典名具有同等的地位。一个词典可以定义多个别名。

- 拥有者

拥有者表明是哪个用户拥有该词典。拥有者与名称（别名）唯一确定一个词典对象。拥有者对该词典能够进行任何操作，而不受其它数据安全机制的制约。

- 类型

按用途划分，TRS 系统有八种类型的词典：

- 分词词典(SEGMENT)
- 附加分词词典(SEGMENTEX)
- 停用词典(STOP)
- 附加停用词典(STOPEX)
- 稀疏词典(SPARSE)
- 主题词典(THESAURUS)
- 同义词典(SYNONYM)
- 反义词典(ANTONYM)

其中，分词词典、附加分词词典、停用词典、附加停用词典和稀疏词典统称为**索引词典**，主题词典、同义词典和反义词典统称为**辅助知识词典**。

● 公共访问权限

公共访问权限是指 TRS 系统中所有用户对该词典所拥有的权限。对词典可设置的公共访问权限共有三种：

- 检索引用权(SELECT)
- 更新维护权(UPDATE)
- 删除权(DROP)

● 使用语言

使用语言是指该词典语言版本。TRS 系统支持的使用语言有：

- 简体中文(GB2312、GBK、GB18030)
- 繁体中文(BIG5)
- 纯英文(ENGLISH)
- 统一码(UTF8)

● 数据存放路径

数据存放路径指明词典数据文件的存储目录。词典创建后，其存储目录任何时候都不允许修改。

● 说明信息

说明信息是对 TRS 词典对象进行说明的信息。

系统安装时，将自动创建以下几部系统缺省词典：分词词典、停用词典、主题词典、同义词典和反义词典。

4.1.1 分词词典

按词索引和检索是 TRS 的主要特点之一，内嵌的分词系统采用以词典为基础的分词算法。分词词典用于对表意文字（中文、日文、韩文）自动分词，由若干个词汇组成，词典中由字母和数字组成的词汇不起作用。每部分词词典可容纳多达 50 万条词汇，每个词汇的最大长度为 48 个字节。

分词词典由词汇组成，TRS 系统提供的缺省分词词典的每个词汇还具有语法属性，以提高分词的准确性。用户定义的分词词典或增加的词汇则没有属性。该词典是经过加密处理的，用户可以浏览词典的词汇，但不能浏览其属性。每部分词词典可容纳多达 10 万条词汇，每个词汇的最大长度为 48 个字节。

分词词典的使用要点：

- 在创建数据库时引用分词词典。分词词典被引用后，一般不允许再对该词典进行维护，除非重新创建数据库的索引。
- 在一般情况下均使用系统提供的缺省分词词典。
- 对英文数据库不使用分词词典。
- 一般不对系统缺省分词词典进行维护，当需要加入用户专业词汇时，可通过创建附加分词词典来实现。
- 分词词典中的英文词汇不起作用，因此不要在分词词典中加入英文词汇。
- 生僻词并不一定要加入到分词词典，这些词即使不在词典中，也能检索。

TRS 系统提供一个特殊空分词词典，其特殊性在于：系统安装或升级时，该词典将被自动创建，不能进行维护，也不能以其它任何手工方式创建（否则不具有自动创建时的功能）；词典中没有任何词项，当数据库引用该词典时，将按字建立索引，并用来指定“所有的单字都建索引”，以便能够进行单个字的查询。因此也把该词典叫做“**全字词典**”。

4.1.2 附加分词词典

附加分词词典是分词词典的补充。TRS 提供的缺省分词词典是一部通用的分词词典，在多数情况下没有包括用户应用的特殊词汇。在这种情况下，一般不是对系统缺省分词词典进行维护，而是通过建立新的附加分词词典来定义新的词汇，系统在自动分词时将同时参考分词词典和附加分词词典中的词汇。

附加分词词典由一系列词汇组成，词典中的英文词汇不起作用。每部附加分词词典可容纳多达 5 万条词汇，每个词汇的最大长度为 48 个字节。

附加分词词典的使用要点：

- 在创建数据库时引用附加分词词典。附加分词词典被引用后，一般不允许再对该词典进行维护，除非重新创建数据库的索引。
- 在一般情况下均使用系统提供的缺省分词词典。即使在数据库不引用任何附加分词词典，用户应用的特殊词汇也能进行检索。
- 对英文数据库不使用分词词典。
- 附加分词词典中由字母和数字组成的词汇不起作用，因此不要在加入这些词汇。
- 附加分词词典一般包含了某个领域的专业词汇。

4.1.3 停用词典

停用词典又称为禁用词典，它是由一系列没有检索意义的高频词组成的，如英文文献中的“the”、“of”、“and”、“to”等，中文文献中的“的”、“关于”、“但是”、“而且”等。从相关性方面讲，文献中的这些词没有检索意义，因为这些词会出现在每篇

文献中。在检索系统中，通常使用停用词典来过滤掉文献中没有检索意义的词，以最大限度地减少数据库的空间膨胀率、加快查询速度。

停用词典中的词汇可以是中英文标点符号、中英文高频词等。每部停用词典可容纳多达 5 千条词汇，每个词汇的最大长度为 48 个字节。

停用词典的使用要点：

- 在创建数据库时引用停用词典。
- 一般不对系统缺省停用词典进行维护，当需要过滤更多的无意义词时，可通过创建附加停用词典来实现。
- 无论是何种语言的数据库，均可引用停用词典。
- 并不是语言中所有的高频词都需要作为停用词，例如“家”、“世界”是高频词，但对大多数社会科学资料数据库来说，它们可能是重要的词汇。一个词是不是需要作为停用词，与数据库的领域特点有关，如在计算机科学文献中，计算机可作为停用词，因为它几乎出现在每一篇文献中，没有检索意义。

4.1.4 附加停用词典

附加停用词典是停用词典的补充。TRS 提供的缺省停用词典是一部通用的停用词典，不包括特殊领域的无检索意义的词汇。在这种情况下，一般不是对系统缺省停用词典进行维护，而是通过建立新的附加停用词典，并在创建数据库时引用该附加停用词典。

附加停用词典中的词汇可以是标点符号、高频词等。每部附加停用词典可容纳多达 5 千条词汇，每个词汇的最大长度为 48 个字节。

附加停用词典的使用要点：

- 在创建数据库时引用附加停用词典。
- 无论是何种语言的数据库，均可引用附加停用词典。
- 附加停用词典一般包含了某个领域的高频词。

4.1.5 稀疏词典

在某些应用领域中，用户并不需要进行全文检索，而希望从文献中自动抽取一些词汇作为检索的关键词，也就是说并不需要索引文献中的所有词汇，而只需索引其中一些关键词。稀疏词典就是存储这些关键词汇的词典。

关键词可以是任何词汇，包括由字母和数字组成的词汇。每部稀疏词典可容纳多达 5 万条词汇，每个词汇的最大长度为 48 个字节。

稀疏词典的使用要点：

- 在创建数据库时引用稀疏词典。稀疏词典被引用后，一般不允许再对该词典进行维护，除非

重新创建数据库的索引。

- 稀疏词典一般单独使用。
- 稀疏词典中的词汇可以是任何词汇，包括由字母和数字组成的词汇。

4.1.6 主题词典

主题词典又称为叙词表，它是一种语义词典，由词及其各种关系组成，能反映某学科领域的语义相关概念。TRS 系统支持 ANSI Thesaurus 标准 (Z39. 19-1980) 所规定的所有 13 种词间关系。这 13 种关系完全包括了我国《汉语主题词表》的“用、代、属、分、参”结构。

主题词典主要用于检索时的后控制和标引时的自动或辅助选择索引词，是提高查全率和查准率、实现多语种检索和智能化概念检索的重要途径。在 TRS 系统中，主题词典是一种特殊的数据库，所以其规模能够达到一般数据库的所能达到的规模。

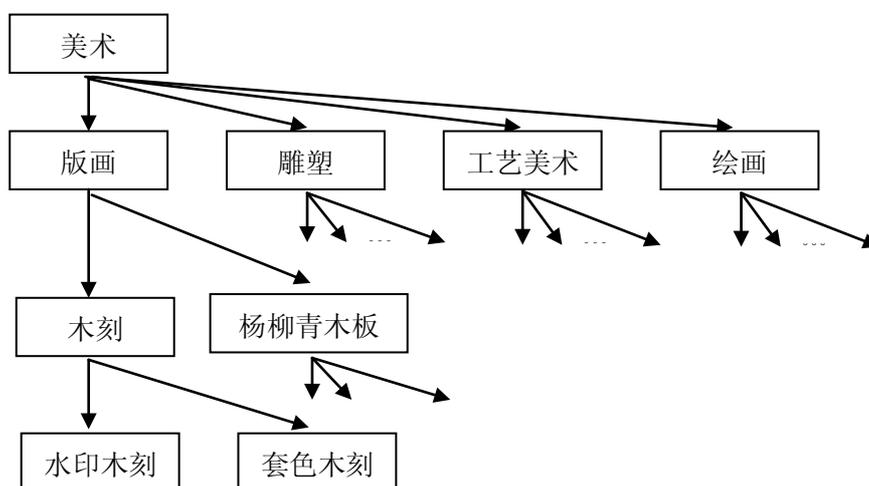
下表描述了 ANSI 主题词表的结构：

关系	说明	举例
族首词 (Lead Term)	概念上外延最广的主题词	“木刻”的族首词是“美术”
上位词 (Broader Term)	概念上外延较广的主题词	“木刻”的上位词是“版画”
下位词 (Narrower Term)	概念上内涵更窄的主题词	“木刻”的下位词有“水印木刻”和“套色木刻”
等同词 (Equivalence, Preferred Term)	也称正式主题词或规范主题词	“电子计算机”是“电脑”的等同词
替代词(用代词) (Used For, Non-Preferred Term)	也称非正式主题词或非规范主题词	“电脑”、“计算机”是“电子计算机”的替代词
相关词 (Related Term)	概念相近但不是等关系的主题词	“计算机应用”和“自动控制”是相关关系
缩略词 (Abbreviation)	主题词的缩略词	“Database Management System”的缩略词为“DBMS”
被所略词 (Abbreviation For)	缩略词的正式名称	“DBMS”的正式名称为“Database Management System”
组合概念 (Combined Concepts)	为将来TRS版本预留	
历史注释 (History Note)	对主题词的历史进行说明	
范围注释 (Scope Note)	对主题词的适用范围进行说明	
外文等同词 (Language Equivalent)	其它语种的等同词	“计算机”的英文等同词是“Computer”
后组配概念 (Post-coordinated Concept)	为将来TRS版本预留	

对应于主题词表的结构，TRS 系统定义了相应的函数以引用结构中的各种关系：

函数名	含义和作用
BT(主题词[,词典])	获得某个主题词的上位词
NT(主题词[,词典])	获得某个主题词的下位词
PT(主题词[,词典])	获得某个主题词的等同词
UF(主题词[,词典])	获得某个主题词的替代词
AB(主题词[,词典])	获得某个主题词的缩略词
ABF(主题词[,词典])	获得某个主题词的被缩略词
RT(主题词[,词典])	获得某个主题词的相关词
LE(主题词[,词典])	获得某个主题词的外文等同词
ALLBT(主题词[,词典])	获得某个检索词的所有层次的上位词
ALLNT(主题词[,词典])	获得某个检索词的所有层次的下位词

下图是一个“下位词”网络的局部示意图：



主题词典的使用要点：

- 通过浏览主题词典进行词汇的交互式选择完成检索。
- 利用主题词典函数，或自动扩展功能进行多语种和智能化概念检索。此时必须保持词典的字符集必须与所检索的数据库的字符集保持一致，否则无效。

在TRS系统中，主题词典是一种特殊的数据库，其最大规模与一般的数据库相同。主题词典主要用于检索时的后控制和标引时的自动或辅助选择索引词，是提高查全率和查准率、实现多语种检索和智能化概念检索的重要途径。为此TRS系统定义了与主题词关系相对应的函数以使能在检索表达式中引用词典中的各种关系。

4.1.7 同义词典

同义词典是对主题词典中替代关系的扩充，以实现范围更大和意义更广的同义词扩展检索，以提高查全率，是 TRS 系统完成智能化检索的又一种手段。比如：假设在同义词典

中“计算机”的同义词为“电脑”，那么在使用该同义词典对“计算机”进行扩展检索时，所有包含“计算机”或“电脑”的文献将被检索出来。

在主题词典中，有一些关系（如替代、缩略等）也具有“同义词”的含义，这些关系与同义词典的区别在于：

- 主题词典中的主题词及其关系是经过严格控制的，一般不能随意地加以变动，而同义词典则没有这么严格，可以比较随意地进行增加、删除和修改等维护。
- 主题词典中一般只收录概念词，不收录或很少收录专用名词，如人名、机构名等，这些词应在同义词典中定义。

在 TRS 系统中，同义词典是一种特殊的数据库，所以其规模能够达到一般数据库的所能达到的规模。同义词典的内部结构非常简单，一般不需要了解。为了便于引用同义词典，TRS 系统定义了一个同义词典函数：ST(词汇[,词典])。通过该函数可在指定的同义词典中获取给定词汇的同义词集。

同义词典的使用要点：

- 通过浏览同义词典进行词汇的交互式选择完成检索。
- 利用同义词典函数，或自动扩展功能进行智能化概念检索。此时必须保持词典的字符集必须与所检索的数据库的字符集保持一致，否则无效。

下表是与同义词典有关的检索函数：

函数名	含义和作用
ST(词汇[,词典])	获得某个词的同义词

4.1.8 反义词典

反义词典主要用来进行反义词的联想检索，也是 TRS 系统实现智能化检索的手段之一。比如：假设在反义词典中“拥护”的反义词为“反对”，那么在使用该同义词典对“拥护”进行扩展检索时，所有包含“拥护”或“反对”的文献将被检索出来。

在 TRS 系统中，反义词典是一种特殊的数据库，所以其规模能够达到一般数据库的所能达到的规模。反义词典的内部结构非常简单，一般不需要了解。为了便于引用反义词典，TRS 系统定义了一个反义词典函数：AT(词汇[,词典])。通过该函数可在指定的反义词典中获取给定词汇的反义词集。

反义词典的使用要点：

- 可以在检索时，通过浏览反义词典进行交互式选择进行检索。
- 利用反义词典函数，或自动扩展功能进行智能化概念检索。

下表是与反义词典有关的检索函数：

函数名	含义和作用
-----	-------

AT(词汇[,词典]) 获得某个词的反义词

4.2 创建新的词典

必须是具有 Resource 权限的用户才可以创建词典。进入创建词典对话框有三种方法。

方法一：选择“文件|新建”命令，在“新建”对话框里双击词典图标；

方法二：当对象管理窗口显示词典列表切焦点也在其上时，按工具栏上的按钮；

方法三：当对象管理窗口显示词典列表切焦点也在其上时，在对象管理框按鼠标右键，出现菜单后选择新建。

选择上述方法之一，将会看到如下对话框：



创建词典对话框的截图显示如下：

- 字典名称 (N): MySegdic
- 拥有者名 (O): SYSTEM
- 字典类型 (T): 主分词词典 (SEGDIC)
- 说明信息 (C): 我的主分词词典
- 别名列表 (A):
- 使用语言 (L): 中文GB2312/GBK/GB18030
- 存放路径 (P): E:\trshome\ibase
- 公共属性:
 - 公共检索 (S)
 - 公共更新 (U)
 - 公共删除 (D)

底部有“确定”和“取消”按钮。

图 4—1 新建词典对话框

对话框里各控制的说明如下：

控制名称	功能简介
词典名称	新建词典的名字，不能包含中文、字母、数字或下划线以外的字符，大小写无关。名字长度在1-31个字符之间，不能是TRS保留字，也不能和拥有者的数据库、视图和词典重名。
拥有者名	新建词典所属的拥有者名，缺省为当前登录用户的用户名。

词典类型	新建词典的类型，可为以下几种类型之一：分词词典、附加分词词典、停用词典、附加停用词典、稀疏词典、主题词典、同义词典、反义词典。	
说明信息	不超过63个字符长度的词典说明信息。	
别名列表	词典的别名列表，多个别名之间以半角分号、逗号或空格分隔，单个别名长度不能超过31个字符，总长度（含分隔符）不能超过63个字符。	
使用语言	词典所用的语言，有如下几种取值（缺省为系统配置中的缺省语言环境）： 中文GB2312/GBK/GB18030、中文BIG5、英语、统一码UTF8。	
存放路径	词典存放的路径，必须是系统配置中的数据和索引路径之一，缺省为系统安装路径下的dict子目录。可以在下拉列表中选择。	
公共属性	公共检索	选中表示所有用户均可对该词典进行检索，总是选中。
	公共维护	选中表示所有用户均可对该词典进行维护。
	公共删除	选中表示所有用户均可删除该词典。

设定好词典名、拥有者、词典类型、说明信息以及其他词典属性后，按[确定]，TRS 管理员将在当前服务器上建立新的词典。如想放弃新建词典，请按[取消]。

4.3 修改词典的属性

用户可方便地查看和修改当前选中词典的属性、结构和权限等信息。如要对词典的属性进行修改，则当前用户必须是 DBA 用户，或为选中词典的拥有者，或对选中词典有 ALTER 权限。

词典的属性包括常规属性和访问权限两个方面，其中常规属性在属性窗口中的属性框里显示，访问权限以对话框显示，用户可对其属性进行修改。修改完毕后，按属性窗口上的按钮  则保存对常规属性的修改。若属性窗口已关闭，可选择菜单上的“文件|显示属性”命令，或按工具栏上的  按钮重新打开属性窗口。

4.3.1 修改词典的常规属性

在对象管理窗口里选中要查看属性的词典对象，属性窗口将会看到如下图：

常规属性	
对象名称	THESAURUS_GB
拥有者	SYSTEM
对象类型	主题词典
说明信息	This is system thesaurus
创建时间	2006-07-24 11:30:29
修改时间	2007-10-29 10:46:52
字段数	14
记录数	9799
锁定状态	共享 (SHARE)
锁定用户	
高级属性	
别名列表	
使用语言	中文GB2312/GBK/GB18030
存放路径	F:\TRS\TRS6\dict
公共属性	
公共检索	<input checked="" type="checkbox"/>
公共更新	<input type="checkbox"/>
公共删除	<input type="checkbox"/>

图 4—2 词典属性

属性框里各控制的说明如下：

	控制名称	功能简介
常规属性	对象名称	TRS词典的对象名，即词典名，是访问、维护和管理该词典的入口。词典名不能与属于同一用户的其它视图、数据库以及各种词典同名。
	拥有者	词典的拥有者。
	对象类型	词典的类型，有八种取值：主分词词典、附加分词词典、停用词典、附加停用词典、稀疏词典、主题词典、同义词典、反义词典。
	说明信息	不超过63个字符长度的词典说明信息。
	创建时间	词典的创建时间。
	修改时间	词典的最后一次修改时间。
	字段数	词典的字段数。
	记录数	词典的记录数。
	锁定状态	词典的当前访问状态，有如下几种取值：共享锁(SHARE)、访问互斥锁 (ACCESS_EXCL)、修改锁 (UPDATE_EXCL)、修改等待锁 (UPDATE_WAIT)、拒绝访问锁 (ACCESS_DENY)。
锁定用户	锁定词典的用户名。	
高级属性	别名列表	词典的别名列表，多个别名之间以半角分号、逗号或空格分隔，长度不能超过63个字符。
	使用语言	词典所用的语言，有如下几种取值：中文GB2312/GBK/GB18030、中文BIG5、英语、统一码UTF8。
	存放路径	词典存放的路径，必须是系统配置中的数据和索引路径之一。不可修改。
公共	公共检索	选中表示所有用户均可对该词典进行检索，总是选中。
	公共维护	选中表示所有用户均可对该词典进行维护。

属性	公共删除	选中表示所有用户均可删除该词典。
----	------	------------------

如对词典的属性作了修改，按  将保存所作的修改。

4.3.2 访问权限

词典的访问权限只有检索权限、更新权限和删除权限三种，词典没有字段级权限。若当前属性窗口显示的对象为词典，直接选择菜单上的“文件|权限”命令，或按工具栏上的

 按钮，将会看到如下对话框：



图 4—3 词典访问权限对话框

“访问权限”属性页里各控制的说明如下：

控制名称	功能简介
用户组/用户	当前服务器上所有用户组/用户的列表。用户组/用户名前面的小图标表明了该用户组/用户的类型：  表示用户组、  表示全域管理员用户、  表示组管理员用户、  表示资源用户、  表示登录用户。
字段	词典没有字段级权限，因此这里的权限操作都是针对整个词典而言。
访问权限	当前选中的用户组/用户对指定对象/字段的访问权限，包括以下所列几项。直接单击组里的复选框可修改所有处于选择状态的用户组/用户对所

	有选择的对象/字段的访问权限，不需按“确定”或“应用”即可立刻生效。
检索权限	选中表示当前用户组/用户对词典有检索权限，总是选中。
更新权限	选中表示当前用户组/用户对词典有记录更新权限。
删除权限	选中表示当前用户组/用户对词典有删除权限。

按[确定]或[取消]将并关闭对话框。

4.4 词典重命名

重命名词典要求操作用户必须拥有 DBA 权限或为当前词典的拥有者。

先选中要改名的词典，然后选择“文件”菜单上的“重命名”，或按工具栏上的  按钮，或直接按空格键，当前词典名字的位置上将出现一编辑框；在编辑框中输入新的名字，然后按 ENTER 键，所选词典将被重新命名为所输入的名字；如要取消重命名操作，可按 ESC 键。此外，还可以在属性窗口中直接修改，点击词典，属性窗口将显示其相应的属性，在“对象名称”一栏可直接对词典名进行修改，修改完毕点击工具栏上的  按钮即可保存新名字。

4.5 删除 TRS 词典对象

如果满足下列条件之一，则有权删除词典：

- 超级用户。
- DBA 用户。
- 词典的拥有者。
- 词典具有公共 DROP 属性。
- 与词典的拥有者同用户组的 GDA 用户。
- 拥有该词典的 DROP 权的用户。

使用 TRS 管理员工具删除 TRS 词典对象，首先要打开对象管理的词典标签，通过四种方法可以完成对词典对象的删除。

方法一：选择要删除的词典，按下鼠标的右键并选择“删除”或“剪切”（如图）

方法二：选择要删除的词典，选择菜单“编辑|删除”或“编辑|剪切”。

方法三：选择要删除的词典，按下“Delete”键或“Ctrl+X”。

方法四：选择要删除的词典，选择工具栏中  或  按钮。

系统会出现如下对话框：

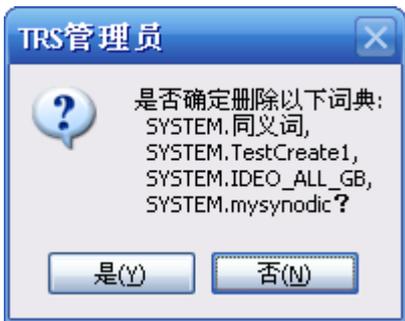


图 4—4 删除词典确认对话框

- 选择“是(Y)”删除词典结构（同时删除所有数据）；
- 选择“取消”放弃删除操作。

“删除”和“剪切”的区别在于剪切操作把所选的词典复制到剪贴板里，而删除则不进行复制。利用“剪切”和“粘贴”功能的配合可快速复制词典。

4.6 词典的维护

与用户自定义的数据库不同，对词典的维护通过对话框来实现，不同类型的词典对应不同的对话框，但各种词典的维护方式是很相似的。在对象管理窗口里切换到词典查看标签，选中需要维护的词典，再选择“文件|检索”菜单或直接按回车键，即可打开该词典进行维护，直接双击词典也可达到同样的效果。

与数据库和视图不同，一次只可打开一个词典进行维护，而且必须关闭词典维护对话框后才能进行其它 TRS 操作。

4.6.1 索引词典的维护

索引词典包括主分词词典、附加分词词典、停用词典、附加停用词典、稀疏词典五种，对这五种词典的维护操作是一致的。在对象管理窗口里双击索引词典对象，或选中需要查询和维护的索引词典，并按回车或选择“文件|检索”命令，将弹出如下对话框：

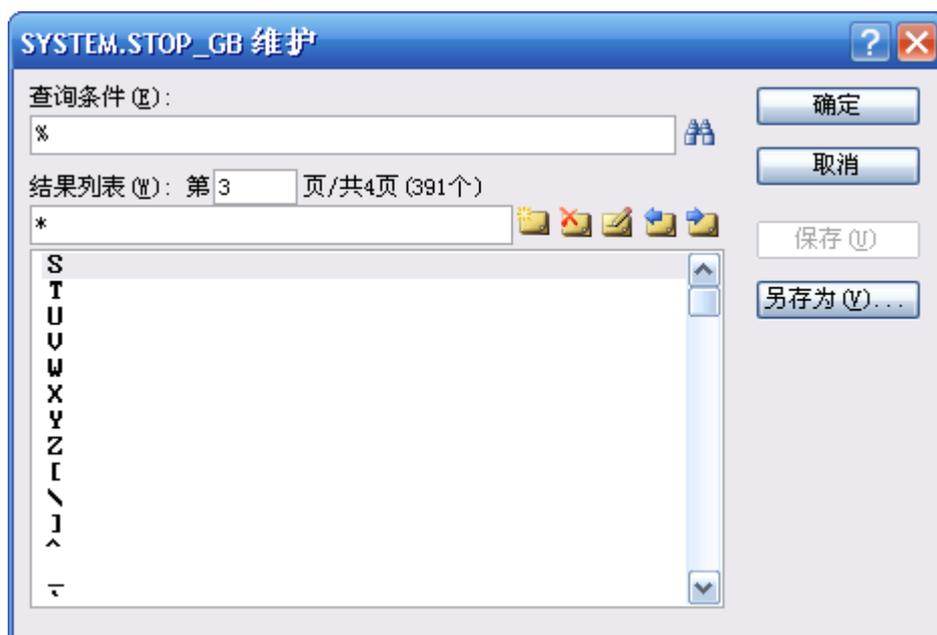


图 4—5 索引词典维护对话框

对话框中各控制的说明如下：

控制名称	功能简介
查询条件	查询表达式。可为“目标词”或“函数名(目标词, ...)”形式，其中“目标词”既可是个确定的词汇或字段值，也可是一个符合TRS规范的模糊词汇或模糊字段值（对分词词典不能使用“%”）。按右边的  按钮进行词典查询，并把结果显示在结果列表里。
结果列表	命中词语的列表。每页显示个数可在“选项 集成环境”里设定，缺省为100个。要看其它页的词语信息，可用右上角的  和  按钮进行前后翻页。按  按钮可把上面输入框内的词语添加到词典里，按  按钮可对当前选中词语进行修改，按  按钮可把当前选中词语从词典里删除，不过添加、修改和删除必须在执行了保存操作以后才会保存到词典里。
当前页面	当前显示的页数。在编辑框中输入想要跳转的页数以后点击回车键，将会跳转到指定的页数。
保存	保存当前所作的修改（包括添加、删除和修改属性操作）。
另存为	打开“词典导出”对话框，把当前查询结果的一部分或全部保存到一个文件里。

修修改完后按[确定]，当前所作的添加、删除、修改操作将会保存到词典里，并关闭对话框。按[关闭]则放弃所作的修改，并关闭对话框。

4.6.2 主题词典的维护

在对象管理窗口里双击主题词典对象，或选中需要查询和维护的主题词典，并按回车或选择“文件|检索”命令，将弹出如下对话框：

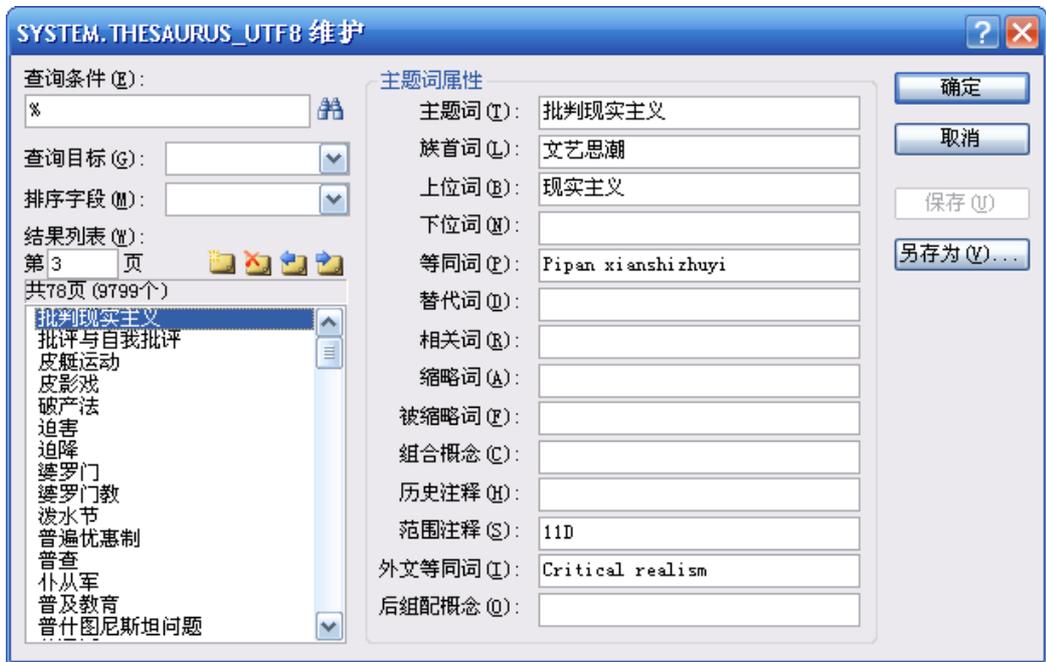


图 4—6 主题词典维护对话框

对话框中各控制的说明如下：

控制名称	功能简介	
查询条件	查询表达式。可为“目标词”或“函数名(目标词,...)”形式，其中“目标词”既可是个确定的词汇或字段值，也可是一个符合TRS规范的模糊词汇或模糊字段值。按右边的  按钮进行词典查询，并把结果显示在结果列表里。	
查询目标	主题词典查询的目标字段，缺省为空，表示按主题词查询。	
排序字段	主题词典查询的排序方式，缺省为空，表示不进行排序。	
结果列表	命中词语的列表。每页显示个数可在“选项 集成环境”里设定，缺省为100个。要看其它页的词语信息，可用右上角的  和  按钮进行前后翻页。按  按钮可向词典里添加新词语，按  可把当前选中词语从词典里删除，不过添加和删除必须在执行了保存操作以后才会保存到词典里。	
当前页面	当前显示的页数。在编辑框中输入想要跳转的页数以后点击回车键，将会跳转到指定的页数。	
主题词属性	主题词	主题词名，是主题词典的关键词。
	族首词	族首词(Lead Term)，概念上外延最广的主题词。如“木刻”的族首词是“美术”。
	上位词	上位词(Broader Term)，概念上外延较广的主题词。如“木刻”的上位词是“版画”。
	下位词	下位词(Narrower Term)，概念上内涵更窄的主题词。如“木刻”的下位词有“水印木刻”和“套色木刻”。

等同词	等同词(Equivalence,Preferred Term), 也称正式主题词或规范主题词。如“电子计算机”是“电脑”的等同词。
替代词	替代词(用代词)(Used For,Non-Preferred Term), 也称非正式主题词或非规范主题词。如“电脑”、“计算机”是“电子计算机”的替代词。
相关词	相关词(Related Term), 概念相近但不是等关系的主题词。如“计算机应用”和“自动控制”是相关关系。
缩略词	缩略词(Abbreviation), 主题词的缩略词。如“Database Management System”的缩略词为“DBMS”。
被缩略词	被所略词(Abbreviation For), 缩略词的正式名称。如“DBMS”的正式名称为“Database Management System”。
组合概念	组合概念(Combined Concepts), 为将来TRS版本预留。
历史注释	历史注释(History Note), 对主题词的历史进行说明。
范围注释	范围注释(Scope Note), 对主题词的适用范围进行说明。
外文等同词	外文等同词(Language Equivalent), 其它语种的等同词“计算机”的英文等同词是“Computer”。
后组配概念	后组配概念(Post-coordinated Concept), 为将来TRS版本预留。
保存	保存当前所作的修改(包括添加、删除和修改属性操作)。
另存为	打开“词典导出”对话框, 把当前查询结果的一部分或全部保存到一个文件里。

修改完毕后按[确定], 当前所作的添加、删除、修改操作将会保存到词典里, 并关闭对话框。按[关闭]则放弃所作的修改, 并关闭对话框。

4.6.3 同义词典的维护

在对象管理窗口里双击同义词典对象, 或选中需要查询和维护的同义词典, 并按回车或选择“文件|检索”命令, 将弹出如下对话框:



图 4—7 同义词典维护对话框

对话框中各控制的说明如下：

控制名称	功能简介
查询条件	查询表达式。可为“目标词”或“函数名(目标词, ...)”形式，其中“目标词”既可是个确定的词汇或字段值，也可是一个符合TRS规范的模糊词汇或模糊字段值。按右边的 按钮进行词典查询，并把结果显示在结果列表里。
结果列表	命中词语的列表。每页显示个数可在“选项集成环境”里设定，缺省为100个。要看其它页的词语信息，可用右上角的 和 按钮进行前后翻页。按 按钮可向词典里添加新词语，按 按钮可把当前选中词语从词典里删除，不过添加和删除必须在执行了保存操作以后才会保存到词典里。
当前页面	当前显示的页数。在编辑框中输入想要跳转的页数以后点击回车键，将会跳转到指定的页数。
同义词属性	当前选中词语的同义词列表。在上面的输入框里输入新的词语，并按右上角的 按钮或按回车键，该词语将被添加到同义词表里。从同义词表里选择要删除的词语，再按 按钮将把该词语从同义词表里删除。
保存	保存当前所作的修改（包括添加、删除和修改属性操作）。
另存为	打开“词典导出”对话框，把当前查询结果的一部分或全部保存到一个文件里。

修改完毕后按[确定]，当前所作的添加、删除、修改操作将会保存到词典里，并关闭对话框。按[关闭]则放弃所作的修改，并关闭对话框。

4.6.4 反义词典的维护

在对象管理窗口里双击反义词典对象，或选中需要查询和维护的反义词典，并按回车或选择“文件|检索”命令，将弹出如下对话框：



图 4—8 反义词典维护对话框

对话框中各控制的说明如下：

控制名称	功能简介
查询条件	查询表达式。可为“目标词”或“函数名(目标词, ...)”形式，其中“目标词”既可是个确定的词汇或字段值，也可是一个符合TRS规范的模糊词汇或模糊字段值。按右边的 按钮进行词典查询，并把结果显示在结果列表里。
结果列表	命中词语的列表。每页显示个数可在“选项 集成环境”里设定，缺省为100个。要看其它页的词语信息，可用右上角的 和 按钮进行前后翻页。按 按钮可向词典里添加新词语，按 可把当前选中词语从词典里删除，不过添加和删除必须在执行了保存操作以后才会保存到词典里。
当前页面	当前显示的页数。在编辑框中输入想要跳转的页数以后点击回车键，将会跳转到指定的页数。
反义词 同义词	当前选中词语的同义词语列表。在上面的输入框里输入新的词语，并按右上角的 按钮或按回车键，该词语将被添加到同义词表里。从同义词表里选择要删除的词语，再按 按钮将该词语从同义词表里删除。

属性	反义词表	当前选中词语的反义词列表。在上面的输入框里输入新的词语，并按右上角的  按钮或按回车键，该词语将被添加到反义词表里。从反义词表里选择要删除的词语，再按  按钮将把该词语从反义词表里删除。
保存		保存当前所作的修改（包括添加、删除和修改属性操作）。
另存为		打开“词典导出”对话框，把当前查询结果的一部分或全部保存到一个文件里。

输入目标文件名，设置词语范围、保存格式等选项以后，按[确定]，TRS 管理员将开始保存查询结果到目标文件里。按[取消]则放弃保存操作并关闭对话框。由于保存查询结果是一个耗时过程，因此保存过程中有“保存进度”对话框显示目前已保存的记录数和保存完成的百分比。如要终止保存过程，可按[终止]按钮。

4.6.5 保存词典查询结果

如果在词典维护窗口里按下[另存为]按钮，将出现如下对话框：



图 4—9 词语导出对话框

对话框中各控制的说明如下：

控制名称	功能简介
目标文件名	保存的目标文件名，可按右边的  按钮来选择目标文件名。
词语范围	需要保存的词语的范围，缺省为所有词语。
当前记录	只保存当前记录。
所有记录	保存检索结果里的所有记录。
指定范围	保存指定范围的记录，可用“起始记录号-终止记录号”的方式表示一定范围的记录，或仅用单个数字表示一条记录，多个分段之间以“;”分隔。
保存格式	结果保存的格式，对索引词典查询无效，缺省为TRS装库数据格式。
记录文件	以记录数据格式保存检索结果，一般只需浏览检索结果时选择此格式。
装库文件	以TRS装库数据格式保存检索结果，所保存的文件可以用来重新进行装库。
ISO2709	以ISO2709国际标准格式保存检索结果。

XML	以XML格式保存检索结果。
-----	---------------

输入目标文件名，设置词语范围、保存格式等选项以后，按[确定]，TRIS 管理员将开始保存查询结果到目标文件里。按[取消]则放弃保存操作并关闭对话框。由于保存查询结果是一个耗时过程，因此保存过程中有“保存进度”对话框显示目前已保存的记录数和保存完成的百分比。如要终止保存过程，可按[终止]按钮。

第5章 使用用户组对象

5.1 TRS 用户组概述

TRS 具有独立于操作系统的用户组管理机制。用户组的设立，使得系统对数据的安全控制更加完善，也更加简单方便。每个 TRS 系统可支持多达 65535 个用户组。

用户组对象具有下列属性：

- 名称

名称是 TRS 用户组对象名，也就是常说的用户组名。用户组名与用户组名之间，以及用户组名与用户名之间不能重复。

- 说明信息

说明信息是对 TRS 用户组对象进行说明的信息。

系统安装时，将自动创建两个用户组：ADMINISTRATOR 和 GUEST。系统还有一个匿名用户组，当一个用户组被删除时，该用户组下的所有用户将自动迁移到这个匿名用户组。

5.2 创建新的用户组

只有 SYSTEM 用户才能创建用户组。进入创建用户组对话框有三种方法。

方法一：选择“文件|新建”命令，在“新建”对话框里双击用户组图标；

方法二：当焦点在导航窗口的“所有用户”节点上时，按工具栏上的新建  按钮；

方法三：当焦点在导航窗口的“所有用户”节点上时，按鼠标右键，出现菜单后选择新建。

通过上述方法之一将会开始执行“创建用户组”，出现下面对话框：



图 5—1 新建用户组对话框

对话框里各控制的说明如下：

控制名称	功能简介
用户组名	新建用户组的名字，不能包含中文、字母、数字或下划线以外的字符，大小写无关。名字长度在1-31个字符之间，不能是TRS保留字，也不能和已存在的用户组/用户重名。
说明信息	不超过63个字符长度的用户组说明信息。
用户列表	新用户组包含的用户列表。
已选用户	已选入新用户组的用户列表。可按  按钮从已选用户列表里清除。此操作将使选中用户脱离当前用户组。
可选用户	没有选入新用户组的用户列表。可按  按钮加到已选用户列表里。此操作将使选中用户脱离原用户组，并加入到当前用户组中。

设定好用户组名、说明信息和用户列表后，按[确定]，TRS 管理员将在当前服务器上建立新的用户组，并把所有已选用户列表里的用户移到新建的用户组里。如想放弃新建用户组，请按[取消]。

5.3 修改 TRS 用户组属性

用户可方便地查看和修改当前选中用户组的属性信息。如要对用户组的属性进行修改，则当前用户必须是 SYSTEM 用户。

用户组的属性在属性窗口中显示，点击要查看属性的用户组对象，其相应的属性便会在下方属性窗口中显示（如下图所示）。若属性窗口已关闭，可选择菜单上的“文件|显示属性”命令，或按工具栏上的按钮重新打开属性窗口。



图 5-2 用户组的属性

“常规信息”属性页里各控制的说明如下：

控制名称	功能简介
内部编号	用户组在系统中的内部标识号。
对象类型	对象的类型，恒为“用户组”。
创建者	对象的创建者，恒为“SYSTEM”。
创建时间	用户组的创建时间。
修改时间	用户组的最后一次修改时间。
说明信息	不超过63个字符长度的用户组说明信息。
用户列表	点击  按钮将弹出“用户列表选择对话框”，对所属于当前用户组的用户进行选择或更改。

属性修改完毕后，按属性窗口上的按钮则保存所做的修改。

5.4 用户组重命名

只有 SYSTEM 用户有权重命名用户组。

先选中要改名的用户组，然后选择“文件”菜单上的“重命名”，或按工具栏上的按钮，或直接按空格键，当前用户组名字的位置上将出现一编辑框；在编辑框中输入新的名字，然后按 ENTER 键，所选用用户组将被重新命名为所输入的名字；如要取消重命名操作，可按 ESC 键。此外，还可以在属性窗口中直接修改，点击用户组，属性窗口将显示其相应的属性，在“对象名称”一栏可直接对词典名进行修改，修改完毕点击工具栏上的按钮即可保存新名字。

5.5 删除 TRS 用户组对象

只有 SYSTEM 用户有权删除用户组：

使用 TRS 管理员工具删除 TRS 用户组对象，首先要打开对象管理的用户组标签，通过四种方法可以完成对用户组对象的删除。

方法一：选择要删除的用户组，按下鼠标的右键并选择“删除”或“剪切”。

方法二：选择要删除的用户组，选择菜单“编辑|删除”或“编辑|剪切”。

方法三：选择要删除的用户组，按下“Delete”键或“Ctrl+X”。

方法四：选择要删除的用户组，选择工具栏中或按钮。

系统会出现如下对话框：



图 5—3 删除用户组确认对话框

选择“是(Y)”删除用户组；选中“删除用户组下的所有用户”单选框表示删除用户组的同时还删除此用户组下的所有用户。选择“否(N)”放弃删除操作。

“删除”和“剪切”的区别在于剪切操作把所选的用户组复制到剪贴板里，而删除则不进行复制。利用“剪切”和“粘贴”功能的配合可快速复制用户组（但不包括用户组里的用户）。

第6章 使用用户对象

6.1 TRS 用户概述

TRS 具有独立于操作系统的用户管理机制。要访问 TRS 系统，必须首先申请一个用户帐号。用户帐号是实现系统及数据安全的主要手段，也是系统计费的主要依据。每个 TRS 系统可支持多达 65535 个用户。

用户对象具有下列属性：

- 用户账号

用户账号是 TRS 用户对象名，也就是常说的用户名。用户名与用户名之间，以及用户名与用户组名之间不能重复。

- 用户密码

用户密码是操作者使用该用户登录到 TRS 系统时必须输入的口令。设置用户密码的目的主要是为了保证用户帐号的私用性。

- 密级

密级定义了该用户在 TRS 系统的接触机密数据的级别。数值越小，级别越高。对于数据库中使用安全级别加密的数据记录，只有当用户的安全级别值不大于记录的安全级别值时，才能看到该记录的内容。

- 用户类型

用户类型实际上是用户在 TRS 系统中的用户权限。TRS 系统有四种用户类型：

- 系统数据库管理员(DBA)
- 用户组数据库管理员(GDA)
- 系统资源用户(RESOURCE)
- 系统登录用户(CONNECT)

- 所属用户组

即该用户所在用户组。用户也可属于匿名用户组。

- **最大检索历史步数**

最大检索历史步数是指用户在进行数据查询时，需要保存的最大检索历史步数。保存的检索历史可以在以后检索中直接引用，而不需要重新检索。

- **最大下载记录数**

最大下载记录数是指用户在进行数据查询时，允许一次下载的最大检索结果记录数。

- **缺省附加分词词典**

该用户在创建物理数据库时，如果没有指定数据库的附加分词词典，则系统自动为数据库指定这里所给出的附加分词词典。

- **缺省附加停用词典**

该用户在创建物理数据库时，如果没有指定数据库的附加停用词典，则系统自动为数据库指定这里所给出的附加停用词典。

- **缺省稀疏词典**

该用户在创建物理数据库时，如果没有指定数据库的稀疏词典，则系统自动为数据库指定这里所给出的稀疏词典。

- **说明信息**

说明信息是对 TRS 用户对象进行说明的信息。

- **登录互斥开关决定**

登录互斥开关决定该用户是否允许同时从多个不同的 IP 地址进行登录，但来自本机 IP 地址的用户登录不计算在内。

系统安装时，将自动创建两个用户：SYSTEM 和 PUB。

SYSTEM 是 TRS 系统的超级用户，其主要特征有：

- 拥有系统数据库管理员权限。
- 属于 ADMINISTRATOR 用户组。
- 不能被注销。
- 不能修改名称。
- 可由自己修改口令。
- 可创建和注销其它任何用户。
- 可清除任何用户的口令。

- 可修改任何用户的除口令以外的任何属性。
- 可授予或收回任何用户的任何权限。

PUB 是 TRS 系统的一个公共用户，其主要特征有：

- 拥有系统登录权限。
- 属于 GUEST 用户组。
- 不能修改名称。
- 没有口令，也不允许设置口令。
- 可被 SYSTEM 用户注销。
- 对 TRS 系统数据对象的操作，必须通过相关的权限检查。

TRS 用户还有其它一些特点：

- 每一个用户组中允许有多个 DBA 用户。
- 每一个用户组中允许有多个 GDA 用户。

任何用户可修改“自己”的如下属性：

- 登录口令
- 最大检索历史步数
- 缺省附加分词词典
- 缺省附加停用词典
- 缺省稀疏词典
- 说明信息
- 唯一登录

其它项只能由 SYSTEM 用户管理。

当一个用户被删除时，该用户所拥有的所有数据库、用户组，以及词典对象的拥有者将改为 SYSTEM 用户，即这些对象自动归为 SYSTEM 用户所拥有。

6.2 创建新的用户

只有 SYSTEM 用户才能创建用户。进入创建用户对话框有三种方法。

方法一：选择“文件|新建”命令，在“新建”对话框里双击用户图标；

方法二：当焦点在导航窗口除根节点以外的其他节点上时，按工具栏上的新建  按钮；

方法三：当焦点在导航窗口除根节点以外的其他节点上时，按鼠标右键，出现菜单后选择新建。

通过上述方法之一将会开始执行“创建用户”，出现下面对话框：



图 6—1 创建新用户对话框

对话框里各控制的说明如下：

控制名称	功能简介
用户帐号	新建用户的帐号名，不能包含中文、字母、数字或下划线以外的字符，大小写无关。名字长度在1-31个字符之间，不能是TRS保留字，也不能和已存在的用户组/用户重名。
用户组名	用户所在的组名，可以是已存在的用户组之一，也可以是空，表示不属于任何用户组。
用户类型	用户的类型，有四种选择：全域管理员用户(DBA)、组管理员用户(GDA)、资源用户(Resource)、登录用户(Connect)。
说明信息	不超过63个字符长度的用户说明信息。
设置密码	新建用户的口令。为保密起见，输入字符以“*”显示。密码的有效长度是1-31个字符。若此控制项为空，即表示没有密码。
确认密码	确认用户的密码，必须与设置密码一致。为保密起见，输入字符以“*”显示。
附加分词	用户缺省的附加分词词典名。创建该用户的数据库时，如不指定附加分词词典，则自动取此词典作为数据库的附加分词词典。缺省值为空。
附加停用	用户缺省的附加停用词典名。创建该用户的数据库时，如不指定附加停用词典，则自动取此词典作为数据库的附加停用分词词典。缺省值为空。

控制名称	功能简介
稀疏词典	用户缺省的稀疏词典名。创建该用户的数据库时，如不指定稀疏词典，则自动取此词典作为数据库的稀疏词典。缺省值为空。
最大检索历史步数	用户的最大检索历史步数，清空或为零时取缺省值，即系统配置中设定的最大检索历史步数。
最大下载记录数	用户的最大下载记录数，清空或为零时取缺省值，即系统配置中设定的最大下载记录数。
密级	密级定义了该用户在TRS系统的接触机密数据的级别。数值越小，级别越高。对于数据库中使用安全级别加密的数据记录，只有当用户的安全级别值不大于记录的安全级别值时，才能看到该记录的内容。
登录互斥开关	选中表示不允许该用户同时从多个不同的IP地址进行登录，但来自本机IP地址的用户登录不计算在内。缺省为清除状态。

设定好用户组名和用户的各项属性后，按[确定]，TRS 管理员将在当前服务器上建立新的用户。如想放弃新建用户，请按[取消]。

6.3 修改 TRS 用户属性

用户可方便地查看和修改当前选中用户的属性信息。如要对用户的属性进行修改，则当前用户必须是 SYSTEM 用户，或为选中用户本身。

用户的属性在属性窗口中显示，点击要查看属性的用户对象，其相应的属性便会在下方属性窗口中显示（如下图所示）。若属性窗口已关闭，可选择菜单上的“文件|显示属性”

命令，或按工具栏上的  按钮重新打开属性窗口：

常规属性	
内部编号	126
对象名称	beijing
所在组	Guest
用户类型	系统登录用户 (CONNECT)
创建时间	2008.03.25 17:35:48
修改时间	2010.05.04 17:52:11
说明信息	
高级属性	
设置密码	*****
确认密码	*****
附加分词词典	SYSTEM.MYSEGEI
附加停用词典	
稀疏词典	SYSTEM.MYSPARSEDIC
最大检索历史步数	10
最大下载记录数	10000
密级	0
登录互斥	<input checked="" type="checkbox"/>

图 6—2 用户的属性

“常规信息”属性页里各控制的说明如下：

控制名称	功能简介	
常规属性	内部编号	用户在系统中的内部标识号。
	对象名称	名称是TRS用户对象名，即用户帐号名，也就是用户名。用户名与用户名之间，以及用户名与用户组名之间不能重复。
	所在组	用户所在的组名，可以是已存在的用户组之一，也可以是空，表示不属于任何用户组。
	对象类型	用户的类型，有四种取值：系统数据库管理员(DBA)、用户组数据库管理员(GDA)、系统资源用户(Resource)、系统登录用户(Connect)。
	说明信息	不超过63个字符长度的用户说明信息。
	创建时间	用户的创建时间。
	修改时间	用户的最后一次修改时间。
高级属性	设置密码	用户的口令。为保密起见，输入字符以“*”显示。密码有效长度是1-31个字符。此控制项初始化为31个“*”，表示保持原密码。若此控制项为空，表示没有密码。
	确认密码	确认用户的密码，必须与设置密码一致。为保密起见，输入字符以“*”显示。
	附加分词词典	用户缺省的附加分词词典名。创建该用户的数据库时，如不指定附加分词词典，则自动取此词典作为数据库的附加分词词典。
	附加停用词典	用户缺省的附加停用词典名。创建该用户的数据库时，如不指定附加停用词典，则自动取此词典作为数据库的附加停用分词词典。
	稀疏词典	用户缺省的稀疏词典名。创建该用户的数据库时，如不指定稀疏词典，则自动取此词典作为数据库的稀疏词典。
	最大检索历史步数	用户的最大检索历史步数。
最大下载记	用户的最大下载记录数。	

录数	
密级	密级定义了该用户在TRS系统的接触机密数据的级别。数值越小，级别越高。对于数据库中使用安全级别加密的数据记录，只有当用户的安全级别值不大于记录的安全级别值时，才能看到该记录的内容。
登录互斥开关	选中表示不允许该用户同时从多个不同的IP地址进行登录，但来自本机IP地址的用户登录不计算在内。

如对用户属性作了修改，按  将保存所作的修改。

6.4 用户重命名

只有 SYSTEM 用户有权重命名用户：

先选中要改名的用户，然后选择“文件”菜单上的“重命名”，或按工具栏上的  按钮，或直接按空格键，当前用户名字的位置上将出现一编辑框；在编辑框中输入新的名字，然后按 ENTER 键，所选用户将被重新命名为所输入的名字；如要取消重命名操作，可按 ESC 键。此外，还可以在属性窗口中直接修改，点击用户，属性窗口将显示其相应的属性，在“对象名称”一栏可直接对词典名进行修改，修改完毕点击工具栏上的  按钮即可保存新名字。

6.5 删除 TRS 用户对象

只有 SYSTEM 用户有权删除用户：

使用 TRS 管理员工具删除 TRS 用户对象，首先要打开对象管理的用户标签，通过四种方法可以完成对用户对象的删除。

方法一：选择要删除的用户，按下鼠标的右键并选择“删除”或“剪切”。

方法二：选择要删除的用户，选择菜单“编辑|删除”或“编辑|剪切”。

方法三：选择要删除的用户，按下“Delete”键或“Ctrl+X”。

方法四：选择要删除的用户，选择工具栏中  或  按钮。

系统会出现如下对话框：



图 6—3 删除用户确认对话框

选择“是(Y)”删除用户；选择“否(N)”放弃删除操作。

“删除”和“剪切”的区别在于剪切操作把所选的用户复制到剪贴板里，而删除则不进行复制。利用“剪切”和“粘贴”功能的配合可快速复制用户。

第7章 管理数据格式

7.1 TRS 数据格式对象概述

数据格式用来说明多媒体数据的格式类型。在数据库加载时，TRS 系统将根据数据格式确定相应的数据存储方式，使用相应的分词方法；在数据输出时，TRS 系统将根据数据格式对数据进行相应的处理，便于浏览。每个 TRS 系统可管理多达 1023 个数据格式对象。

数据格式对象具有下列属性：

- **格式名称**

格式名称是 TRS 数据格式的对象名，是访问、维护和管理该数据格式的入口。数据格式名不能重复。

- **所属分类**

即该数据格式属于哪种格式类型。TRS 系统把数据格式划分成六种类型：

- 纯文本(TEXT)
- 格式文档(DOCS)
- 图像(IMAGE)
- 音频(AUDIO)
- 视频(VIDEO)
- 自定义(BIT)

- **文件后缀**

文件后缀是指该数据格式的数据文件通常使用的文件名后缀。

- **压缩开关**

压缩开关决定该数据格式的数据是否需要进行无损压缩。

- **说明信息**

说明信息是对数据格式对象进行说明的信息。

TRS 系统在安装时，将自动创建下列数据格式对象(名称):

- BIT (不确定格式的二进制数据)
- TEXT (文本数据格式)
- WORD (Microsoft Word 数据格式)
- WS (WordStar 数据格式)
- WP (WordProfect 数据格式)
- XCL (Microsoft Excel 数据格式)
- WPS (金山 WPS 数据格式)
- S2 (方正、华光二扫数据格式)
- DOCS (不确定格式的文档数据格式)
- AUDIO (不确定格式的音频数据格式)
- IMAGE (通用图像数据格式)
- VIDEO (不确定格式的视频数据格式)
- AVI (AVI 动画数据格式)
- MPEG (MPEG 压缩动画数据格式)
- WAVE (Wave 音频数据格式)
- MIDI (MIDI 音频数据格式)
- CDA (CD 音频数据格式)
- PS (PostScript 数据格式)
- GIF (GIF 图像数据格式)
- TIF (TIF 图像数据格式)
- PCX (PCX 图像数据格式)
- BMP (BMP 图像数据格式)
- JPEG (JPEG 图像数据格式)
- HTML (HTML 超文本数据格式)
- PDF (Adobe PDF 数据格式)
- PPT (Microsoft Power Point 数据格式)
- RTF (Rich Text 数据格式)
- XML (XML 数据格式)
- MP3 (MP3 数据格式)
- MP4 (MP4 文件格式)
- ZIP (ZIP 压缩文件格式)
- RAR (RAR 压缩文件格式)
- TAR (TAR 包文件格式)

TRS 系统允许用户自己定义新的数据格式，或者修改已有的数据格式。数据格式一旦定义，就不允许删除，所以必须谨慎对待。只有 SYSTEM 用户有权创建新的数据格式。

7.2 创建新的数据格式

方法一：选择“文件|新建”命令，在“新建”对话框里双击数据格式图标；

方法二：当对象管理窗口显示数据格式列表且焦点也在其上时，按工具栏上的新建按钮；

方法三：当对象管理窗口显示数据格式列表且焦点也在其上时，在对象管理框按鼠标右键，出现菜单后选择新建。

通过上述方法之一将会开始执行“创建数据格式”，出现下面对话框：



图 7—1 创建数据格式对话框

对话框里各控制的说明如下：

控制名称	功能简介
格式名称	新建数据格式的名字，不能包含英文字母或数字以外的字符，大小写无关。名字长度在1-15个字符之间，不能是TRS保留字，也不能和已存在的数据格式重名。
所属分类	数据格式的分类名，可以从已存在的分类名中选择，或置为空。
后缀列表	数据格式所对应的文件的后缀名列表，多个后缀名间以半角分号、逗号或空格分隔。其长度不超过31个字符，缺省值为空。
说明信息	不超过63个字符长度的数据格式说明信息。
无损压缩	选中表示对该格式的文件进行无损压缩。缺省为选中。

设定好格式名称、所属分类、后缀列表、说明信息等属性后，按[确定]，TRS 管理员将在当前服务器上建立新的数据格式。如想放弃新建数据格式，请按[取消]。

注意：由于创建后的数据格式不能被删除，请在创建前考虑清楚是否应该创建该数据格式。

7.3 修改数据格式的属性

用户可方便地查看和修改当前选中数据格式的属性信息。如要对数据格式的属性进行

修改，则当前用户必须是 SYSTEM 用户。

数据格式的属性在属性窗口中显示，点击要查看属性的数据格式对象，其相应的属性便会在下方属性窗口中显示（如下图所示）。若属性窗口已关闭，可选择菜单上的“文件|显示属性”命令，或按工具栏上的按钮重新打开属性窗口。

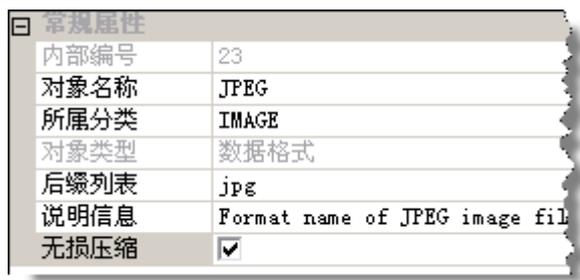


图 7-2 数据格式属性

“常规信息”属性页里各控制的说明如下：

控制名称	功能简介
内部编号	数据格式在系统中的内部标识号。
对象名称	数据格式的对象名，即数据格式名，是访问、维护和管理该数据格式的入口。数据格式名不能重复，并且 只能由英文字母和数字组成 。
所属分类	数据格式的分类名，可以从已存在的分类名中选择，或置为空。分类名长度不能超过15个字符。
对象类型	对象的类型，恒为“数据格式”。
后缀列表	据格式所对应的文件的后缀名列表，多个后缀名间以半角分号、逗号或空格分隔， 只能由英文字母和数字组成 。
说明信息	不超过63个字符长度的数据格式说明信息。
无损压缩	选中表示对该格式的文件进行无损压缩。

如对数据格式属性作了修改，按将保存所作的修改。

7.4 重命名数据格式

只有 SYSTEM 用户有权重命名数据格式。

先选中要改名的数据格式，然后选择“文件”菜单上的“重命名”，或按工具栏上的按钮，或直接按空格键，当前数据格式名字的位置上将出现一编辑框；在编辑框中输入新的名字，然后按 Enter 键，所选数据格式将被重新命名为所输入的名字；如要取消重命名操作，可按 Esc 键。此外，还可以在属性窗口中直接修改，点击数据格式，属性窗口将显示其相应的属性，在“对象名称”一栏可直接对词典名进行修改，修改完毕点击工具栏上的按钮即可保存新名字。

第8章 TRS 管理员的设置

TRs 管理员包含很多选项，用户可以通过改变这些选项达到自己定制 TRs 的目的。对选项的修改将保存到服务器端或本地计算机，对以后启动时同样有效。

选择“工具|选项”命令后，TRs 管理员弹出“选项”对话框。该对话框共有四页：缺省登录、集成环境、系统配置和知识词典。其中前两页为 TRs 管理员的集成环境选项，后两页为当前连接的服务器的 x 系统配置，只有 SYSTEM 用户才能修改系统配置。如当前没有登录到服务器，则对话框将只含第一页。

单击对话框左部图片框中的图标可分别查看各页面的具体选项的设置。如对选项作了修改，按[确定]按钮可保存所作的修改并关闭对话框，按[应用]按钮可保存所作的修改而不关闭对话框，按[取消]按钮则放弃所作的修改并关闭对话框。

8.1 缺省登录

“缺省登录”属性页如下所示：

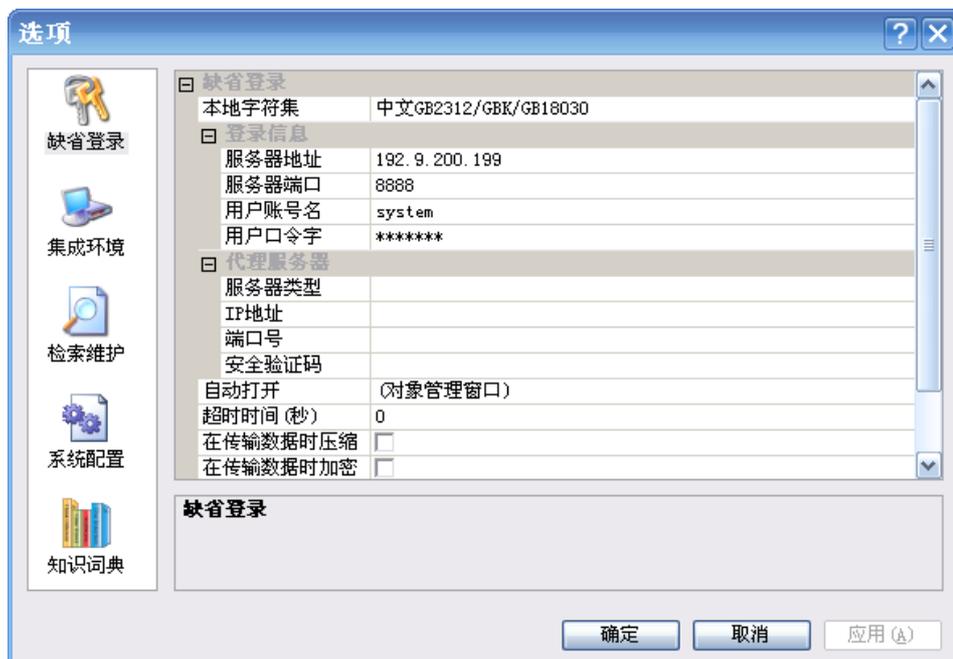


图 8—1 选项的缺省登录属性页

属性页中各选项的含义如下：

控制名称	功能简介	
本地字符集	本地所用的字符集，可为“中文GB2312/GBK/GB18030”、“中文BIG5”、“统一码UTF8”或“不设置（与原来兼容）”，缺省为第一个选项。	
登录信息	服务器地址	服务器所在的主机，既可以是TRS服务器在网络中的IP地址（如“192.9.200.80”），也可以是服务器所在主机的名称（如“ALPHA”）。如果输入的是服务器主机的名称则要求该主机名是当前所使用的网络驱动系统的数据库文件hosts的一个入口项（如“192.9.200.80 alpha”）。服务器地址可以从TRS的系统管理员处获得。
	服务器端口	服务器端口既可以是TRS服务器的INTERNET端口（如“8000”），也可以是TRS服务器的名称（如“trserver”）。如果输入的是服务器的名称，则要求该名称是当前所使用的网络驱动系统的数据库文件services的一个TCP协议入口项（如“trserver 8000/tcp”）。服务器端口可以从TRS的系统管理员处获得。
	用户帐号名	用户的帐号名，必须是指定服务器中已经存在的用户名。用户的帐号由SYSTEM用户创建。
	用户口令字	用户的密码，区分大小写。为防止泄密，输入字符将用“*”号显示。用户的密码可由用户自己修改。
代理服务	服务器类型	不为空表示将通过代理连接到目的服务器，目前仅支持“TRS”的代理服务器。
	IP地址	如果服务器类型不为空，则此属性必须指定代理服务器的有效IP地址。
	端口号	如果服务器类型不为空，则此属性必须指定代理服务器的有效端口号。
	安全验证码	如果服务器类型不为空，则此属性可以指定一个代理服务器所需要的安全验证码。
自动打开	登录成功后自动打开的窗口，可为以下选项之一：对象管理窗口、数据库/视图。	
超时时间	设置登录时的超时时间，以秒为单位，超过指定时间仍未连接上则报错。为零表示使用缺省值（为10秒）。	
在传输数据时压缩	选中表示在传输数据时进行压缩，压缩算法由内部确定，对于本地客户端无效。	
在传输数据时加密	选中表示在传输数据时进行加密，加密算法由内部确定，对于本地客户端无效。	
自动显示对话框	启动时是否显示“启动”对话框标志。如清除该选项，则下次启动时将不再显示“启动”对话框。	
最大化窗口	是否最大化子窗口标志。如设置该选项，则所有子窗口都将占满整个主窗口的空间。	

8.2 集成环境

“集成环境”属性页如下所示：

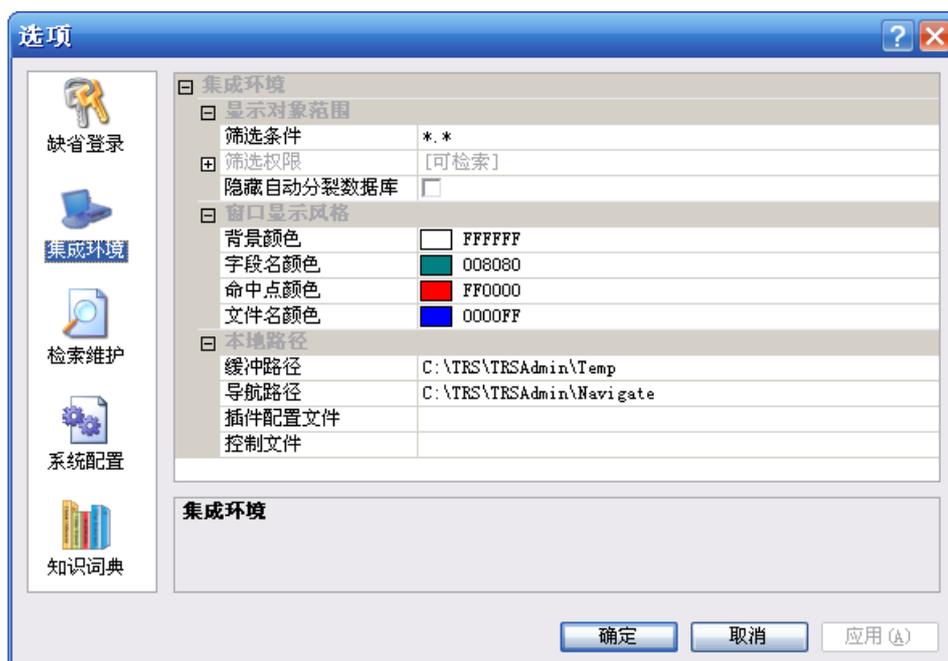


图 8—2 选项的集成环境属性页

属性页中各选项的含义如下：

控制名称		功能简介	
显示对象范围	筛选条件	显示对象的筛选条件，可使用通配符：用“?”匹配单个字符，用“*”匹配任何字符。缺省为“*.*”，表示所有对象。	
	筛选权限	可检索的	如选中表示只显示可以检索的对象，仅对数据库和视图有效。
		可更新的	如选中表示只显示可以更新的对象，仅对数据库、自动模式视图和词典有效。
		可索引的	如选中表示只显示可以建立索引的对象，仅对数据库和自动模式视图有效。
		可修改的	如选中表示只显示可以修改结构的对象，仅对数据库和视图有效。
		可删除的	如选中表示只显示可以删除的对象，仅对数据库、视图和词典有效。
隐藏自动分裂数据库	如选中表示不显示自动分裂视图对应的物理数据库。		
窗口显示风格	背景颜色	概览/细览窗口的背景颜色，按下右边的下拉按钮可打开“颜色”对话框以选择背景颜色。	
	字段名色	细览窗口的字段标题显示颜色，按下右边的下拉按钮可打开“颜色”对话框以选择字段名显示颜色。	
	命中点色	细览窗口的命中点显示颜色，按下右边的下拉按钮可打开“颜色”对话框以选择命中点显示颜色。	
	文件名色	细览窗口的文件名显示颜色，按下右边的下拉按钮可打开“颜色”对话框以选择文件名显示颜色。	
本地	缓冲路径	TRS管理员存放检索结果源文件的临时路径，按右边的...按钮可选择缓冲文件存放的目录，缺省为TRS管理员安装路径下的Temp子目录。	

控制名称		功能简介
路径	导航路径	TRS管理员存放导航文件的路径,按右边的...按钮可导航文件存放的目录,缺省为TRS管理员安装路径下的Navigate子目录。
	插件配置	插件配置文件名,文件的格式参见《TRS全文数据库系统 用户手册》。缺省为空,按右边的...按钮可选择新的插件配置文件。
	控制文件	数据加载的控制文件,用于添加/更新记录中(因集成环境选项里已有设置,其中的多值分隔符不起作用),以及作为装入记录的缺省控制文件。缺省为空,表示取系统缺省值。按...按钮可选择装库控制文件。

8.3 检索维护

“检索维护”属性页如下所示:

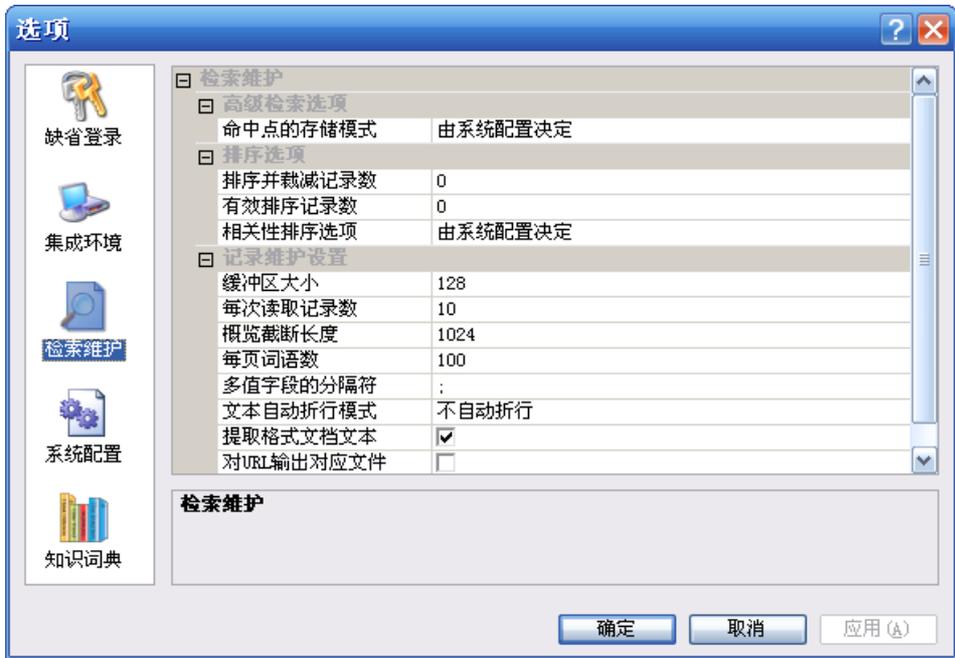


图 8—3 选项的检索维护属性页

属性页中各选项的含义如下:

控制名称		功能简介
高级检索选项	命中点的存储模式	不同的命中点存储模式,不但影响检索速度,还会影响相关性排序结果。
	估算分段单位	使用估算检索模式,其值为数据的分段单位(记录数)。
	估算终止条件	使用估算检索模式,其值为检索终止条件(记录数)。
	标题加权系数(倍)	命中点在标题(第0段/第0句)位置上的加权系数(倍数)。

控制名称		功能简介
	知识扩展权重 (%)	知识扩展检索(主题/同义/反义)词的权重系数(百分比)。
	命中分支权重	命中分支数的权重系数。
	检索超时时间 (秒)	检索的超时时间(秒数, 不能大于系统的的超时时间设置)。如果值为0, 表示使用系统的的超时时间设置。
	中断时输出部分结果	选中表示当检索过程被中断时也输出已经得到的部分检索结果, 否则按出错处理。
	忽略不存在的索引属性	对于不存在的索引属性, 简单地忽略, 不按出错处理。
	集群裁减设置	集群对后台检索结果的裁减设置。0表示不作限制; 1~9表示集群至少需要等待相应数量的后台返回结果; 10~99表示集群至少需要等待相应比例的后台返回结果; 100及以上表示至少需要返回的结果记录数。
排序选项	排序并裁减记录数	存储之前完成排序, 并只存储排在前面的部分记录, 同时自动设置“非精确命中点”模式。0表示由系统配置决定, 缺省是不进行裁减。如果值等于65535, 则表示忽略系统配置, 关闭排序裁减功能。
	有效排序记录数	即部分排序, 只保证前一部分记录是有序的。0表示由系统配置决定, 缺省是全部排序。如果值等于65535, 则表示则表示忽略系统配置, 关闭部分排序功能。只有在排序裁减记录数为0的情况下才可能有效。
	相关性排序选项	不同的相关性的计算模式和IDF模式组合下会有不同的排序结果。分为7种情况: 由系统配置决定; 忽略词的位置和使用基于单个数据库范围的IDF值; 忽略词的位置和词IDF值不参与相关性计算; 考虑词的位置和使用基于单个数据库范围的IDF值; 考虑词的位置和词IDF值不参与相关性计算; 用一个记录中命中词的单位向量长度(每个命中词只计算一次)作为记录的相关; 用一个记录中命中词的单位向量长度, 以及命中词的词频和作为记录的相关度。
	加权表示优先顺序	相关性排序时表达式中多字段的加权表示优先顺序。
	中文排序方式	中文的排序方式, 可选择按内码、拼音、多音字拼音或笔画数排序。
记录维护设置	缓冲区大小	检索结果缓冲区大小。缓冲区大小以记录为单位, 缺省为128。
	每次读取记录数	读取检索结果时每次读取的记录数, 不能大于缓冲区的大小, 仅对取概览内容时有效, 缺省为16。
	概览截断长度	PHRASE、DOCUMENT和BIT字段的截断长度, 为零表示不进行截断。缺省为1024。
	每页词语数	查询数据库的索引词和词典的时候, 每页显示的词语数, 其数值不能大于上面设置的缓冲大小, 缺省值为100。
	多值字段的分隔符	多值字段的分隔符, 仅对DATE、NUMBER和CHAR类型字段有效。为空串表示取系统缺省值 ‘;’。
	文本自动折行模式	只对DOCUMENT和PHRASE字段有效, 其中“所有文本不分行”选项一般用于细览字段以便于控制其显示格式, “段落文本不分行”一般用于细览字段以便于自动换行。
	提取格式文档文本	只对DOCUMENT字段有效, 表示只存储了单个格式化文档字段值的字段, 尽量提取文本并给出命中点。
	对URL输出对应文件	选中表示对URL(统一资源定位)输出其对应的文件(如HTML等), 而不是输出URL本身。缺省值为清除。
	输出原始的相关度值	输出记录的原始(绝对)相关度值, 而不是缺省方式的相对相关度值。

8.4 系统配置

“系统配置”属性页如下所示：

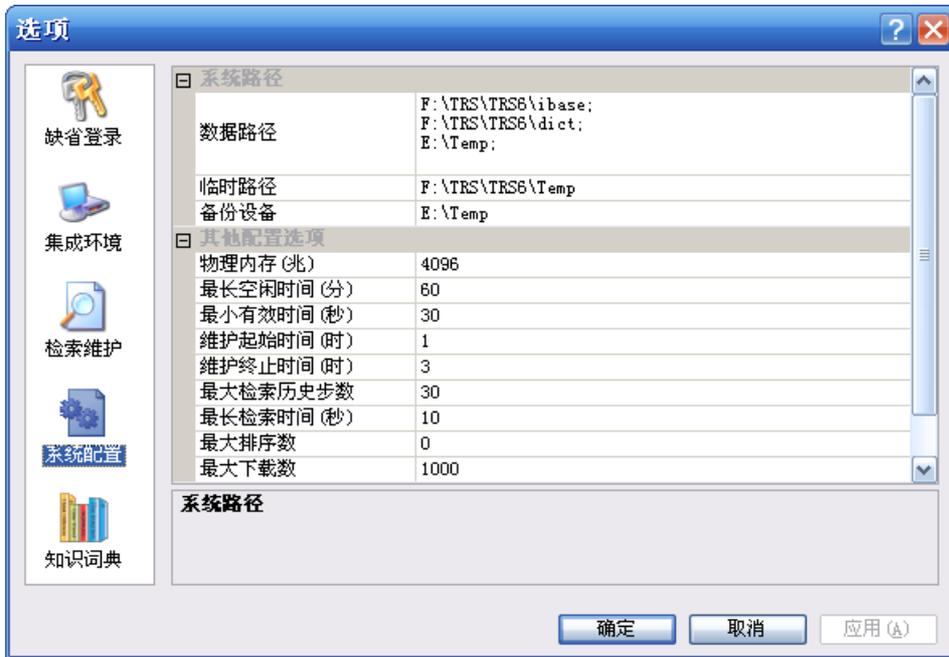


图 8—4 选项的系统配置属性页

属性页中各选项的含义如下：

控制名称	功能简介	
系统路径	数据路径	允许存放数据库数据和索引文件的目录集，至少必须包含一个目录路径。缺省为TRS安装目录下的ibase和dict目录。创建数据库或索引时如不指定路径，则缺省为目录集的第一个目录。当登录本地服务器时，可点击...按钮选择本地路径，自己编辑时可按Ctrl+Enter输入回车。
	临时路径	系统缺省的临时文件路径。缺省为TRS安装目录下的temp目录。当登录本地服务器时，可点击...按钮选择本地路径。
	备份设备	系统缺省的备份设备，可以是磁带名或文件名。缺省为TRS安装目录下的temp目录。当登录本地服务器时，可点击...按钮选择本地路径。
其他配置选项	物理内存	服务器的物理内存总量，以兆字节计算。缺省值根据服务器端配置决定。
	最长空闲时间	允许一个用户实例最长的空闲时间，以分计，超时将自动Logout。为零表示不限制，缺省为30分钟。
	最小有效时间	一个用户应该保持的最小有效时间(除非退出登录)，以秒计，在此期间内不允许自动退出登录。为零表示不限制，缺省为10秒。
	维护起始时间	当系统空闲时，允许自动进行系统或数据库维护的起始时间，以小时计，在0到23之间。24表示禁止系统进行自动维护，缺省值为1。
	维护终止时间	允许自动进行系统或数据库维护的截止时间，以小时计，在0到23之间。

控制名称	功能简介	
	缺省值为3。	
最大检索历史步数	系统缺省的最大检索历史步数，缺省值为20。创建用户时作为用户最大检索历史步数的缺省值。	
最长检索时间	一次检索操作的最长执行时间，以秒计算，零表示不限制，缺省值为10。	
最大排序数	对检索结果进行排序的最大记录数，为零表示不限制，缺省值为10000。	
最大下载数	系统缺省的一次最大下载记录数，缺省值为1000。创建用户时作为用户最大下载记录数的缺省值。	
缺省建库选项	词根索引	系统缺省的只建立词根索引标志，缺省为清除。启动Admin后第一次新建数据库时若没有指定此标志，则自动取配置的值。其后在没有重新启动Admin以前再新建数据库的话，则自动继承上一次新建时的设置，下同。
	大小写敏感	系统缺省的大小写敏感标志，缺省为清除。
	忽略数字	系统缺省的忽略数字串标志，缺省为清除。

8.5 知识词典

“知识词典”属性页如下所示：



图 8—5 选项的知识词典属性页

属性页中各选项的含义如下：

控制名称	功能简介
语言环境	系统缺省的语言环境，有四种取值：中文GB2312/GBK/GB18030、中文BIG5、英语、统一码UTF8，缺省为选项一。

控制名称		功能简介
知识词典集	分词词典	当前语言对应的缺省主分词词典名，建库时作为缺省的数据库主分词词典。缺省为系统安装时建立的分词词典。
	停用词典	当前语言对应的缺省的停用词典名，建库时作为缺省的数据库停用词典。缺省为系统安装时建立的停用词典。
	主题词典	当前语言对应的缺省的主题词典名，检索时作为缺省的自动扩展主题词典。缺省为系统安装时建立的主题词典。
	同义词典	当前语言对应的缺省的同义词典名，检索时作为缺省的自动扩展同义词典。缺省为系统安装时建立的同义词典。
	反义词典	当前语言对应的缺省的反义词典名，检索时作为缺省的自动扩展反义词典。缺省为系统安装时建立的反义词典。

附录 A 检索表达式构造

数据库的检索工具多种多样，这里说明的不是这些应用工具的使用方法，而是说明 TRS 对数据库检索的支持，如运算符、函数、表达式以及键值的构成等。

A.1 运算符

TRS 的运算符有 7 类，分别是：算术运算符、比较运算符、逻辑运算符、重复逻辑运算符、属性运算符、限定运算符和加权运算符。算术运算符只能在统计运算表达式中使用，而其他运算符只能在检索表达式中使用。所有的运算符都不能作用于 BIT 字段，运算符不区分大小写。

1、算术运算符 (作用于 NUMBER 字段)

符号	含义和作用	举例
*	数值型数据相乘	sum(salary) * 0.10
/	数值型数据相除	sum(salary) / 20
%	数值型数据求余数	sum(salary) % 20
+	数值型数据相加	avg(age) + 10
-	数值型数据相减或求负数	avg(age) - 10

2、比较运算符

符号	含义和作用	举例
=	等于	作者=李力
!=	不等于	版面!=国际
>	大于	工资>1000
<	小于	工资<1000
>=	大于等于	工资>=1000
<=	小于等于	工资<=1000

- 比较运算符作用于 NUMBER、DATE、CHAR 字段。但 '=' 和 '!=' 可作用于 PHRASE 和 DOCUMENT 字段。
- 比较运算符的左值必须是一个的检索入口（字段名或其别名）。
- 只有当字段类型为 DATE、NUMBER 和 CHAR 时，比较运算符的右值才可以是一个值列表，并

且只有比较运算符 ‘=’ 和 ‘!=’ 支持右值为一个值列表。

- 如果检索条件不是单个词，而是一个表达式，则该表达式应该用圆括号 “(” 和 “)” 括起来。

3、逻辑运算符

符号	含义和作用	举例
and	两者必须同时满足	中国 and 美国
xor	两者只能满足其一	中国 xor 美国
not	两者中只能出现前者	中国 not 美国
or	两者至少满足其一	中国 or 美国

- 逻辑运算符可作用于除BIT之外的所有字段。
- 逻辑运算符不仅能够连接两个检索词，也可以连接两个检索子表达式，从而构造出非常复杂的检索表达式。

4、重复逻辑运算符

符号	含义和作用	举例
*=	检索各个字段上满足同一条件的记录	c1,c2 *= (中国 * 香港)
^=	检索只有一个字段上满足条件的记录	c1,c2 ^= (中国 * 香港)
-=	检索只有第一个字段满足条件的记录	c1,c2 -= (中国 * 香港)
+=	检索任意字段上同时满足条件的记录	c1,c2 += (中国 * 香港)

- 重复逻辑运算符实际上是在各个指定的字段上用相同的条件式进行检索后，再在各个字段之间进行指定的逻辑运算，可作用于除BIT之外的所有字段。
- 重复逻辑运算符的左值必须是由单字节的逗号 ‘,’ 分隔的两个或两个以上的检索入口。如果右值不是单个词，而是一个表达式，则该表达式应该用圆括号 “(” 和 “)” 括起来。

5、属性运算符

符号	含义和作用	举例
adj	两者邻接，前者先出现	中国 adj 香港
equ/n	两者正好相距n个字先后同时出现	中国 equ/10 香港
pre/n	两者相距至多n个字先后同时出现	中国 pre/10 香港
pre/n#	两者正好相距n个字先后同时出现	中国 pre/10# 香港
pre/nL	两者相距至少n个字先后同时出现	中国 pre/10L 香港
pre/pos=n	两者正好相距n个字先后同时出现	中国 pre/pos=10 香港
pre/sen	两者在同一句中先后同时出现	中国 pre/sen 香港
pre/sen=n	两者在指定的句中先后同时出现	中国 pre/sen=1 香港
pre/seg	两者在同一段中先后同时出现	中国 pre/seg 香港
pre/seg=n	两者在指定的段中先后同时出现	中国 pre/seg=1 香港
pre/tit	两者在标题中先后同时出现	中国 pre/tit 香港
pre/txt	两者在正文中先后同时出现	中国 pre/txt 香港

pre/att	两者在指定属性的节点先后同时出现	中国 pre/文 香港
pre/att=v	两者在指定属性值的节点先后同时出现	中国 pre/章=2 香港
pre/att\$sub=v	两者在指定特征值的节点先后同时出现	中国 pre/文\$色=红 香港
and/n	两者相距至多n个字同时出现	中国 and/10 香港
and/n#	两者正好相距n个字同时出现	中国 and/10# 香港
and/nL	两者相距至少n个字同时出现	中国 and/10L 香港
and/pos=n	两者正好相距n个字同时出现	中国 and/pos=10 香港
and/sen	在同一句中两者同时出现	中国 and/sen 香港
and/sen=n	在指定的句中两者同时出现	中国 and/sen=2 香港
and/seg	在同一段中两者同时出现	中国 and/seg 香港
and/seg=n	在指定的段中两者同时出现	中国 and/seg=2 香港
and/tit	在标题中两者同时出现	中国 and/tit 香港
and/txt	在正文中两者同时出现	中国 and/txt 香港
and/att	在指定属性的节点两者同时出现	中国 and/文 香港
and/att=v	在指定属性值的节点两者同时出现	中国 and/章=2 香港
and/att\$sub=v	在指定特征值的节点两者同时出现	中国 and/文\$色=红 香港
xor/n	同一句中不能在n个字之内同时出现	中国 xor/10 香港
xor/n#	同一句中不能正好相距n个字同时出现	中国 xor/10# 香港
xor/nL	同一句中不能在n个字之外同时出现	中国 xor/10L 香港
xor/pos=n	同一句中不能正好相距n个字同时出现	中国 xor/pos=10 香港
xor/sen	在同一句中只能出现其一	中国 xor/sen 香港
xor/sen=n	在指定的句中只能出现其一	中国 xor/sen=1 香港
xor/seg	在同一段中只能出现其一	中国 xor/seg 香港
xor/seg=n	在指定的段中只能出现其一	中国 xor/seg=1 香港
xor/tit	在标题中只能出现其一	中国 xor/tit 香港
xor/txt	在正文中只能出现其一	中国 xor/txt 香港
xor/att	在指定属性的节点只能出现其一	中国 xor/文 香港
xor/att=v	在指定属性值的节点只能出现其一	中国 xor/章=2 香港
xor/att\$sub=v	在指定特征值的节点只能出现其一	中国 xor/文\$色=红 香港
not/n	同一句中不能在n个字之内出现后者	中国 not/10 香港
not/n#	同一句中不能正好相距n个字出现后者	中国 not/10# 香港
not/nL	同一句中不能在n个字之外出现后者	中国 not/10L 香港
not/pos=n	同一句中不能正好相距n个字出现后者	中国 not/pos=10 香港
not/sen	在同一句中只能出现前者	中国 not/sen 香港
not/sen=n	在指定的句中只能出现前者	中国 not/sen=1 香港
not/seg	在同一段中只能出现前者	中国 not/seg 香港
not/seg=n	在指定的段中只能出现前者	中国 not/seg=1 香港
not/tit	在标题中只能出现前者	中国 not/tit 香港
not/txt	在正文中只能出现前者	中国 not/txt 香港
not/att	在指定属性的节点只能出现前者	中国 not/文 香港
not/att=v	在指定属性值的节点只能出现前者	中国 not/章=2 香港
not/att\$sub=v	在指定特征值的节点只能出现前者	中国 not/文\$色=红 香港
or/sen=n	在指定的句中至少出现一个	中国 or/sen=1 香港
or/seg=n	在指定的段中至少出现一个	中国 or/seg=1 香港
or/tit	在标题中至少出现一个	中国 or/tit 香港
or/att	在指定属性的节点至少出现一个	中国 or/文 香港
or/att=v	在指定属性值的节点至少出现一个	中国 or/章=2 香港
or/att\$sub=v	在指定特征值的节点至少出现一个	中国 or/文\$色=红 香港

- 属性运算符只作用于DOCUMENT字段。adj、X/n、X/n#、X/nL和X/pos=n运算符可作用于PHRASE字段。‘X’代表‘pre’、‘and’、‘or’、‘not’和‘xor’，以下同，不再特别说明。
- 属性运算符X/att、X/att=v以及X/att\$sub=v中的att对应一个具体的格式索引属性别名（如果没有别名，则为格式索引属性名，其本身不是一个法定的语法符号，而只是用来说明语法的合法属性的表示符）；sub对应一个具体的格式索引属性的特征别名（如果特征没有别名，则为特征名，其本身不是一个法定的语法符号，而只是用来说明语法的合法特征的表示符），如果这个特征是该格式索引属性所定义的唯一特征，或者是第一个特征，则\$sub部分可以省略；v是一个格式索引属性或其特征的值，值的形式与具体的编码方式有关。如果att之前加上“~”符，则表示是属性限定条件的“非”运算，如“and/~att\$sub=v”的意义是“两者同时出现，且不满足‘att的sub特征值为v’的条件”。
- 运算符pre/att、pre/att=v和pre/att\$sub=v不是合法的，但att、att=v和att\$sub=v可与pos、sen、seg、tit、txt等一起使用，如属性运算符pre/txt/att\$sub=v则是合法的。
- 属性运算符中的“=”可以用其他任意比较运算符替代。但X/att=v和X/att\$sub=v中的“=”可以用哪些比较运算符替代，取决于格式索引属性或其特征的编码方式。
- 属性运算符X/sen=n和X/seg=n中的n，都从0开始计数。
- 属性运算符X/n[#/L]的隐含条件是它们在同一句中。N的计算方法是位置在前的检索词的词尾到位置在后的检索词的词首之间的字母单词或中文字的个数。其中字母L大小写无关，表示Leastwise。这里X/n等价于X/pos<=n；X/n#等价于X/pos=n；X/nL等价于X/pos>=n。
- 属性运算符adj等价于pre/0；equ/n等价于pre/n#。运算符adj和equ/n只是为了向下兼容而设计的，应尽量避免使用。
- 属性运算符X/txt不等价于X/seg>0(或X/seg!=0)，并且不能与/sen[=n]和/n[#/L]属性同时使用。X/txt不要求两个运算对象满足“段号相等”的条件。
- 属性运算符X/tit等价于X/seg=0。

6、限定运算符

表示符号	含义和作用	使用举例
/pos=n	检索词在指定的位置出现	中国/pos=0
/sen=n	检索词在指定的句中出現	中国/sen=12
/seg=n	检索词在指定的段中出现	中国/seg=10
/tit	检索词在标题中	中国/tit
/txt	检索词在正文中	中国/txt
/att	检索词具有att属性	中国/文
/att=v	检索词具有att属性，且值为v	中国/章=2
/att\$sub=v	检索词具有att属性，且sub特征值为v	中国/文\$色=红
/rec=n	输出满足命中数条件的索引词（值）	%/rec>=1000000
/fre=n	输出满足词频数条件的索引词（值）	%/fre>=1000000

- 限定符/rec=n和/fre=n只能用于输出索引词（或字段值），而不能用于检索，可作用于除BIT之外的所有字段。
- 限定符/pos=n只用于检索，可作用于DOCUMENT和PHRASE字段，

- 其它限定符只能用于在DOCUMENT字段上的检索。
- 限定符/att、/att=v以及/att\$sub=v中的att对应一个具体的格式索引属性别名（如果没有别名，则为格式索引属性名，其本身不是一个法定的语法符号，而只是用来说明语法的合法属性的表示符）；sub对应一个具体的格式索引属性的特征别名（如果特征没有别名，则为特征名，其本身不是一个法定的语法符号，而只是用来说明语法的合法特征的表示符），如果这个特征是该格式索引属性所定义的唯一特征，或者是第一个特征，则\$sub部分可以省略；v是一个格式索引属性或其特征的值，值的形式与具体的编码方式有关。如果att之前加上“~”符，则表示是限定条件的“非”运算，如“/~att\$sub=v”的意义是“检索词不满足‘具有att属性，且sub特征值为v’的条件”。
- 限定符中的“=”可以用其他任意比较运算符替代。但/att=v和/att\$sub=v中的“=”可以用哪些比较运算符替代，取决于格式索引属性或其特征的编码方式。
- 限定符中的n，都从0开始计数。其中/pos=n中的n是指检索词在句中出现的起始位置，一个字/词在句中的位置计算规则请参见3.4.1节说明。限定符/pos=n与属性运算符X/pos=n是有差别的。
- 限定符/txt等价于/seg>0(或/seg!=0)，并且可以与/sen=n和/pos=n同时使用，这一点与运算符X/txt不同。
- 限定符/tit等价于/seg=0。

7、加权运算符

表示符号	含义和作用	使用举例
/n	指定某个检索词在查询中的权重	计算机/10 * 科学

- 检索词在查询中的缺省权重是1。
- 权重n是一个不小于1的实数，且不能用科学记数法表示。
- 权重只对用于进行相关性排序的字段（DOCUMENT或PHRASE），并进行相关性排序时有效。在进行相关性排序前，系统需要根据固定的算法自动计算每个检索词对一个命中记录的相关度的作用因子（即贡献率），权重的作用就是用来倍增（放大）这个词的“作用因子”的。

A.2 等价运算符

常规运算符	!=	ADJ	PRE	AND	XOR	NOT	OR
等价运算符	<>	Space	&	*	^	-	+

A.3 运算符与字段的关系

运算符	DATE	NUMBER	CHAR	PHRASE	DOCUMENT
=	✓	✓	✓	✓	✓
!=	✓	✓	✓	✓	✓

>	✓	✓	✓		
<	✓	✓	✓		
>=	✓	✓	✓		
<=	✓	✓	✓		
*=	✓	✓	✓	✓	✓
^=	✓	✓	✓	✓	✓
=	✓	✓	✓	✓	✓
+=	✓	✓	✓	✓	✓
adj				✓	✓
equ/n				✓	✓
pre/n[#/L]				✓	✓
pre/pos=n				✓	✓
pre/sen[=n]					✓
pre/seg[=n]					✓
pre/tit					✓
pre/txt					✓
pre/[~]att[[\$sub]=v]					✓
and	✓	✓	✓	✓	✓
and/n[#/L]				✓	✓
and/pos=n				✓	✓
and/sen[=n]					✓
and/seg[=n]					✓
and/tit					✓
and/txt					✓
and/[~]att[[\$sub]=v]					✓
xor	✓	✓	✓	✓	✓
xor/n[#/L]				✓	✓
xor/pos=n				✓	✓
xor/sen[=n]					✓
xor/seg[=n]					✓
xor/tit					✓
xor/txt					✓
xor/[~]att[[\$sub]=v]					✓
not	✓	✓	✓	✓	✓
not/n[#/L]				✓	✓
not/pos=n				✓	✓
not/sen[=n]					✓
not/seg[=n]					✓
not/tit					✓
not/txt					✓
not/[~]att[[\$sub]=v]					✓
or	✓	✓	✓	✓	✓
or/sen=n					✓
or/seg=n					✓
or/tit					✓
or/txt					✓
or/[~]att[[\$sub]=v]					✓
/pos=n				✓	✓
/sen=n					✓
/seg=n					✓
/tit					✓
/txt					✓
/[~]att[[\$sub]=v]					✓
/rec=n	✓	✓	✓	✓	✓

/fre=n	✓	✓	✓	✓	✓
/n				✓	✓

- 表中的“✓”表示该字段支持这种运算符。
- 算术运算符不能用于检索表达式，所以任何字段都不支持。
- 所有运算符都不能作用于BIT字段。
- 运算符中的‘[’和‘]’如 `and/[~]att[[$sub]=v]` 不是运算符的一部分，表示被包括的部分可以省略。但省略与不省略，其运算符的含义是有差别的。

A.4 统计函数

TRS 系统所支持的函数分为两种：统计函数和检索函数。统计函数是统计表达式的一部分，用于对检索结果的统计；而检索函数则是检索表达式的一部分，以函数的形式描述检索结果必须满足的条件。

函数名	参数	含义和作用
AVG()	字段名	求数据库中满足条件的记录中某个数值型字段的平均值
MAX()	字段名	求数据库中满足条件的记录中某个数值型字段的最大值
MIN()	字段名	求数据库中满足条件的记录中某个数值型字段的最小值
SUM()	字段名	求数据库中满足条件的记录中某个数值型字段的和
COUNT()	*	用于统计数据库中满足条件的记录有多少

- 统计函数只能用于统计表达式。
- 函数参数中的“字段名”也可以使用字段的别名。
- 只有 NUMBER 型字段的字段名才能作为统计函数的参数。

A.5 检索函数

函数名	参数	含义和作用
BT()	主题词[, 词典名]	获得某个主题词的上位词
NT()	主题词[, 词典名]	获得某个主题词的下位词
PT()	主题词[, 词典名]	获得某个主题词的等同词
UF()	主题词[, 词典名]	获得某个主题词的替代词
AB()	主题词[, 词典名]	获得某个主题词的缩略词
ABF()	主题词[, 词典名]	获得某个主题词的被缩略词
RT()	主题词[, 词典名]	获得某个主题词的相关词
LE()	主题词[, 词典名]	获得某个主题词的外文等同词
ALLBT()	主题词[, 词典名]	获得某个检索词的所有层次的上位词
ALLNT()	主题词[, 词典名]	获得某个检索词的所有层次的下位词
ST()	主题词[, 词典名]	获得某个检索词的广义同义词
AT()	主题词[, 词典名]	获得某个检索词的反义词
BETWEEN()	值1, 值2	获取在指定范围之内的字段值
BEYOND()	值1, 值2	获取在指定范围之外的字段值
EXACT()	值	对值不进行任何处理，直接匹配索引
INCLUDE()	词串	对词串不进行分词处理，并加上前后模糊

AMB()	词串	用扩展查找的方法来解决由于分词错误所引起的漏检问题，或者满足非词检索的要求，又称邻接检索（一般不鼓励使用）
ALL()	词串	对词串进行分词处理，并包含词串本身
RANGE()	词串=值	在同一句进行数值范围检索（一般不鼓励使用）
LOWER()	词串	转化为小写，只对字母串
UPPER()	词串	转化为大写，只对字母串
TOCHINESE()	词串	把词串中的半角字母数字转化为全角
TOASCII()	词串	把词串中的全角字母数字转化为半角
CASE()	词串	进行大小写无关检索，只对字母串
LIKE()	词串[, 开关]	对词串进行分词处理，忽略词之间的位置关系，进行简单的逻辑运算
DATA()	值列表	直接扫描数据进行精确匹配，无参数时表示空值查找

- 检索函数BETWEEN和BEYOND只能作用于DATE、NUMBER以及CHAR字段；EXACT、LOWER、UPPER、TOCHINESE、TOASCII和CASE检索函数可作用于除BIT之外的所有字段；DATA函数可作用于包括BIT字段在内的所有字段；其他检索函数则只能作用于PHRASE和DOCUMENT两个可进行全文检索的字段。其中LOWER、UPPER、TOCHINESE、TOASCII和CASE检索函数在以英文为字符集的数据库中无效，即不进行任何转换。
- TRS系统支持辅助知识自动扩展检索，这一功能涵盖了大部分的检索函数，包括：BT、NT、PT、UF、AB、ABF、RT、LE、ALLBT、ALLNT、ST、AT以及TOCHINESE和TOASCII。自动扩展检索可以对检索表达式里的所有词自动施加这些函数调用，无须显式调用。
- 检索函数BETWEEN和BEYOND的参数表的圆括号‘(’和‘)’可以分别使用方括号‘[’和‘]’替换。‘(’表示值1不包含在范围之内，而‘[’表示值1包含在范围之内；‘]’表示值2不包含在范围之内，而‘)’表示值2包含在范围之内，并且值1和值2都不支持模糊字段值。
- 函数BT、NT、PT、UF、AB、ABF、RT、LE、ALLBT、ALLNT、ST、AT参数中的‘[’和‘]’不是参数的一部分，表示词典可以省略。当词典省略时，系统将根据被检索的数据的语言类型，获取系统缺省的相应类型的词典作为被引用的词典。其中函数ST所引用的词典必须是同义词典，函数AT所引用的词典必须是反义词典，其它函数所引用的词典必须是主题词典。
- RANGE函数中比较运算符‘=’可以用其他任意比较运算符替代。
- LIKE函数主要用来改变全文检索的方式。缺省的全文检索方式是：对于给定的短语（词串），经过分词处理后，各词之间要进行严格的位置运算。而对以LIKE函数参数给定的短语（词串），经过分词处理后，各词之间不进行位置运算，而只进行简单的逻辑运算：如果参数表是以圆括号对‘(’和‘)’表示的，则各词之间的逻辑关系是“与”；如果参数表是以方括号对‘[’和‘]’表示的，则各词之间的逻辑关系是“或”。第二个“开关”参数表示是否忽略二元索引：取值‘TRUE’（大小写无关）或者‘1’、或者缺省时，表示忽略二元索引，以提高查全率（更多的结果）；取值‘FALSE’（大小写无关）或者‘0’时，表示不忽略二元索引，以提高检索速度。LIKE函数等价于‘and’和‘or’运算符而允许进行属性运算，其书写格式如：
LIKE(...)/[~]att[[\$sub]=v]/seg[=n]/sen[=n]。
- DATA函数的参数形式与5.4.2节中“检索键值列表”的形式是一致的。参数表的圆括号对‘(’和‘)’可以使用方括号对‘[’和‘]’替换，表示作用于CHAR型字段时，不区分字母的大小写。如果DATA函数作用于PHRASE、DOCUMENT和BIT字段，则只允许无参数调用。
- 函数EXACT、LOWER、UPPER、TOCHINESE、TOASCII和CASE检索函数，以及参数列表

只有单个值的DATA函数的结果值可用于进行字段值大小比较的表达式中，如：“字段名>function(...)”，其它函数则不能出现在这样的表达式中，而只能用于“等于”或“不等于”的表达式中。

A.6 统计表达式

统计表达式分为两类：一类是计算总的命中结果记录数。这是缺省形式，其表示非常简单，只有一个单字节的星号‘*’。另一类是对 NUMBER 型字段在命中结果集上进行算术运算。这类统计表达式较为复杂，下面将详细说明其表达式的结构。

- “统计表达式”的组成：

- 子统计表达式
- 子统计表达式, 子统计表达式, …… , 子统计表达式

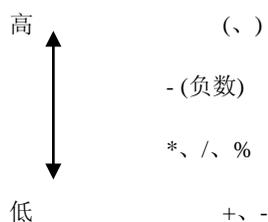
其中，单字节的逗号‘,’是一种语法符号，是“统计表达式”的一部分，用于分隔多个“子统计表达式”。省略号“……”表示可有多个“子统计表达式”。

- “子统计表达式”的组成：

- 数值
- 统计函数
- 子统计表达式
- (子统计表达式)
- 子统计表达式 ArithOP 子统计表达式

“子统计表达式”的组成具有自递归的特点。其中“ArithOP”代表所有的算术运算符。

- 统计表达式中算术运算符的优先级为：(同级运算遵循左结合律)



A.7 检索表达式

- “检索表达式”的组成：

- 逻辑运算表达式
- 限定运算表达式

- “逻辑运算表达式”的组成：

- 检索键值
- 检索函数
- @检索历史号
- DOCID CompOP 记录号
- DOCID = 记录号列表
- DOCID != 记录号列表
- ENTRY CompOP 检索键值
- ENTRY = (字段逻辑表达式)
- ENTRY != (字段逻辑表达式)
- ENTRY, ENTRY, ……, ENTRY RepeatOP 检索键值
- ENTRY, ENTRY, ……, ENTRY RepeatOP (字段逻辑表达式)
- 逻辑运算表达式 LogicOP 逻辑运算表达式
- (逻辑运算表达式)
- 限定运算表达式

“检索表达式”的组成具有自递归的特点。其中，“ENTRY”是字段名或字段别名，单字节的逗号‘,’是语法符号，是“检索表达式”的一部分，用于分隔多个“ENTRY”，省略号“……”表示可有多个“ENTRY”。“CompOP”代表所有的比较运算符。“RepeatOP”代表所有的重复逻辑运算符。“LogicOP”代表所有的逻辑运算符。

- “字段逻辑表达式”的组成：

- 检索键值
- 检索函数
- 检索键值列表
- 字段逻辑表达式 LogicOP 字段逻辑表达式
- (字段逻辑表达式)
- 限定运算表达式

“字段逻辑表达式”的组成具有自递归的特点。其中，“LogicOP”代表所有的逻辑运算符。

- “检索键值列表”的组成：

- 精确检索键值

- 精确检索键值 TO 精确检索键值
- 检索键值列表, 检索键值列表, …… , 检索键值列表

“检索键值列表”的组成具有自递归的特点。其中，单字节的逗号 ‘,’ 是语法符号，是“检索键值列表”的一部分，用于分隔多个“检索键值列表”，省略号 “……” 表示可有多个“检索键值列表”。“TO” 表示一个连续的检索键值区间。

● “记录号列表”的组成：

- 记录号
- 记录号 TO 记录号
- 记录号列表, 记录号列表, …… , 记录号列表

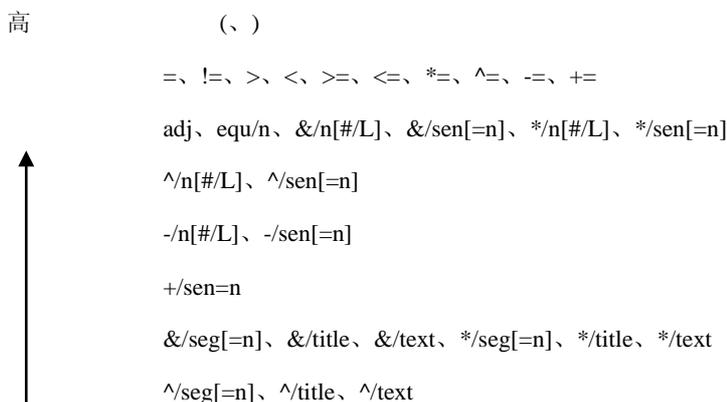
“记录号列表”的组成具有自递归的特点。其中，单字节的逗号 ‘,’ 是语法符号，是“记录号列表”的一部分，用于分隔多个“记录号列表”，省略号 “……” 表示可有多个“记录号列表”。“TO” 表示一个连续的记录号区间。

● “限定运算表达式”的组成：

- 检索键值
- 检索函数
- 检索键值/Limit
- 检索函数/Limit (除 Range、between、beyond 外)
- 限定运算表达式 LimitOP 限定运算表达式
- (限定运算表达式)

“限定运算表达式”的组成具有自递归的特点。其中，“LimitOP” 代表所有的属性运算符。“Limit” 代表所有的限定运算符和加权运算符。

● 检索表达式中各种运算符的优先级为：(同级运算遵循左结合律)



`-/seg[=n]、-/title、-/text`

`+/seg=n、+/title、+/text`

`*/attr[=v]`

`^/attr[=v]`

`-/attr[=v]`

`+/attr=v`

`*`

`^`

`-`

低

+

A.8 检索键值

TRS 既支持在 DATE、NUMBER 和 CHAR 字段上通过键值进行数据查询，同时也支持在 PHRASE 和 DOCUMENT 字段上用任意自由词、词串以及数值进行检索。请参见特殊键值的检索。

1. DATE 型字段上的键值

在检索表达式中，DATE 键值的格式有通用格式、英文格式、无分隔符格式和中文格式四种类型。

- 通用格式

`Year[/Month[/Day[/Hour[/Minute[/Second]]]]][A[M]]`

`[Year]/Month[/Day[/Hour[/Minute[/Second]]][A[M]]`

`[Year]/[Month]/Day[/Hour[/Minute[/Second]]][A[M]]`

`[Year]/[Month]/[Day]/Hour[/Minute[/Second]][A[M]]`

`[Year]/[Month]/[Day]/[Hour]/Minute[/Second][A[M]]`

`[Year]/[Month]/[Day]/[Hour]/[Minute]/Second[A[M]]`

其中：

- Year 为一到四位表示的完全年份(如 1998)。
- Month、Day、Hour、Minute、Second 分别为二位或一位表示的月、日、时、分、秒。
- A/[M]可用 P/PM 替换。A/AM 表示上午，P/PM 表示下午，大小写无关。如果它们出现在格式中，则“Hour”不能却省。
- 单字节除号 ‘/’ 分隔符可分别用单字节的逗号 ‘,’、句号 ‘.’、冒号 ‘:’ 以及减号 ‘-’ 任意替代，在两个明确（不省略）的日期成份之间，空格也可作为分隔符。
- 用方括号引起的部分表示可以省略（模糊日期）。
- Year、Month、Day、Hour、Minute、Second 都可以用一个或多个单字节的问号 ‘?’、百分号 ‘%’ 替代，表示模糊日期。

● 英文格式

Month[Day[,Year[Hour[:Minute[:Second]]]]][A[M]]

Month[Day],Year[Hour[:Minute[:Second]]][A[M]]

Month[Day],Year[Hour]:Minute[:Second][A[M]]

Month[Day],Year[Hour]:[Minute]:Second[A[M]]

[Day-]Month[-Year[Hour[:Minute[:Second]]]]][A[M]]

[Day-]Month[-Year[Hour]:Minute[:Second]]][A[M]]

[Day-]Month[-Year[Hour]:[Minute]:Second][A[M]]

DayMonthYear[Hour[:Minute[:Second]]][A[M]]

DayMonthYear[[Hour]:Minute[:Second]]][A[M]]

DayMonthYear[[Hour]:[Minute]:Second][A[M]]

其中：

- Year 为一到四位表示的完全年份(如 1998)。
- Month 为英文表示的月份：Jan. Feb. Mar. Apr. May Jun. Jul. Aug. Sept. Oct. Nov. Dec.；或：January February March April May June July August September October November December。
- Day、Hour、Minute、Second 分别为二位或一位表示的月、日、时、分、秒。
- A/[M]可用 P/PM 替换。A/AM 表示上午，P/PM 表示下午，大小写无关。如果它们出现在格式中，则“Hour”不能却省。
- Year 与 Hour 之间用单字节空格分隔；‘,’ ‘:’ ‘-’ 分别为单字节的逗号、冒号、减号。
- 用方括号引起的部分表示可以省略（模糊日期）。

● 无分隔格式

YearMonth[Day[Hour[Minute[Second]]]]]

其中：

- Year 为固定四位表示的完全年份(如 1998)。
- Month、Day、Hour、Minute、Second 分别为二位表示的月、日、时、分、秒。
- 用方括号引起的部分表示可以省略(模糊日期)。

● 中文格式

中文格式包括所有用全角字符表示的上述几种格式，另外还支持如下的格式：

[Year年][Month月][Day日][Hour时][Minute分][Second秒][上午]

其中：

- Year 为一到四位表示的完全年份(如 1998 或一九九八或壹玖玖捌)。其中数字既可为半角，也可为全角。
- Month、Day、Hour、Minute、Second 分别为二位或一位表示的月、日、时、分、秒。
- 格式中“上午”可用“下午”替换。如果“上午”或“下午”出现在格式中，则“Hour 时”不能却省。
- 用方括号引起的部分表示可以省略(模糊日期)。

另外，还支持\$DATE 和\$TIME 特殊日期(不区分大小写)的检索，其中前者表示当前日期的“年月日”，后者表示当前日期的“年月日时分秒”。

在阅读《TRS 数据库加载格式规范》时会发现，其日期格式中的“年份”与这里所要求的不一样：加载时“年份”可以省略“世纪数”，而检索时的“年份”没有缺省“世纪数”，无论多少位都表示的是完全年份，如“98”表示公元 98 年，而不是公元 1998 年。在检索时，无分隔格式中的“年份”必须是四位，如日期“190610”表示 1906 年 10 月，但在数据库加载时则表示 1919 年 6 月 10 日(假设缺省世纪数为 19)。之所以存在这种差别，主要原因是在数据库加载时，往往是大批量的日期数据，通过加载控制文件指定年份的缺省世纪数可以支持按习惯方法表示的日期，而检索时则没有必要使用其他方法指定年份的缺省世纪数，直接写出完全的年份更为简捷方便。

2. NUMBER 型字段上的键值

在检索表达式中，NUMBER 型键值是一个实数。NUMBER 型字段上不支持模糊检索。

3. CHAR 型字段上的键值

在检索表达式中，CHAR 键值是一个字符串。这个字符串可以单引号引起来，也可以

不用。对于普通的键值，是否被引号引起来，或是否被转义，都具有相同的含义。

TRIS 支持对键值的模糊匹配。用通配符“?”表示一个英文字符或汉字，用“%”表示 0 到多个英文字符或汉字。通配符可出现在键值的首部、尾部和中间的任意位置。组合‘??...?’和‘%??...?’的意义相同，都表示在模糊匹配的字符串中至少有 n 个英文字符或汉字，其中 n 是‘?’的个数。例如：

- 表达式“作者=%国庆”将检索出所有作者为“张国庆”、“欧阳国庆”、“国庆”等的记录；
- 表达式“作者=张%”将检索出所有作者为“张国庆”、“张三”、“张”等的记录；
- 表达式“作者=张%庆”将检索出所有作者为“张国庆”、“张大庆”、“张庆”等的记录；
- 表达式“作者=张?庆”将检索出所有作者为“张国庆”、“张大庆”等的记录，而作者为“张庆”的记录不会被检索出来。

另外，还支持\$USER 特殊键值（不区分大小写）的检索，表示检索字段值等于当前操作的用户名。

4. PHRASE 和 DOCUMENT 型字段上的键值

在检索表达式中，PHRASE 和 DOCUMENT 的键值与 CHAR 键值的具有相同的形式，区别在于 PHRASE 和 DOCUMENT 的键值须经过分词处理后在进行检索。

A.9 特殊键值

TRIS 检索表达式中有四类语法符号：

第一类：‘(’、‘)’、‘;’、‘/’、‘@’、‘=’、‘>’、‘<’、‘!’、‘&’、‘*’、‘^’、‘-’、‘+’。

第二类：‘ADJ’、‘EQU’、‘PRE’、‘AND’、‘XOR’、‘NOT’、‘OR’、‘TO’、空格、函数名。

第三类：‘%’、‘?’。

第四类：单引号和转义符‘\’。

检索表达式的键值必须对这些符号进行特殊处理：

- 如果键值中含有第一类语法符号，则必须将该键值用单引号引起来，或者用转义符将这些语法符号进行转义。如：日期键值 ‘.,8’、\.,8。
- 如果第二类语法符号是一个键值，则必须将该键值用单引号引起来，或者用转义符将这些语法符号进行转义。如：‘AND’、\AND。
- 如果第三类语法符号是键值的值的一部分，而不是模糊匹配符，则必须用转义符将这些语法符号进行转义；否则无论是被引号引起来，还是没有被引起来，都将解释为模糊匹配符。如：10\%。
- 如果第四类语法符号是键值的值的一部分，则必须用转义符将它们再转义一次。

A.10 保留字

以下是 TRS 所用的保留字，这些保留字不能作为对象名使用：

AB	ABF	ACCOUNT	ACTIVE
ADD	ADJ	ALIAS	ALL
ALLBT	ALLNT	ALONE	ALTER
AMB	AND	ANTONYM	AS
ASC	AT	AVG	
BACKUP	BASE	BATCH	BETWEEN
BEYOND	BIG5	BIT	BROWSE
BT	BY		
CASE	CHAR	CLASS	CLEAR
CONDITION	COLUMN	COMMENT	CONNECT
CONTROL	COUNT	CREATE	
DATABASE	DATA	DATE	DBA
DEFAULT	DELETE	DESC	DIR
DISPLAY	DOCID	DOCUMENT	DOWNLOAD
DROP	DUMB		
ECHO	EDIT	ENFORCE	ENGLISH
ENTRY	EQU	EXCLUSIVE	EXACT
EXIT			
FOR	FORMAT	FRE	FREC
FROM			
GB18030	GB2312	GBK	GDA
GRANT	GROUP		
HELP	HIGHT	HISTORY	HOST
IDENTIFIED	IN	INCLUDE	INDEX
INSERT	INTERVAL	INTO	INVERT
ISO2709			
LANGUAGE	LE	LENGTH	LIKE
LINK	LOAD	LOCK	LOWER
MAIL	MAX	MIN	MODE
MODIFY	MULTIVALUE	MULTILINGUAL	
NO	NOT	NOTINDEX	NOWAIT
NT	NULL	NUMBER	

OFF	ON	OPTIMIZE	OR
ORDER			
PATH	PASSWORD	PHRASE	PRE
PRECISION	PRINT	PT	PUBLIC
QUIT			
RANGE	REAL	REC	RECORD
RECOVER	RELEVANCE	REM	RENAME
RESOURCE	RESTORE	REVOKE	RT
RUN			
SAVE	SECURITY	SEGMENT	SEGMENTEX
SELECT	SEN	SEG	SET
SHARE	SHOW	SIZE	SLEEP
SPARSE	SPOOL	ST	STEM
STOP	STOPEX	STRUCTURE	SUFFIX
SUM	SYNONYM		
TEXT	THESAURUS	TIME	TILL
TITLE	TO	TOASCII	TOCHINESE
TRS	TYPE		
UF	UNICODE	UNIQUE	UPDATE
UPPER	URL	USER	UTF8
VALIDATE	VALUEMUST	VALUES	VIEW
WIDTH	WHERE	WORDS	
XOR	XML		
YES			

附录 B 数据库加载格式规范

B.1 数据记录文件

数据记录文件实际上是由一个或多个记录数据组成的 TRS 置标文本文件。每个记录都以标记 “<REC>” 开始，即标记 “<REC>” 是记录之间的分隔符。第一个记录前和最后一个记录后的该标记可以省略；如果文件中只有一个记录，则该标记可以省略。

数据记录文件的一般格式是：

```
<REC>  CRYPT
字段1
字段2
.....
字段n
<REC>  CRYPT
字段1
字段2
.....
字段n
其中：
```

- 标记<REC>必须顶格书写。
- 字段 n 的描述共有 5 种不同的格式，分别为“标准格式”、“字段内部标识号标记格式”、“字段顺序号标记格式”、“无字段标记格式”和“全文格式”。
- 标记<REC>后面的 CRYPT 是记录的加密描述符。CRYPT 必须与标记<REC>写在同一行。加密描述符可以省略。

CRYPT 由下列一个或多个分项组成：

- SecurityCode = CodeValue。其中“SecurityCode”可使用缩写“SC”，不区分大小写。“CodeValue”是一个有效的用户密级（数字），表示只有用户的密级不大于该值时才能读出该记录。
- UserGroup = GroupName。其中“UserGroup”可使用缩写“UG”，不区分大小写。“GroupName”是一个有效的用户组名（不区分大小写），表示只有该用户组的用户才能读出该记录。
- Password = PasswordString。其中“Password”可使用缩写“PWD”，不区分大小写。“PasswordString”是一个不超过 8 个字符的字符串（大小写敏感），表示该记录读出时必须提供相同的口令。

CRYPT 中包含多个分项时，分项之间用单字节的逗号“,”或分号“;”或空格分隔，分项之间没有顺序限制，但每个分项必须是完整的。如：

```
<REC> SecurityCode = 10; UserGroup = Administrator;  
.....  
<REC> PWD = OnlyMe  
.....  
<REC>  
.....
```

1. 标准格式

标准格式的 TRS 数据记录文件的组织格式为：

```
<REC>  
<字段名1>=字段值1.....  
.....  
<字段名2>=字段值2.....  
.....  
<字段名n>=字段值n.....  
.....  
<REC>  
<字段名1>=字段值1.....  
.....  
<字段名2>=字段值2.....  
.....  
<字段名n>=字段值n.....  
.....
```

其中：

- 标记引导符“<”必须顶格书写，标记结束符“>=”是一个整体，中间不能有任何其他字符。这两个标记符号都是 ASCII 码，必须成对出现。
- 标记引导符“<”和标记结束符“>=”之间是一个字段名，或者是一个字段的别名。“<字段名 n>=”必须书写在同一行。
- 标记结束符“>=”与下一个标记引导符“<”之间的一行或多行文本为该字段的值。文本前后的空字符不计入字段值，但 PHRASE 和 DOCUMENT 字段可通过控制文件的设置（KEEP_TEXT_ACTUAL 变量）保持其文本原样。字段值可以为空。
- 字段之间没有顺序限制。多值字段的值可以用一个“<字段名 n>=字段值 n”的方式书写，多值之间用特定的字符或字符串分隔（控制文件中 MULTI_VALUE_SEPARATER 变量的值必须与其相一致）；也可以用多个“<字段名 n>=字段值 n”的方式书写；这两种方式也可以同时使用。
- 对于 PHRASE 和 DOCUMENT 字段，如果在一个记录中出现多次，这根据控制文件的设置（CHECK_TEXT_ITERATIVE 变量）进行如下处理：如果控制文件的设置不要求检查这两种字段的重复值，则忽略前面出现的字段值，而只保留最后出现的字段值；否则认为这是一个错误的记录数据。

下面是一个包含两条记录的“专利数据库”的标准格式的数据记录文件：

```
<REC>  
<公告日>=920101  
<分类号>=A01B49/02
```

<公告号>=CN_2091520U
 <申请号>=91208300.X
 <申请日>=910514
 <申请人>=宝鸡县机电设备维修厂
 <地址>=721304陕西省宝鸡县钓渭乡
 <设计人>=赵志栋
 <专利代理机构>=宝鸡市专利事务所
 <代理人>=翟中平

<摘要>=耕耙犁

耕耙犁是用于农田深耕、碎土的农具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

<REC>

<公告日>=920101
 <分类号>=A01B49/04
 <公告号>=CN_2091521U
 <申请号>=91206137.5
 <申请日>=910411
 <申请人>=张岳
 <地址>=475000河南省开封市南太平街3号楼2单元付3号
 <设计人>=张岳
 <设计人>=张龙
 <设计人>=范文增
 <设计人>=石建峡
 <设计人>=阎唯一
 <专利代理机构>=开封市专利事务所
 <代理人>=黄永武
 <代理人>=乔玉萍

<摘要>=牵引式多功能农用机

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农用机。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔的市场。

2. 字段内部标识号标记格式

字段内部标识号标记格式与标准格式的唯一区别在于：标准格式中用标记“<字段名>=”来指示一个字段，而这种格式则用标记“<字段内部标识号>=”来指示一个字段。字段内部标识号是指该字段在数据词典中的唯一标识号。

下面是这个“专利数据库”的字段内部标识号标记格式的数据记录文件组织：

<REC>
 <1>=920101
 <2>=A01B49/02

<3>=CN_2091520U
 <4>=91208300.X
 <5>=910514
 <7>=宝鸡县机电设备维修厂
 <8>=721304陕西省宝鸡县钓渭乡
 <9>=赵志栋
 <10>=宝鸡市专利事务所
 <11>=翟中平
 <12>=耕耙犁

耕耙犁是用于农田深耕、碎土的农机具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农机具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

<REC>
 <1>=920101
 <2>=A01B49/04
 <3>=CN_2091521U
 <4>=91206137.5
 <5>=910411
 <7>=张岳
 <8>=475000河南省开封市南太平街3号楼2单元付3号
 <9>=张岳
 <9>=张龙
 <9>=范文增
 <9>=石建峡
 <9>=阎唯一
 <10>=开封市专利事务所
 <11>=黄永武
 <11>=乔玉萍
 <12>=牵引式多功能农农机

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农农机。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔的市场。

这里假设各个字段的内部标识号为：

字段名	字段内部标识号	字段名	字段内部标识号
公告日	1	申请人	7
分类号	2	地址	8
公告号	3	设计人	9
申请号	4	专利代理机构	10
申请日	5	代理人	11
颁证日	6	摘要	12

3. 字段顺序号标记格式

字段顺序号标记格式与字段内部标识号标记格式非常相似，仅有两点差别：

- 字段内部标识号标记格式则用标记“<字段内部标识号>=”来指示一个字段，而这种格式是用标记“<字段顺序号>=”来指示一个字段。
- 字段顺序号是字段在控制文件中（COLUMN_NAME_SEQUENCE 变量）给出的字段列表中的次序。第一个字段的顺序号为 1，依次类推。

下面是“专利数据库”的字段顺序号标记格式的数据记录文件组织：

```
<REC>
<1>=920101
<2>=A01B49/02
<3>=CN_2091520U
<4>=91208300.X
<5>=910514
<6>=宝鸡县机电设备维修厂
<7>=721304陕西省宝鸡县钓渭乡
<8>=赵志栋
<9>=宝鸡市专利事务所
<10>=翟中平
<11>=耕耙犁
```

耕耙犁是用于农田深耕、碎土的农具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

```
<REC>
<1>=920101
<2>=A01B49/04
<3>=CN_2091521U
<4>=91206137.5
<5>=910411
<6>=张岳
<7>=475000河南省开封市南太平街3号楼2单元付3号
<8>=张岳
<8>=张龙
<8>=范文增
<8>=石建峡
<8>=阎唯一
<9>=开封市专利事务所
<10>=黄永武
<10>=乔玉萍
<11>=牵引式多功能农用机
```

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农用机。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便，适

用于个体户、农场、农机站的需要，具有广阔的市场。

这里假设控制文件中给出的字段列表为：公告日;分类号;公告号;申请号;申请日;申请人;地址;设计人;专利代理机构;代理人;摘要。

4. 无字段标记格式

无字段标记格式是指在 TRS 数据记录文件的一个记录中，字段之间不是用显式的标记来分隔，而是通过自然的文本行来标示不同的字段。这种格式的特点是：

- 字段的顺序由控制文件（COLUMN_NAME_SEQUENCE 变量）定义。
- 除最后一个字段外，每个字段的值只能占用一行。没有值的字段，如果不是最后一个字段，则也必须用一空行来表示。
- 最后一个字段的值可以是多行，直到下一个记录的标记引导符“<”。也就是说，字段值多于一行的字段必须是每个记录的最后一个字段。
- 如果有多个字段的值多于一行，则不能采用这种格式组织数据记录文件。

下面是“专利数据库”的无字段标记格式的数据记录文件组织：

```
<REC>
920101
A01B49/02
CN_2091520U
91208300.X
910514
宝鸡县机电设备维修厂
721304陕西省宝鸡县钓渭乡
赵志栋
宝鸡市专利事务所
翟中平
耕耙犁
```

耕耙犁是用于农田深耕、碎土的农机具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农机具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

```
<REC>
920101
A01B49/04
CN_2091521U
91206137.5
910411
张岳
475000河南省开封市南太平街3号楼2单元付3号
张岳;张龙;范文增;石建峡;阎唯一
开封市专利事务所
黄永武;乔玉萍
牵引式多功能农动机
```

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能

农用车。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔的市场。

这里假设控制文件中给出的字段列表为：公告日;分类号;公告号;申请号;申请日;申请人;地址;设计人;专利代理机构;代理人;摘要。

5. 全文格式

全文格式的特点是：

- 数据记录文件的每个记录中只有一个字段。
- 记录中没有字段标记。
- 加载目标字段由控制文件（COLUMN_NAME_SEQUENCE 变量）定义，并且必须是一个 DOCUMENT 或 PHRASE 字段。

下面是一个包含两条记录的“红楼梦数据库”的全文格式的数据记录文件：

<REC>

第一回 甄士隐梦幻识通灵 贾雨村风尘怀闺秀

此开卷第一回也。作者自云：因曾历过一番梦幻之后，故将真事隐去，而借“通灵”之说，撰此《石头记》一书也。故曰“甄士隐”云云。但书中所记何事何人？自又云“今风尘碌碌，一事无成，忽念及当日所有之女子，一一细考较去，觉其行止见识，皆出于我之上。何我堂堂须眉，诚不若彼裙钗哉？实愧则有余，悔又无益之大无可如何之日也！当此，则自欲将已往所赖天恩祖德，锦衣纨裤之时，饫甘餐肥之日，背父兄教育之恩，负师友规谈之德，以至今日一技无成，半生潦倒之罪，编述一集，以告天下人：我之罪固不免，然闺阁中本自历历有人，万不可因我之不肖，自护己短，一并使其泯灭也。虽今日之茅椽蓬牖，瓦灶绳床，其晨夕风露，阶柳庭花，亦未有妨我之襟怀笔墨者。虽我未学，下笔无文，又何妨用假语村言，敷演出一段故事来，亦可使闺阁昭传，复可悦世之目，破人愁闷，不亦宜乎？故曰“贾雨村”云云。

.....

<REC>

第二回 贾夫人仙逝扬州城 冷子兴演说荣国府

诗云

一局输赢料不真，香销茶尽尚逡巡。
欲知目下兴衰兆，须问旁观冷眼人。

却说封肃因听见公差传唤，忙出来陪笑启问。那些人只嚷：“快请出甄爷来！”封肃忙陪笑道：“小人姓封，并不姓甄。只有当日小婿姓甄今已出家一二年了，不知可是问他？”那些公人道：“我们也不知什么‘真’‘假’，因奉太爷之命来问，他既是你女婿，便带了你去亲见太爷面禀，省得乱跑。”说着，不容封肃多言，大家推拥他去了。封家人个个都惊慌，不知何兆。

那天约二更时，只见封肃方回来，欢天喜地。众人忙问端的。他乃说道：“原来本府新升的太爷姓贾名化，本贯胡州人氏，曾与女婿旧日相交。方才在咱门前过去，因见娇杏那丫头买线，所以他只当女婿移住于此。我一将原故回明，那太爷倒伤感叹息了一回，又问外孙女儿，我说看灯丢了。太爷说：‘不妨，我自使番役务必探访回来。’说了一回话，临走倒送了我二两银子。”甄家娘子听了，不免心中伤感。一宿无话。

.....

这里假设控制文件中给出的字段列表为：COLUMN_NAME_SEQUENCE=正文。

6. ISO2709 国际标准格式

TRS 能够将 ISO2709 格式的数据文件的直接加载到数据库中。ISO2709 格式是一种国际标准，这里就不再进行说明。

下面是一个包含两条记录的“专利数据库”的 ISO2709 格式的数据文件：

```
00640000000000181000550000100011000000020001000011003000120002100400011
00033005000110004400600011000550070000700066008000380007300900007001110
10000010011801100001001190120033800120#1993.01.06#A01K83/04#CN_2126505Y
#92203792.2#1992.03.06#1993.01.03#王卫平#100045北京市西城复外真武庙6号楼1604
号#王卫平###
```

防逃式钓鱼器

本实用新型属于一种防逃式钓鱼器。它的作用原理：当鱼咬钩时，把连接固定钩（7）的销（4）从套管（3）中拉出，活动钩连线（8）被拉直，使活动钩（9）张开，活动钩与固定钩共同作用，牢牢卡住鱼嘴。本防逃式钓鱼器结构简单，动作灵敏，鱼一旦上钩不易逃脱。可实现在无人看管的情况下钓鱼。本钓鱼器制作工艺简单，便于成批系列化生产。

```
##005760000000001810005500001000110000000200009000110030001200020004000
11000320050001100043006000110005400700013000650080002700078009000210010
5010000170012601100015001430120023600158#1993.01.06#A01M9/00#CN_2126506
Y#92226804.5#1992.07.09#1993.01.03#北京农业大学#100094北京市圆明园西路二号#巨
修练；崔泌；孙继新#农业部专利事务所#林祥明；苟显伦#
```

手动静电喷粉机及其电晕带电喷头

本实用新型提供了一种手动静电喷粉机及其电晕带电喷头，该手动静电喷粉机可广泛用于施用农药粉剂，防治病虫害。试验证明，使用本实用新型施用农药，其沉积量可提高 50—80%，靶标反面沉积量可提高 3—4 倍。

##

这里假设控制文件中给出的字段列表为：公告日；分类号；公告号；申请号；申请日；颁证日；申请人；地址；设计人；专利代理机构；代理人；摘要。

需要注意的一点：格式中的“字段标识符”域是该字段在控制文件字段列表中的序号（从 1 开始计数）。参照上例，如某“字段标识符”是 12，则该字段就是数据库中的“摘要”字段。

7. XML 格式

和 ISO2709 一样，XML 格式也是一种国际标准，具体说明请参考有关资料，其中子字段也不会被解析。

下面是一个包含两条记录的“专利数据库”的 XML 格式的数据文件：

```

<?xml version="1.0" encoding="gb2312"?>
<!DOCTYPE TRS [
<!ELEMENT TRS (REC)*>
<!ELEMENT REC (公告日, 分类号, 公告号, 申请号, 申请日, 颁证日, 申请人, 地址, 设计人, 专利代理机构, 代理人, 摘要)>
<!ELEMENT 公告日 (#PCDATA)>
<!ELEMENT 分类号 (#PCDATA)>
<!ELEMENT 公告号 (#PCDATA)>
<!ELEMENT 申请号 (#PCDATA)>
<!ELEMENT 申请日 (#PCDATA)>
<!ELEMENT 颁证日 (#PCDATA)>
<!ELEMENT 申请人 (#PCDATA)>
<!ELEMENT 地址 (#PCDATA)>
<!ELEMENT 设计人 (#PCDATA)>
<!ELEMENT 专利代理机构 (#PCDATA)>
<!ELEMENT 代理人 (#PCDATA)>
<!ELEMENT 摘要 (#PCDATA)>
]>
<TRS>
<REC>
<公告日>1994. 01. 05</公告日>
<分类号>A01B49/02</分类号>
<公告号>CN_2151607Y</公告号>
<申请号>92243586. 3</申请号>
<申请日>1992. 12. 08</申请日>
<颁证日>1993. 10. 03</颁证日>
<申请人>刘殿富</申请人>
<地址>154552黑龙江省勃利县双河镇永平村</地址>
<设计人>刘殿富</设计人>
<摘要><![CDATA[
铲耧锄草耕作机
本实用新型公开了一种垄帮、苗间锄草，起垄深松破茬一次完成的铲耧锄草耕作机，是通过牵引带动向位传轮的伞形齿与苗间锄草器立式转动轴的伞型齿相啮合，使苗间锄草器在苗间水平旋转，同时通过动力带动安装在顺梁上的垄帮锄草器、双犁头深松器和齐刃犁头一起工作，就可达到垄帮、苗间锄草，垄沟深松、破茬起垄一次完成的目的。苗间和垄帮锄草可通过调节杆调节入土深度，苗间锄草部可单独与各种牵引，悬挂式三铧犁配套使用，也可用畜力牵引。
]]></摘要>
</REC>
<REC>
<公告日>1994. 01. 05</公告日>
<分类号>A01C23/04</分类号>
<公告号>CN_2151608Y</公告号>
<申请号>93202243. X</申请号>
<申请日>1993. 01. 20</申请日>
<颁证日>1993. 09. 12</颁证日>
<申请人>刘忠厚</申请人>
<地址>161005黑龙江省齐齐哈尔市龙沙区农委</地址>
<设计人>刘忠厚</设计人>
<专利代理机构>齐齐哈尔市专利事务所</专利代理机构>
<代理人>叶仲刚</代理人>
<摘要><![CDATA[

```

恒压等浓度施肥罐

一种恒压等浓度施肥罐，由罐体、输水管、进水管、施肥管及各种控制阀、压力表和水表组成。在罐体内设置一隔板，将罐体分为上下罐，隔板上有一个气体连通管将上罐连通。当上罐加入化肥或农药溶液密封后，由进水管向下罐注入清水，罐中产生气压，使上罐内的化肥或农药溶液经过施肥管注入输水管。本装置具有施肥均匀度好，肥浓度变化可控制，便于自动化实施，成本低且运行费用少，可制成固定式或移动式。

]]></摘要>

</REC>

</TRS>

TRIS 目前只检查 XML 的 DTD 中 DOCTYPE 以及 DOCTYPE 的元素两项，忽略 DTD 中的实体定义。XML 的 DTD 部分可以省略，缺省的 DOCTYPE 为 TRIS，缺省的 DOCTYPE 的元素为 REC。

DOCTYPE 的元素 (REC) 支持三个属性：SC 或 SecurityCode；UG 或 UserGroup；PWD 或 Password。这三个属性用于对记录的加密。

8. 数据记录文件中注释

在 TRIS 数据记录文件的“标准格式”、“字段内部标识号标记格式”和“字段顺序号标记格式”中可以插入注释信息。注释信息实际上是一个标记为单字节的星号 ‘*’ 的虚字段的值。格式为：<*>=注释信息。

B.2 批处理数据文件

对数据库进行加载时，数据文件既可以是单个 TRIS 数据记录格式（“标准格式”，或“字段内部标识号标记格式”，或“字段顺序号标记格式”，或“无字段标记格式”，或“全文格式”之一）文件，或者单个 ISO2709 格式或 XML 格式文件；也可以是一批相同格式（上述 7 种数据格式文件之一）的数据文件。

批处理数据文件实际上是存储一批数据记录文件名的文件。其基本规则是：

- 批处理数据文件是一个文本文件。文件的每一自然行存储一个数据记录文件名。
- 数据记录文件名中既可以包含路径，也可以不包含路径。如果不包含路径，并且在当前工作目录下也不存在指定的数据记录文件，则系统将取批处理数据文件的路径作为数据记录文件的缺省路径。其中路径的子目录分隔符没有 ‘\’ 和 ‘/’ 之分，这两种符号系统能够自动纠正。
- 文件中所列的全部数据记录文件必须具有相同数据记录格式。

下面是一个批处理数据文件的例子：

```
D:\DATA\人民日报\97ALL.TRS
D:\DATA\人民日报\9801.TRS
D:\DATA\人民日报\9802.TRS
D:\DATA\人民日报\9803.TRS
```

```
D:\DATA\人民日报\9804.TRS
D:\DATA\人民日报\9805.TRS
D:\DATA\人民日报\9806.TRS
D:\DATA\新华社\9806.TRS
D:\DATA\人民日报\9807.TRS
D:\DATA\新华社\9807.TRS
D:\DATA\人民日报\9808.TRS
D:\DATA\新华社\9808.TRS
D:\DATA\人民日报\9809.TRS
D:\DATA\新华社\9809.TRS
D:\DATA\人民日报\9810.TRS
D:\DATA\新华社\9810.TRS
D:\DATA\人民日报\9811.TRS
D:\DATA\新华社\9811.TRS
D:\DATA\人民日报\9812.TRS
D:\DATA\新华社\9812.TRS
D:\DATA\新华社\99ALL.TRS
```

B.3 加载多媒体数据

1. 加载 DOCUMENT 字段的多媒体数据

DOCUMENT 字段的值一般是文本数据，可直接放在 6 种数据记录文件中。除此之外，DOCUMENT 字段还可以接受其他数据格式类型的数据作为字段值，诸如 HTML、WORD 等等。由于这些类型的数据本身是格式化的，不能当作文本数据直接放在数据记录文件中，往往需要将其置于单独的文件中，然后在数据记录文件中再引用这些文件名。引用的方法是在文件名的前面加上字符 ‘@’ 后作为字段值放在数据记录文件中。

DOCUMENT 字段的值一般是需要对其进行索引的，但是 TRS 目前只能对文本格式和 HTML 格式的数据进行自动索引。至于其他格式类型的数据，如果需要建立索引，则必须事先将其转换成文本格式或 HTML 格式的数据文件，并与原数据文件一并提供给数据记录文件引用。由原数据文件通过格式转换得到的文件称为 DOCUMENT 字段值的附加文件。

综合起来，DOCUMENT 字段的多媒体数据的加载有以下三种格式：

- 简单格式： @Source_Name^Format^Store
- 典型格式： @Source_Name | Affix_Name^Format^Store
- 完全格式： @Source_Name^Format^Store | Affix_Name^Format^Store

其中：

- 符号 ‘@’ 为单字节 ASCII 码，是字段值文件引用的引导符。

- 符号 ‘|’ 为单字节 ASCII 码，是字段值文件与其附加文件之间的分隔符。
- 符号 ‘^’ 为单字节 ASCII 码，是字段值文件及其附加文件的控制符引导符。字段值文件及其附加文件的控制符有两种类型：数据格式控制符，即格式中的 Format 符号；存储方式控制符，即格式中的 Store 符号。

在被引用的字段值文件及其附加文件的文件名中，既可以包含路径，也可以不包含路径，系统能够自动确定其正确的路径。方法是：

- 直接检查文件名所对应的文件是否存在，如果存在，则不对文件名做任何修改，检查结束，否则继续。
- 检查文件名中是否存在路径，如果存在，则进行错误处理，否则继续。
- 检查控制文件中是否设置了 DATA_PATH_DEFAULT 变量，如果没有，则进行错误处理，否则继续。
- 检查控制文件的 DATA_PATH_DEFAULT 变量所指定的目录下是否存在这个文件，如果不存在，则进行错误处理，否则继续。
- 确定 DATA_PATH_DEFAULT 变量值为文件的路径，检查结束。

2. 数据格式控制符

数据格式控制符用来描述一个文件的具体数据格式，在 TRS 系统中表现为一部词典，用户可以对这部词典进行必要的增补或修改，以达到实际使用的要求。对于同一个文件，不能同时使用多个数据格式控制符来控制。

TRS 系统最基本的数据格式控制符有：

控制符	数据格式	文件后缀
BIT	不确定格式的二进制数据	bit
TEXT	文本数据格式	txt
WORD	Microsoft Word文件格式	doc
WS	WordStar文件格式	ws
WP	WordProfect文件格式	wp
XCL	Microsoft Excel文件格式	xls
WPS	金山WPS文件格式	wps
S2	方正、华光二扫文件格式	s2
DOC	不确定格式的文档文件格式	doc
AUDIO	不确定格式的音频文件格式	aud
IMAGE	通用图像文件格式	img
VIDEO	不确定格式的视频文件格式	vid
AVI	AVI动画文件格式	avi
MPEG	MPEG压缩动画文件格式	mpg
WAVE	Wave音频文件格式	wav
MIDI	MIDI音频文件格式	mid
CDA	CD音频文件格式	cda
PS	PostScript文件格式	ps

GIF	GIF图像文件格式	gif
TIF	TIF图像文件格式	tif
PCX	PCX图像文件格式	pcx
BMP	BMP图像文件格式	bmp
JPEG	JPEG图像文件格式	jpg
HTML	HTML超文本文件格式	htm;html
PDF	Adobe PDF文件格式	pdf
PPT	Microsoft Power Piont文件格式	ppt;pps
RTF	Rich Text文件格式	rtf
XML	XML文件格式	xml
MP3	MP3文件格式	mp3

数据格式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件或其附加文件的数据格式的步骤如下：

- 步骤1：检查控制文件中 FILE_SUFFIX_PRIOR 变量的值是否被设置成 TRUE，如果是，则继续步骤 2，否则转到步骤 4。
- 步骤2：检查文件是否有文件名后缀，如果有，则继续步骤 3，否则转到步骤 4。
- 步骤3：通过文件名后缀，获取其对应的数据格式控制符。如果文件名后缀存在对应的数据格式控制符，则该数据格式控制符就确定了字段值文件或其附加文件的数据格式，否则进行出错处理。
- 步骤4：默认该字段值文件或其附加文件为文本数据。

3. 存储方式控制符

存储方式控制符用来控制字段值文件及其附加文件在数据库中的存储方式，如 CLASS 和 DUMB 等等（不区分大小写）。对于同一个文件，不能同时使用多个存储方式控制符来控制。

TRS 系统所支持的存储方式控制符有：

控制符	存储方式
ALONE	以独立的文件形式存储
CLASS	将文件内容集中存储到数据库的按字段类型分类的文件中
URL	文件名为一个URL，在数据库中存储这个URL
LINK	文件名为一个URL，在数据库中存储这个URL（与URL相同）
BASE	说明一个URL为HTML文件中BASE元素的HREF属性值
DUMB	只存储文件的数据格式等信息，不存储数据

除此之外，系统内部还有一种缺省的存储方式，即集中存储方式。这种方式是将所有数据全部存储在数据库的主文件中。

BASE 控制符只在 DOCUMENT 字段的“典型格式”和“完全格式”两种加载格式中有效，在“简单格式”或 BIT 字段中出现时，实际上将忽略被控制的 URL。BASE 控制符

只对 HTML 数据格式的附加文件有效，它表示将被控制的 URL 自动插入到 HTML 文件中作为其 BASE 元素（如果原文件中已经存在 BASE 元素，则不作任何改动），被控制的 URL 本身不存入数据库。

存储方式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件的存储方式的步骤如下：

步骤1： 检查控制文件中 CLASS_DOCUMENT_STORE 变量的值是否被设置成 TRUE，如果是，则按分类方式存储字段值文件，否则转到步骤 2。

步骤2： 按系统缺省的存储方式来存储字段值文件。

在省略存储方式控制符的情况下，确定一个字段值附加文件的存储方式的步骤如下：

步骤1： 检查控制文件中 WAIVE_DOCUMENT_REFERER 变量的值是否被设置成 TRUE，如果是，则按 DUMB 方式字段值附加文件，否则转到步骤 2。

步骤2： 检查控制文件中 CLASS_DOCUMENT_STORE 变量的值是否被设置成 TRUE，如果是，则按分类方式存储字段值附加文件，否则转到步骤 3。

步骤3： 按系统缺省的存储方式来存储字段值附加文件。

4. 加载格式说明

DOCUMENT 字段多媒体数据的加载有三种格式，下面对这三种格式进行详细说明。

● 简单格式 @Source_Name^Format^Store

在这种格式中，Format 可以是所有数据格式控制符之一，Store 可以是所有存储方式控制符之一。

当存储方式控制符为 URL 时，如果 Source_Name 对应的数据格式是 TEXT 或者 HTML，则系统将自动从远程获取该 URL 的所对应的实际文件作为其附加文件，以便建立索引。系统以 URL 的方式存储 Source_Name 本身，但并不存储这个临时的附加文件。

系统能够自动识别一个 URL，即系统默认以不区分大小写的字符串“www.”或“http://”或“//”引导的 Source_Name 就是一个 URL，另外在 Win32 下以“\”引导的 Source_Name 也是一个 URL（网络邻居下的资源）。这样，如果存储方式控制符不是 URL，而 Source_Name 给出的却是一个 URL，则系统将自动从远程获取该 URL 的所对应的实际文件来替换 Source_Name，然后完成数据的加载。

需要说明的是：系统能够自动从远程获取的 URL 目前仅限于在 Win32 下以“\”引导的“网络邻居”上的 URL 资源，还不支持以“http://”和“//”引导的 Internet 上的 URL 资源。

- **典型格式** @Source_Name | Affix_Name^Format^Store

在这种格式中，Format 可以是所有数据格式控制符之一，Store 可以是所有存储方式控制符之一。并且，数据格式控制符和存储方式控制符的控制对象都是 Source_Name，而不是 Affix_Name。

如果存储方式控制符为 URL，则系统存储 Source_Name 本身，否则当系统自动识别 Source_Name 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Source_Name，然后完成数据的加载。

附加文件 Affix_Name 的数据格式的确定规则是：如果 Affix_Name 没有文件名后缀，则默认其数据格式为 TEXT，否则进一步判断这个文件名后缀是否对应了一个数据格式，如果有这样的数据格式，则 Affix_Name 的数据格式即为该格式，否则进行错误处理。

附加文件 Affix_Name 的存储方式省略存储方式控制符的情况下确定一个字段值附加文件的存储方式的步骤进行确定。

与 Source_Name 一样，如果系统自动识别 Affix_Name 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Affix_Name，然后完成数据的加载。

需要说明的是：系统能够自动从远程获取的 URL 目前仅限于在 Win32 下以“\\”引导的“网络邻居”上的 URL 资源，还不支持以“http://”和“//”引导的 Internet 上的 URL 资源。

- **完全格式** @Source_Name^Format^Store | Affix_Name^Format^Store

在这种格式中，Source_Name 和 Affix_Name 具有各自的数据格式控制符和存储方式控制符。控制 Source_Name 和 Affix_Name 的 Format 可以是所有数据格式控制符之一；控制 Source_Name 的 Store 可以是所有存储方式控制符之一；但控制 Affix_Name 的 Store 只能是 ALONE、CLASS 和 DUMB 三种存储方式控制符，URL 或 LINK 不能用来控制 Affix_Name。

如果 Source_Name 的存储方式控制符为 URL，则系统存储其本身，否则当系统自动识别 Source_Name 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Source_Name，然后完成数据的加载。

与 Source_Name 一样，如果系统自动识别 Affix_Name 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Affix_Name，然后完成数据的加载。

需要说明的是：系统能够自动从远程获取的 URL 目前仅限于在 Win32 下以“\\”引导的“网络邻居”上的 URL 资源，还不支持以“http://”和“//”引导的 Internet 上的 URL 资源。

5. 加载 BIT 字段的多媒体数据

对于 BIT 字段，虽然可以接受直接放在 TRS 数据记录文件中的文本数据作为字段值，但可作为字段值的其它格式化的数据是不能直接放在 TRS 数据记录文件中的，所以 DOCUMENT 字段引用文件名作为字段值的方法同样适合加载 BIT 字段的值。

BIT 字段对文件的引用格式要比 DOCUMENT 字段简单得多，只有一种形式：

```
@Source1_Name^Format^Store; Source2_Name^Format^Store; ...
```

其中：

- 符号 '@'、'^' 'Format' 以及 'Store' 与 DOCUMENT 字段的引用格式中的相应符号具有相同的意义。
- 分号 ';' 为单字节 ASCII 码，是多个引用文件之间的分隔符，即多值字段值之间的分隔符，也可以用其他字符或者字符串来代替。但不管用什么作为多值之间的分隔符，都要求控制文件中 MULTI_VALUE_SEPARATER 变量的值必须与其相一致。

格式中 Format 可以是所有数据格式控制符之一。数据格式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件的数据格式的步骤与 DOCUMENT 字段基本相同，只是字段值文件的默认数据格式为 BIT 格式，而不是 TEXT 格式。

格式中 Store 可以是所有存储方式控制符之一。存储方式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件的存储方式的方法是：如果控制文件中 CLASS_BIT_STORE 变量的值不为 TRUE，则按系统缺省的存储方式来存储字段值文件，否则按分类方式存储字段值文件。

如果 Source_NameX 的存储方式控制符为 URL，则系统存储其本身，否则当系统自动识别 Source_NameX 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Source_NameX，然后完成数据的加载。

与 DOCUMENT 字段一样，在被引用的字段值文件名中，既可以包含路径，也可以不包含路径，系统能够自动确定其正确的路径，确定方法与 DOCUMENT 字段完全一样。

需要说明的是：系统能够自动从远程获取的 URL 目前仅限于在 Win32 下以 “\\” 引导的“网络邻居”上的 URL 资源，还不支持以 “http://” 和 “//” 引导的 Internet 上的 URL 资源。

下面是一个加载多媒体数据的例子：

```
<REC>  
<正文>= @file1.txt  
<REC>  
<正文>= @\\server\data\file2.htm ^URL  
<REC>
```

```

<正文>= @http://server/data/ | file3.html ^BASE
<REC>
<正文>= @http://server/data/file4.doc ^WORD^CLASS
<REC>
<正文>= @file5.doc | file5.txt ^WORD^CLASS
<REC>
<正文>= @file6.doc ^CLASS | file6.txt ^DUMB
<REC>
<正文>= @//server/data/file7.doc | file7.txt ^URL
<REC>
<图片>= @file8.doc^CLASS; file9.txt
<REC>
<图片>= @\\server\data\file10.doc ^URL; www.server.com\data\file11.html

```

这里，文件路径中的子目录分隔符没有 ‘\’ 和 ‘/’ 之分，这两种符号系统能够自动纠正。但对于一个 URL，其路径中的子目录分隔符系统不会进行纠正，必须使用实际的目录分隔符。

B.4 控制文件

控制文件是指存放那些用以控制数据加载过程的变量配置文件。这个文件是在进行数据加载前由用户编写的。

控制文件的一些基本特性是：

- 控制文件是一个 ASCII 文本文件。
- 一行只能指定一个变量，其格式为：变量名 = 变量值。其中“变量名”、“=”不能省略，“变量名”和“变量值”的字符串不包括其首尾的空白字符。如果“变量值”省略，则等同于该变量未出现在控制文件中。
- 变量在文件中的顺序无关。
- 变量名不区分大小写。除了特别说明的外，变量值也不区分大小写。
- 文件中可以只包含一部分变量。文件中未包含的变量，系统将自动取其缺省值。
- 以单字节字符 ‘#’ 开始的行为注释行。注释行可以放在任何位置，但必须是单独一行。
- 如果用户在数据加载时未指定控制文件，系统将取所有变量的缺省值。

控制文件中的变量包括：

变量名	说明	缺省值
RECORD_FILE_FORMAT	数据记录文件的格式	标准
BIT_FORMAT_DEFAULT	数据记录文件中BIT字段值文件的缺省数据格式名	BIT
HTML_BASE_DEFAULT	HTML数据文件中BASE元素的缺省HREF属性值	无
DATA_PATH_DEFAULT	数据记录文件中字段值所引用的数据文件的缺省路径	当前数据文件
DATE_CENTURY_DEFAULT	数据记录文件中DATE字段值的年份的缺省世	当前世

	纪数	纪数
IGNORE_RECORD_CRYPT	是否忽略数据记录文件中的记录加密描述符的开关	FALSE
FILE_SUFFIX_PRIOR	当数据记录文件中所引用的字段值文件未指定数据格式控制符时,用文件名的后缀来确定其数据格式的开关	TRUE
SYNC_CREATE_INDEX	在加载数据的同时建立索引的开关	TRUE
CLASS_BIT_STORE	是否分类存储BIT字段值的开关	TRUE
CLASS_DOCUMENT_STORE	是否分类存储DOCUMENT字段值的开关	TRUE
WAIVE_DOCUMENT_REFERER	在数据记录文件中, DOCUMENT字段值的附加文件是否存储在数据库中的开关	FALSE
WAIVE_OEMFILTER_TEXT	是否丢弃由OEM过滤器所生成的文本数据的开关	TRUE
KEEP_TEXT_ACTUAL	是否保持PHRASE和DOCUMENT字段值文本中首尾空白行的开关	TRUE
CHECK_TEXT_ITERATIVE	是否允许在数据记录文件中同一个PHRASE或DOCUMENT字段在同一个记录中重复出现的开关	FALSE
MULTI_VALUE_SEPARATER	数据记录文件中多值字段DATE、NUMBER、CHAR、BIT的各个字段值之间的分隔符	单字节 ','
START_RECORD_FROM	在数据记录文件中数据加载的起始记录号	1
SECTION_TAG_PREFIX	TRS标记引导符前缀	无
PROCESS_RECORD_NUMBER	需要加载的最大记录数	全部
MAX_COMMIT_NUMBER	加载时分批提交的最大记录数	5000
MAX_ERROR_NUMBER	加载时允许错误记录的最大个数	最大
COLUMN_NAME_SEQUENCE	需要加载数据的字段名列表	全部
XML_SEGMENT_MARK	XML数据作为一个字段值入库时,过滤器进行自动分段的标记	无
XML_NEWLINE_MARK	XML数据作为一个字段值入库时,过滤器进行自动换行的标记	无
XML_HITSHOW_MARK	XML数据作为一个字段值入库后,检索命中点的指示标记	无
XML_FILTER_CDATA	XML数据作为一个字段值入库时,过滤器是否处理标记中的内容	FALSE
XML_OVERWRITE_PI	XML数据作为一个字段值入库时,是否需要覆盖数据中已有的关于过滤器的处理指令	FALSE
COLUMN_AUTO_MATCH	是否允许自动匹配入库字段	FALSE
NAMIZE_ISO2709_FIELD	是否以ISO2709格式的目录标记(以及数据标识符)映射入库字段	FALSE

在控制文件中,并不是所有的控制变量都必须出现,实际上只需给出那些实际变量值与其缺省值不一致的控制变量。以下各节将对每个变量进行详细说明。

● RECORD_FILE_FORMAT

该变量指明数据记录文件的组织格式。TRS 支持的数据记录文件格式包括:

名称	数据记录格式
----	--------

STANDARD	标准格式
IDENTISTD	字段内部标识号标记格式
SIMPLESTD	字段顺序号标记格式
SEQUENCE	无字段标记格式
FULLTEXT	全文格式
ISO2709	ISO2709国际标准格式
XML	XML格式

该变量的值必须为上述 7 种格式名之一。

如果控制文件中未包含该变量，则系统取其缺省值，即 STANDARD。

● BIT_FORMAT_DEFAULT

该变量指明数据记录文件中，BIT 字段引用的字段值文件的缺省数据格式控制符。该变量的值必须是一个当前系统中数据格式词典的一个合法数据格式控制符。如果控制文件中未包含该变量，则系统取其缺省值，即 BIT。

● HTML_BASE_DEFAULT

该变量指明 DOCUMENT 字段引用的 HTML 格式的字段值文件中 BASE 元素的缺省 HREF 属性值。该变量只对 DOCUMENT 字段的 HTML 字段值有效，显式 BASE 控制符优先于该变量。

该变量系统没有缺省值。

● DATA_PATH_DEFAULT

该变量指明数据记录文件中，DOCUMENT 和 BIT 字段引用的字段值文件的缺省路径。该变量的值必须是一个当前用户有权访问的合法的目录。缺省路径名是大小写敏感的，并且路径的子目录分隔符没有 ‘\’ 和 ‘/’ 之分，这两种符号系统能够自动纠正。

如果控制文件中未包含该变量，则系统取其缺省值，即数据记录文件所在的路径。对于批处理数据加载，其缺省值总是当前正在被处理的数据记录文件所在的路径。因此，该变量总是设置了有效值。

● DATE_CENTURY_DEFAULT

TRS 支持多种日期字段的入库格式。有关日期字段的入库格式详见下一节。

有的日期格式允许只给出两位数表示年份（如 98）。在这种情况下，需要确定该年份所在的世纪数，通过该变量即可设置这一世纪数。如将变量的值设置为 19，则年份 98 表示 1998 年。如果设置变量的值为 0，则表示日期格式中的年份为绝对年份，如 98 即表示公元 98 年。该变量的值必须是一个数字。

如果控制文件中未包含该变量，则系统取其缺省值，即当前世纪数。该变量只对那些小于 100 的年份有效。必须注意的是：如果年份（如 80）大于当前实际年份（如 07），则默认位是上一世纪的年份（如 1980），而不是本世纪的年份（如 2080）。

- **IGNORE_RECORD_CRYPT**

TRIS 支持对数据记录进行加密。

如果数据记录文件中有加密描述的记录，而又想在本次入库时忽略这些加密描述，即不对记录进行加密，则可通过设置这一变量来完成。如果将变量值设置为 **TRUE**（或 **YES**、或 **ON**、或 **1**），则表示忽略所有记录的加密描述；如果变量值设置为 **FALSE**（或 **NO**、或 **OFF**、或 **0**），则表示对那些有加密描述的记录进行加密。

如果控制文件中未包含该变量，则系统取其缺省值，即 **FALSE**。

- **FILE_SUFFIX_PRIOR**

在 TRIS 数据记录文件中，支持以引用文件名的形式给出 **DOCUMENT** 和 **BIT** 字段值，同时支持对这些文件显式地指定其数据格式。

当未指定数据格式，而又需要系统通过文件名的后缀来自动识别其文件的数据格式时，可通过设置该变量来触发系统的自动识别功能。如果将变量值设置为 **TRUE**（或 **YES**、或 **ON**、或 **1**），则表示可由文件后缀来确定其数据格式；如果变量值设置为 **FALSE**（或 **NO**、或 **OFF**、或 **0**），则表示按字段类型确定其缺省数据格式。**DOCUMENT** 字段值的缺省数据格式为 **TEXT**，**BIT** 字段值的缺省数据格式为 **BIT_FORMAT_DEFAULT** 变量的值。

如果控制文件中未包含该变量，则系统取其缺省值，即 **TRUE**。该变量只对那些没有显式指定其数据格式的字段值文件有效。

- **SYNC_CREATE_INDEX**

TRIS 有两种方式来创建数据库的索引：一种是在加载数据本身的同时建立数据的索引，即同步索引方式；另一种是先只加载数据本身，以后再单独建立数据的索引，既异步索引方式。

如果将变量值设置为 **TRUE**（或 **YES**、或 **ON**、或 **1**），则表示以同步方式建立索引；如果变量值设置为 **FALSE**（或 **NO**、或 **OFF**、或 **0**），则表示以异步方式建立索引。

如果控制文件中未包含该变量，则系统取其缺省值，即 **TRUE**。

- **CLASS_BIT_STORE**

由于 **BIT** 字段的数据一般都是数据量较大的诸如声音、图像以及动画之类的多媒体数据，对于 **BIT** 字段的数据，TRIS 允许将其存储到数据库的 **BIT** 分类文件中，这样就可以缩小数据库主文件的规模，从而有利于提高检索结果的概览速度。

在数据记录文件中，当以引用文件的形式给出字段值时，如果对这些文件没有显式地定义相应的存储方式控制符，而又要求系统按分类方式存储这些字段值，则可通过设置该变量实现这一功能。该变量同样能够控制那些直接放在数据记录文件中文本类型的字段值的存储方式。

如果将变量值设置为 **TRUE**（或 **YES**、或 **ON**、或 **1**），则表示将字段值存储到数据库

的 BIT 分类文件中；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示将字段值集中存储到数据库的主文件中。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。该变量只对那些没有明确指定存储方式的 BIT 字段的字段值有效。

● CLASS_DOCUMENT_STORE

虽然 DOCUMENT 字段的数据的数据量一般不是很大，但通常不用于概览，所以从有利于提高检索结果的概览速度的角度出发，TRS 也允许将其存储到数据库的 DOCUMENT 分类文件中，以缩小数据库主文件的规模，达到加速概览的目的。

在数据记录文件中，当以引用文件的形式给出的字段值时，如果没有对这些文件显式地定义相应的存储方式控制符，而又要求系统按分类方式存储这些字段值，则可通过设置该变量实现这一功能。该变量同样能够控制那些直接放在数据记录文件中文本类型的字段值的存储方式。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示将字段值存储到数据库的 DOCUMENT 分类文件中；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示将字段值集中存储到数据库的主文件中。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。该变量只对那些没有明确指定存储方式的 DOCUMENT 字段的字段值有效。

● WAIVE_DOCUMENT_REFERER

为了完成对非 TEXT 和 HTML 格式的数据进行索引，TRS 要求外部为其提供一个 TEXT 和 HTML 格式的附加数据文件。对于那些没有明确指定存储方式的附加文件，其存储方式系统将参考该变量和 CALSS_DOCUMENT_SOTRE 变量的设置情况综合确定。

如果将该变量的值设置为 TRUE（或 YES、或 ON、或 1），则表示数据库中不存储这个附加文件；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示数据库中存储这个附加文件，存储方式由 CALSS_DOCUMENT_SOTRE 变量的设置确定。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。该变量只对那些没有明确指定存储方式的 DOCUMENT 字段值的附加文件有效。

● WAIVE_OEMFILTER_TEXT

TRS 支持配置 OEM 过滤器自动抽取 OFFICE 文档的文本数据以建立索引。该变量确定将这些文本数据是否存储到数据库中。

如果将该变量的值设置为 TRUE（或 YES、或 ON、或 1），则表示数据库中不存储这些文本数据；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示数据库中存储这些文本数据，存储方式与源数据的相同。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。

- **KEEP_TEXT_ACTUAL**

对于 PHRASE 和 DOCUMENT 的文本类型字段值而言，位于其首尾的空白行是无用的，如果需要在数据加载时将它们去掉，则可通过设置该变量来触发系统完成对首尾空白行的过滤。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示保持字段值文本的原样；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示过滤字段值文本的首尾空白行。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。

- **CHECK_TEXT_ITERATIVE**

由于 PHRASE 和 DOCUMENT 字段支持全文检索，所以这两个字段不能有多个值。但在数据记录文件中，可能会由于各种各样的原因，使得在一个记录中两个字段出现多次。如果出现这种情况，系统缺省的处理方式是只保留最后出现的字段值。通过设置该变量可以强制系统进行检查，当发现这种情况时进行错误处理。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示一个记录中如果多次出现 PHRASE 或 DOCUMENT 字段，则认为该记录有问题，并进行错误处理；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示按系统的缺省方式进行处理。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。

- **MULTI_VALUE_SEPARATOR**

该变量用来给出数据记录文件中，DATE、NUMBER、CHAR 以及 BIT 多值字段的各个值之间所采用的分隔符。多值分隔符既可以是单个字符，也可以是多个字符（如全角字符 ‘；’），但最多为 8 个字符，并且大小写敏感。空格不能作为多值分隔符。

如果控制文件中未包含该变量，则系统取其缺省值，即单字节分号 ‘;’。

- **START_RECORD_FROM**

该变量指定从数据记录文件中第几个记录开始加载。对于批处理方式，该变量只对第一个数据记录文件有效，对于其他数据记录文件，总是从第一个记录开始加载。

如果控制文件中未包含该变量，则系统默认从第 1 个记录开始加载。

- **SECTION_TAG_PREFIX**

该变量指定 TRS 格式文件中 TRS 标记引导符的前缀。TRS 支持标记引导符前缀的目的就是要排除数据中的干扰，所以标记引导符前缀应该选择一个在数据中不会出现的字符序列，至少不会在一行的开头出现。

该变量系统没有缺省值。

- **PROCESS_RECORD_NUMBER**

该变量指定要处理的总记录数，包括加载成功的记录和错误记录。如果控制文件中未

包含该变量，则系统默认加载给出的所有数据记录文件中的全部记录。

● MAX_COMMIT_NUMBER

一个数据记录文件往往包含许多记录，系统总是将这些记录分批地加载到数据库中，该变量就是用来设定每批加载的最大记录数的。每一批实际加载多少个记录，系统内部还需要根据机器内存的大小综合考虑，但不会超过这一数。

如果控制文件中未包含该变量，则每一批实际加载的记录数由系统完全根据机器内存的大小自动确定。

● MAX_ERROR_NUMBER

该变量指定：如果在加载过程中所遇到的错误记录数超过了这一数，则结束加载过程。如果控制文件中未包含该变量，则系统默认加载过程不受错误记录数的控制。

● COLUMN_NAME_SEQUENCE

该变量用来指定需要加载数据的字段名（或别名）列表，字段名之间用单字节的分号‘;’、逗号‘,’或空格分隔。在下列几种情况下需要设置该变量：

- 只需从数据记录文件中加载部分数据库字段的值。
- 数据记录文件的记录中存在不属于当前加载的数据库的字段。
- 数据记录文件的格式为 SIMPLESTD 或 SEQUENCE 或 ISO2709，而数据记录文件的记录中字段的顺序与数据库字段的物理顺序（即输出时的顺序）不一致。
- 数据记录文件格式为 FULLTEXT，而加载的目标字段不是数据库的最后一个 DOCUMENT 字段；或者数据库没有 DOCUMENT 字段，而加载的预期字段不是数据库的最后一个 PHRASE 字段。

对于同一个字段，如果这里给出的是字段名，则数据文件中只能用字段名作为“字段标记”；如果这里给出的是字段别名，则数据文件中不能用其它字段别名作为“字段标记”，但字段名始终可以“字段标记”；除非将变量 COLUMN_AUTO_MATCH 设置为 TRUE。

如果控制文件中未包含该变量，则系统取其缺省值为：如果数据记录文件不为 FULLTEXT，则取数据库的所有字段，并按其物理顺序排列。否则，当数据库存在 DOCUMENT 字段时，取数据库的最后一个 DOCUMENT 字段；当数据库存在 PHRASE 字段，取数据库的最后一个 PHRASE 字段；其他情况下进行错误处理，不能进行数据加载。

● XML_SEGMENT_MARK

该变量用来指定 XML 数据作为一个字段值入库时，通知过滤器进行自动分段的标记，它必须是这个 XML 数据中的一个有效的以成对方式出现的标记。在这一标记对之间的数据为一自然段落。一般情况下，为了控制 XML 数据的显示效果，通常会在 XML 数据中插入必要的标记以使相应的解释器进行分段处理，这样的标记就可作为变量 XML_SEGMENT_MARK 的值，注意该标记的大小写是敏感的，最大长度是 60 个字符。

在一个段落中，过滤器将忽略所有的回车换行符，以避免一个词被 TRS 分词系统中的换段规则所打断而无法检索。

该变量系统没有缺省值。

下面举一个例子：

```
<paragraph id="d11">
```

```
    存储器可以存放最后一次接受的参考生物特征图像。存储器可以<br/>
    存放过去输入的多个的参考生物特征图像，其包括以前接受的生物特<br/>
```

```
    <!-- 1 -->
```

```
    <dp n="d1"/>
```

```
    征图像。确定器可以计算输入生物特征图像和各参考生物特征图像之<br/>
    间的相似程度，当所有计算的相似程度小于预定阈值时，确定器确定<br/>
    输入的生物特征图像可以接受。<br/>
```

```
</paragraph>
```

在这个例子中，标记“`paragraph`”是 XSL 用来进行换段控制的，所以我们就将 `XML_SEGMENT_MARK` 变量指向这个标记。如果不设置该变量，过滤器无论如何处理，其中第三个“特征”将不能被检索。

● XML_NEWLINE_MARK

该变量用来指定 XML 数据作为一个字段值入库时，通知过滤器进行自动换行的标记，它必须是该 XML 数据中的一个有效标记，否则过滤器将不会输出任何换行符（因为不可能遇到这个标记）。一般情况下，为了控制 XML 数据的显示效果，通常会在 XML 数据中插入必要的标记以使相应的 XML 浏览器进行换行，这样的标记就可作为变量 `XML_NEWLINE_MARK` 的值，注意该标记的大小写是敏感的，最大长度是 60 个字符。如 4.2.22 节例子中的“`br`”标记。

如果设置了变量 `XML_NEWLINE_MARK`，过滤器将忽略 XML 数据中的回车换行符，而在遇到由该变量所指定的标记时才进行自动换行，否则将以 XML 数据中原有的回车换行符作为过滤器的换行符。但正如 4.2.22 节所说明的，过滤器将忽略处在变量 `XML_SEGMENT_MARK` 所指定的段落标记中的换行标记，以及 XML 数据中原有的回车换行符。

与指定分段标记一样，指定换行标记也有避免一个词被 TRIS 分词系统中的换段规则所打断而无法检索的作用。

该变量系统没有缺省值。

● XML_HITSHOW_MARK

该变量用来指定 XML 数据作为一个字段值入库后，指示检索命中点的标记。这个标记有以下几个特点：

- 其大小写是敏感的，并且被浏览器所识别，最大长度是 60 个字符。
- 在入库前，XML 数据中并不存在这个标记。
- 在检索时，有 TRS 系统自动在输出的 XML 数据中插入。
- 标记以成对方式出现：在命中点之前插入起始标记，在命中点之前插入结束标记。
- 在起始标记中可能会出现一个名为“ID”的属性，其值为命中点的从“1”开始计数的序号。没有出现这个属性的起始标记，说明是一个命中点被其它标记所打断后的重新开始。属性值也有为“-1”的情况，参见 4.2.25。

该变量系统没有缺省值。

下面举一个例子：

```
<paragraph id="d11">
```

```
    存储器可以存放最后一次接受的参考生物<TRSHL ID="1">特征</TRSHL>图像。存储器可以
    <br/>存放过去输入的多个的参考生物<TRSHL ID="2">特征</TRSHL>图像，其包括以前接受的
    生物<TRSHL ID="3">特</TRSHL><br/>
```

```
    <!-- 1 -->
```

```
    <dp n="d1"/>
```

```
    <TRSHL>征</TRSHL>图像。确定器可以计算输入生物<TRSHL ID="4">特征</TRSHL>图像和各
    参考生物<TRSHL ID="5">特征</TRSHL>图像之<br/>间的相似程度，当所有计算的相似程度小
    于预定阈值时，确定器确定<br/>输入的生物<TRSHL ID="6">特征</TRSHL>图像可以接受。
    <br/>
```

```
</paragraph>
```

在这个例子中，“TRSHL”是由 XML_HITSHOW_MARK 变量指定标记。注意其中第三个“特征”命中点被其它标记所打断。“TRSHL”标记在 XSL 中可以用下面的方式进行处理（用红色反显）：

```
<xsl:template match="TRSHL">
```

```
    <font color="red"><xsl:value-of select="."/></font>
```

```
</xsl:template>
```

● XML_FILTER_CDATA

该变量用来指定 XML 数据作为一个字段值入库时，过滤器是否处理 CDATA 标记部分的内容。一般情况下，CDATA 数据不需要（或者说不能）进行索引，但有的 XML 数据中的 CDATA 部分也要求能被检索，有时还要求标记命中点。

- 如果将变量 XML_FILTER_CDATA 设置成 FALSE(或 NO、或 OFF、或 0),则表示不过滤 CDATA 数据中的任何标记,所有内容 & 标记都建立索引。检索结果输出时,如果其中有命中点,则由 XML_HITSHOW_MARK 变量所指定的命中点标记不是直接在 CDATA 数据中插入,而是在其结束符后插入一个空标记(标记之间没有任何数据),其“ID”属性的值为“-1”。
- 如果设置成 TRUE(或 YES、或 ON、或 1),则表示过滤 CDATA 数据中的所有标记,只索引 CDATA 数据中非标记部分。检索命中点的输出方式与取值为 FALSE 时相同。
- 如果设置成 HIT,则表示过滤 CDATA 数据中的所有标记,只索引 CDATA 数据中非标记部分。检索结果输出时,如果其中有命中点,则由 XML_HITSHOW_MARK 变量所指定的命中点标记直接在 CDATA 数据中的准确位置插入。
- 如果设置成 ROW,则表示不过滤 CDATA 数据中的任何标记,但去掉所有回车换行符 & 多余的空白符,全部有效内容 & 标记都在一行上进行索引。检索命中点的输出方式与取值为 FALSE 时相同。
- 如果设置成 ANY,则表示去掉 CDATA 数据中的全部内容,即任何内容(标记 & 非标记)都不建立索引。

该变量的缺省值为 FALSE。

● XML_OVERWRITE_PI

该变量用来指定 XML 数据作为一个字段值入库时,是否需要覆盖 XML 数据中已有的关于过滤器的处理指令。

为了能够在 XML 中准确指示检索命中点的位置,命中点跟踪器也需要使用过滤器完全一致的参数,所以 TRIS 系统会根据控制文件给出的各个相关变量,在 XML 数据入库时,自动插入一条 XML 处理指令。XML 过滤器以及跟踪器都将引用这条处理指令,其格式如下:

```
<?trs-parser ATTRIBUTE?>
```

ATTRIBUTE 由下列一个或多个分项组成:

- SegmentMark = “Mark”。其中“Mark”是由变量 XML_SEGMENT_MARK 指定的段落标记,大小写有关,最大长度是 60 个字符。没有缺省值。
- NewLineMark = “Mark”。其中“Mark”是由变量 XML_NEWLINE_MARK 指定的换行标记,大小写有关,最大长度是 60 个字符。没有缺省值。
- HitShowMark = “Mark”。其中“Mark”是由变量 XML_HITSHOW_MARK 指定的命中点指示标记,大小写有关,最大长度是 60 个字符。没有缺省值。
- FilterCDATA = “Mode”。其中“Mode”的取值包括:“no”、“yes”、“hit”、“row”以及“any”,分别对应于变量 XML_FILTER_CDATA 的取值:“FALSE”、“TRUE”、“HIT”、“ROW”和“ANY”。缺省值为“no”。其中“no”可用“FALSE”、“OFF”、“0”替代,“yes”可用“TRUE”、“ON”、“1”替代。

你也可以手工创建或修改这条指令,但必须注意 ATTRIBUTE 的书写要遵循 XML 的

规范，分项之间用空格分隔，给出的每个分项都必须是完整的。指令中，作为属性值的标记名是区分大小写的，其他字符则是大小写无关的。指令必须出现在 XML 的 DTD 部分之前，并只允许出现一次。下面是一个例子：

```
<?xml version="1.0" encoding="GB2312"?>

<?xml-stylesheet type="text/xsl" href="xsl0117.xsl"?>

<?trs-parser segmentmark=" paragraph" newlinemark="br" hitshowmark="trshl"?>

<!DOCTYPE patent SYSTEM "dtd0117.dtd">

.....
```

在一般情况下，如果过滤器发现 XML 数据中已经有了这条处理指令，无论其属性如何，都会优先使用指令中给出的属性参数，而忽略控制文件给出的变量值。但是，有的情况可能需要用控制文件中给出的变量值重新构造这条处理指令。

如果将变量 XML_OVERWRITE_PI 设置成 FALSE（或 NO、或 OFF、或 0），则表示不覆盖 XML 数据中已有的过滤器处理指令；如果将其设置成 TRUE（或 YES、或 ON、或 1），则表示重新构造并覆盖 XML 数据中已有的过滤器处理指令。

该变量的缺省值为 FALSE。

● COLUMN_AUTO_MATCH

为了方便，无论 COLUMN_NAME_SEQUENCE 变量设置与否，有时都需要 TRS 在数据入库时，能够根据数据中的实际给出的“字段标记”自适应地与数据库字段进行映射。这个功能只要将 COLUMN_AUTO_MATCH 变量设置成 TRUE（或 YES、或 ON、或 1）即可实现，此时 COLUMN_NAME_SEQUENCE 变量的唯一作用是指明哪些字段需要加载数据。但是这个功能有些情况会影响系统性能：在数据中存在错误的“字段标记”；同一个字段的多个别名混合使用，同时作为“字段标记”。因此一般情况不建议使用该功能。

如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示不允许 TRS 系统自适应地进行字段映射：当 COLUMN_NAME_SEQUENCE 变量被设置时，只有出现在列表中的“字段标记”所标记的数据才会被加载；当 COLUMN_NAME_SEQUENCE 变量缺省时，只有数据库的字段名所对应“字段标记”才是合法的。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。

● NAMIZE_ISO2709_FIELD

对于 ISO2709 格式的入库，缺省情况是：ISO2709 格式中的“目录项标记”对应于 COLUMN_NAME_SEQUENCE 变量（或缺省值）给出的字段列表中的字段所在的顺序号。通过 NAMIZE_ISO2709_FIELD 的设置可以改变这种映射关系。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示将 ISO2709 格式中的“目

录项标记”（如“010”）将对应于数据库字段的字段名或别名（如字段名或别名为“010”的字段）；同时还支持抽取数据域的各个子项分别单独入库，即将“目录项标记”（如“010”）与数据域的各个子项标识符（如“a”）进行结合，形成一个“子字段名”（如“010a”），如果在入库目标字段列表中存在该字段名，则对应的数据子项将单独加载到这个数据库字段中。另外，由于数据库的字段名是不区分大小写的，当子项标识符是一个大写字母时，必须按特殊规则处理：将“目录项标记”（如“010”）与数据子项标识符（如“A”）进行结合时用下划线进行连接（如“010_A”），因此其对应的数据库字段名或别名也必须遵循这一规则。

如果使用字段的别名，则必须在控制文件中用 `COLUMN_NAME_SEQUENCE` 变量列出所有需要加载数据的字段所对应的别名，除非设置 `COLUMN_AUTO_MATCH` 变量为 `TRUE`。

如果变量值设置为 `FALSE`（或 `NO`、或 `OFF`、或 `0`），则表示格式中的“目录项标记”域是该字段在控制文件字段列表中的序号（从 1 开始计数）。

如果控制文件中未包含该变量，则系统取其缺省值，即 `FALSE`。

下面是一个控制文件的例子：

```
#This is the control file of loading PeopleDaily database.
RECORD_FILE_FORMAT = STANDARD
BIT_FORMAT_DEFAULT = IMAGE
HTML_BASE_DEFAULT = http://192.9.200.1:6666/doc
DATA_PATH_DEFAULT = D:\DATA
DATE_CENTURY_DEFAULT = 19
IGNORE_RECORD_CRYPT = NO
FILE_SUFFIX_PRIOR = OFF
SYNC_CREATE_INDEX = YES
CLASS_BIT_STORE = TRUE
CLASS_DOCUMENT_STORE = TRUE
WAIVE_DOCUMENT_REFERER = FALSE
KEEP_TEXT_ACTUAL = FALSE
CHECK_TEXT_ITERATIVE = FALSE
MULTI_VALUE_SEPARATOR = ;
SECTION_TAG_PREFIX = $$TRSTAG$$
START_RECORD_FROM = 1
PROCESS_RECORD_NUMBER = 50000
MAX_COMMIT_NUMBER = 1000
MAX_ERROR_NUMBER = 1000
COLUMN_NAME_SEQUENCE = 日期;版次;版名;作者;标题;正文;图像
```

B.5 日期的入库格式

日期的入库格式有通用格式、英文格式、无分隔符格式和中文格式四种类型。

1. 通用格式

Year[/Month[/Day[/Hour[/Minute[/Second]]]]][A[M]]

[Year]/Month[/Day[/Hour[/Minute[/Second]]]]][A[M]]

[Year]/[Month]/Day[/Hour[/Minute[/Second]]][A[M]]

[Year]/[Month]/[Day]/Hour[/Minute[/Second]][A[M]]

[Year]/[Month]/[Day]/[Hour]/Minute[/Second][A[M]]

[Year]/[Month]/[Day]/[Hour]/[Minute]/Second[A[M]]

其中：

- Year 为四位表示的年(如 1998)或者二位表示的年(如 98)。
- Month、Day、Hour、Minute、Second 分别为二位或一位表示的月、日、时、分、秒。
- A/[M]可用 P/PM 替换。A/AM 表示上午，P/PM 表示下午，大小写无关。如果它们出现在格式中，则“Hour”不能却省。
- 单字节除号 ‘/’ 分隔符可分别用单字节的逗号 ‘,’、句号 ‘.’、冒号 ‘:’ 以及减号 ‘-’ 任意替代，在两个明确（不省略）的日期成份之间，空格也可作为分隔符。
- 用方括号引起的部分表示可以省略（模糊日期）。
- Year、Month、Day、Hour、Minute、Second 都可以用一个或多个单字节的问号 ‘?’、百分号 ‘%’ 替代，表示模糊日期。

2. 英文格式

Month[Day[,Year[Hour[:Minute[:Second]]]]][A[M]]

Month[Day],Year[Hour[:Minute[:Second]]][A[M]]

Month[Day],Year[Hour]:Minute[:Second][A[M]]

Month[Day],Year[Hour]:[Minute]:Second[A[M]]

[Day-]Month[-Year[Hour[:Minute[:Second]]]]][A[M]]

[Day-]Month[-Year[Hour]:Minute[:Second]][A[M]]

[Day-]Month[-Year[Hour]:[Minute]:Second][A[M]]

DayMonthYear[Hour[:Minute[:Second]]][A[M]]

DayMonthYear[[Hour]:Minute[:Second]][A[M]]

DayMonthYear[[Hour]:[Minute]:Second][A[M]]

其中:

- Year 为四位表示的年(如 1998)或者二位表示的年(如 98)。
- Month 为英文表示的月份: Jan. Feb. Mar. Apr. May Jun. Jul. Aug. Sept. Oct. Nov. Dec.; 或: January February March April May June July August September October November December。
- Day、Hour、Minute、Second 分别为二位或一位表示的月、日、时、分、秒。
- A/[M]可用 P/PM 替换。A/AM 表示上午, P/PM 表示下午, 大小写无关。如果它们出现在格式中, 则“Hour”不能却省。
- Year 与 Hour 之间用单字节空格分隔; ‘,’ ‘:’ ‘-’ 分别为单字节的逗号、冒号、减号。
- 用方括号引起的部分表示可以省略(模糊日期)。

3. 无分隔格式

Year[MonthDay[Hour[Minute[Second]]]]

YY[MonthDay]

其中:

- Year 为四位表示的年(如 1998); YY 为二位表示的年份(如 98)。
- Month、Day、Hour、Minute、Second 分别为二位表示的月、日、时、分、秒。
- 用方括号引起的部分表示可以省略(模糊日期)。

4. 中文格式

中文格式包括所有用全角字符表示的上述几种格式, 另外还支持如下的格式:

[Year 年][Month 月][Day 日][Hour 时][Minute 分][Second 秒][上午]

其中:

- Year 为四位表示的年(如 1998 或一九九八或壹玖玖捌)或者二位表示的年(如 98 或九八或玖捌)。其中数字既可为半角, 也可为全角。
- Month、Day、Hour、Minute、Second 分别为二位或一位表示的月、日、时、分、秒。
- 格式中“上午”可用“下午”替换。如果“上午”或“下午”出现在格式中, 则“Hour 时”不能却省。
- 用方括号引起的部分表示可以省略(模糊日期)。

另外, 还支持\$DATE 和\$TIME 特殊日期(不区分大小写)的入库, 其中前者表示当

前日期的“年月日”，后者表示当前日期的“年月日时分秒”。这里\$DATE和\$TIME还可以用作函数来表示一个相对于“现在”的日期，即：\$DATE(...)、\$TIME(...)，它们的参数都是一个正整数、或者负整数，\$DATE(...)的参数单位是天数，\$TIME(...)的参数单位是秒数。如：\$DATE(7)、\$DATE(-7)、\$DATE(+7)等价于\$DATE(7)；\$TIME(60)、\$TIME(-60)、\$TIME(+60)等价于\$TIME(60)；\$DATE(0)、\$DATE(-0)和\$DATE()等价于\$DATE，\$TIME(0)、\$TIME(-0)和\$TIME()等价于\$TIME。