

# 自定义中文输入法的自动安装

胡遇杰 宋军强 杨煜普 (上海交通大学自动化系 200030)

**摘要:** 本文给出了一种自定义中文输入法自动安装的方法。文章详细介绍了其工作原理和实现过程应注意的几个问题,最后介绍了利用本文所阐述的知识解决其他问题的两个实例。

**关键词:** 中文输入法 Windows API 注册表

在WINDOWS系统(包括WINDOWS 95、98、2000、NT4.0及NT5.0)中可以自定义输入法,这使得我们可以设计个性化的输入法,当然也可以修改别人的输入法使其个性化,从而大大提高汉字输入的效率。当我们设计了一个效率很高的输入法时,可能想将其发布给他人使用。这时最简单的做法是发布我们的输入法码表源文件(\*.txt),他人利用输入法生成器安装这一输入法。这一过程首先要求使用者的系统中装有输入法生成器,然而操作系统缺省安装时并不安装输入法生成器,因此许多情况下得首先安装输入法生成器。而安装输入法生成器又要求有WINDOWS系统盘(包括网络和硬盘)在手边。其次要求使用者懂得如何使用输入法生成器。以上两项要求无疑阻碍了自定义输入法的发布,这使得开发一个自定义输入法的自动安装程序成为必要。

众所周知,在WINDOWS系统中注册表有举足轻重的地位。通过对注册表的研究,我们找到了解决问题的方法。

## 1 WINDOWS 系统注册表简介

WINDOWS系统中注册表由键、键值和数据三部分组成。

### 1.1 键

WINDOWS注册表中的键数以千计,但根键只有六个:

- HKEY — CLASSES — ROOT

该键下面包括有关DDL和OLE的信息。

- HKEY — CURRENT — USER

该键下面包括登录到机器上的用户信息和有关应用程序的信息。

- HKEY — LOCAL — MACHINE

该键下面包括本地机与用户无关的硬件和软件的配

置信息。

- HKEY — USER

该键下面包括所有的用户说明信息。

- HKEY — CURRENT — CONFIG

该键下面包括用户程序的访问信息。

- HKEY — DYN — DATA

该键下面包括存储在RAM中的信息,这些信息均是动态的。

### 1.2 键值

注册表中的键值相当于一个变量,和数据相对应。每一个键值都有与它对应的数据。

### 1.3 数据

注册表中的数据相当于变量的值。数据有三种类型:字符串、二进制、DWORD型。

## 2 与自动安装程序相关的注册表信息

注册表中与自动安装程序相关的注册表键有

### 2.1 HKEY — USERS/. DEFAULT/Keyboard Layout/preload/.

这个键下的子键名为1、2、3...n(其中n为自然数),表示系统启动时自动装载的几种输入法。假设某系统只安装了智能ABC这一种中文输入法,则该键下有2个子键,子键名分别是1、2。1子键下的键值为“(默认)”,数据为00000409,表示的是英文输入法。2子键下的键值为“(默认)”,数据为E0040804,表示的是智能ABC输入法。

### 2.2 HKEY — USERS/. DEFAULT/software/microsoft/windows/Current Version/

这个键下的子键名为“智能ABC”以及其他系统自带的输入法。智能ABC子键中键值有“(默认)”(字符

串型)、“标调”(DWORD 型)、“纯笔形输入”(DWORD 型)、“词频调整”(DWORD 型)、“风格”(DWORD 型)、“双打键盘类型”(DWORD 型)、“双打提示”(DWORD 型)、“自动转换”(DWORD 型)。不同的设置上面的键值对应的数据不同。

### 2.3 HKEY — LOCAL — MACHINE/SYSTEM/Current ControlSet/control/keyboard layouts/.

该键下的子键为八位十六进制数。代表智能 ABC 的子键是“E0040804”键,其中中文输入法后面四位十六进制数都是 0804,而前面的四位十六进制数不同。

“E0040804”子键下有四个键值,其中第一个为空,对应的数据也为空。后面三个有效的键值是:“IME file”、“layout file”、“layout text”。它们是智能 ABC 输入法的三个重要信息,后文将进一步解释。

## 3 自动安装程序的实现

为便于后文的说明,现假设有一个待安装的自定义输入法名为“形音法”,其码表源文件为 xingyin.txt。

### 3.1 程序流程

(1) 利用输入法生成器从 xingyin.txt 文件生成 xingyin.mb 和 xingyin.ime 文件。

(2) 利用 Visual Studio 中微软的 compress.exe 程序将 xingyin.mb 和 xingyin.ime 文件压缩为 xingyin.mb\_ 和 xingyin.in\_ 文件。

(3) 在 HKEY\_LOCAL\_MACHINE/System/CurrentControlSet/Control/Keyboard Layouts 键下增加一名为“E08F0804”的子键。在“E08F0804”子键下增加下面三个键值:“IME file”、“layout file”、“layout text”。它们对应的数据分别是“xingyin.ime”、“kbdus.kbd”、“形音”。

(4) 在 HKEY\_USERS/DEFAULT/Software/Microsoft/Windows/CurrentVersion 键下增加一名为“形音”的子键。在此子键下创建如下表所示键值,其对应的数据和类型亦如下表所示。

(5) 在 HKEY\_USERS/.DEFAULT/keyboard layout/preload 键下添加一子键。名为 n+1(其中 n 为原来此键下的最大子键名)。然后在此子键下添加一空键值,对其赋值“E08F0804”。

(6) 把文件 xingyin.mb 和 xingyin.ime 复制到系统的 system 目录下(通常是 C:/WINDOWS/SYSTEM 目录)。

(7) 重新启动系统即可使用。

键值名	键值	类型
NULL(默认)	NULL	字符串
词语联想	1	DWORD
词语输入	1	WWORD
逐渐提示	1	DWORD
插空格	0	DWORD
光标跟随	1	DWORD
<SPACE>	1	DWORD
<ENTER>	1	DWORD
GB/GBK	1	DWORD

(8) 输入法卸载程序中只需删去上面所添加的信息即可。

### 3.2 程序设计中的注意事项

这个程序是用 Visual C++ 6.0 编制的。程序设计中应注意以下事项:

#### (1) 对系统类型进行判断

笔者发现 WINDOWS 98 和 WINDOWS 95 的码表源文件有一点点不同,因此,对两种系统所使用的 xingyin.mb 和 xingyin.ime 文件也不同,这要求自动安装程序能根据不同的系统使用不同的源文件。因此要求对系统进行判别。

#### (2) 键值和数据在 C 语言中的数据类型

如前文所述键值和数据有三种类型,分别是字符串、二进制及 DWORD 型。在 Visual C++ 中键值名有如下一些类型: REG\_BINARY、REG\_DWORD、REG\_DWORD\_LITTLE\_ENDIAN、REG\_DWORD\_BIG\_ENDIAN、REG\_EXPAND\_SZ、REG\_LINK、REG\_MULTI\_SZ、REG\_NONE、REG\_RESOURCE\_LIST、REG\_SZ

因此,在给某一键值赋值时要求正确地指出键值类型。笔者在调试程序时发现,在 Visual C++ 6.0 中如果类型为字符串型,使用 .REG\_EXPAND\_SZ 类型比使用 .REG\_SZ 类型更好。如果是 DWORD 型键值,使用 REG\_DWORD 一般能满足要求。

#### (3) 关键的 API 函数及用途

程序中关键 API 及其用途如下: GetVersionEx 用来取得系统版本信息; GetSystemDirectory 用来取得系统目录; ExitWindowsEx 用于重新启动系统; LZOpenFile、LZCopy 和 LZClose 用来完成被 compress.exe 压缩过的文件的复制。RegCreateKeyEx 用于打开和创建注册表中的键。

RegCloseKey 用于关闭打开的注册表中的键。  
RegSetValueEx 用于在特定键下创建键值及对其赋值。  
RegEnumKey 用于对某一键下的子键进行枚举。

### 3.2.4 确保系统启动时能正确装载输入法

安装程序在 HKEY\_USERS/.DEFAULT/keyboard layout/preload 键下添加某一子键时, 要求添加子键后的各子键名是从 1 开始的连续的自然数。同时, 卸载程序中删除 HKEY\_USERS/.DEFAULT/keyboard layout/preload 键下的某一子键时, 要对删除后的各子键的键名进行检查, 以保证删除某一子键后其各子键名是从 1 开始的连续的自然数。这可以通过对 HKEY\_USERS/.DEFAULT/keyboard layout/preload 键下的子键进行枚举来获得所需的信息。可以利用 RegEnumKey 来实现。

安装程序中可以用下面代码检测出 HKEY\_USERS/.DEFAULT/keyboard layout/preload 键下的子键个数。定义如下变量: int I; long IRET; HKEY hkey; DWORD dwDisposition; char szsubkey [256];

```
RegCreateKeyEx(HKEY_USERS, ".DEFAULT//  
keyboard layout//preload", 0, "  
", REG_OPTION_NON_VOLATILE, KEY_ALL_ACCESS, NULL,  
&hkey, &dwDisposition);
```

```
for(I=0; IRET==ERROR_SUCCESS; I++)
```

```
IRET=RegEnumKey(hkey, I, szsubkey, sizeof(szsubkey));
```

其中 I-1 就是已有的子键数目。

按以上步骤可以生成自动安装程序, 至此可将 xingyin.mb\_、xingyin.im\_ 以及安装程序一起发布, 使用者只需运行安装程序, 并按提示中步骤操作即可。这样就解决了前文提出的自定义输入法发布过程中不便安装的问题。

## 4 结束语

本文给出了一种自定义中文输入法的自动安装方法, 并详细分析了方法的实现要点和工作原理。理解本文提出的方法, 不仅可以加深对 WINDOWS 操作系统和中文输入法的理解, 而且可以利用这些知识解决一些系统中中文输入法的问题。比如说有的 WINDOWS 系统中虽然安装了几种输入法, 但是其中一种显示为码表文件坏或其他, 安装在其后的中文输入法也都不能使用。这时只要将注册表中 HKEY\_USERS/.DEFAULT/Keyboard Layout/preload/ 键下对应于码表文件坏的输入法的子键删去, 并将其他输入法的子键名重新按序命名, 使各子键名为从 1 开始的连续的自然数, 然后重新启动系统即可解决问题。更改 HKEY\_USERS/.DEFAULT/Keyboard Layout/preload/ 键下各子键的名字, 可以更改各种中文输入法的排列顺序。比如说将自己最喜欢的输入法对应的子键名命名为 2, 其他的子键也按序重命名。重新启动系统后, 只要按 CTRL+SPACE 键就可以在英文输入法和自己最喜欢的中文输入法间切换, 避免了要在诸多中文输入法中选出自己最喜欢的输入法。如果对其他与输入法有关的注册表键进行修改的话, 还可以获得其他功能。■

### 参考文献

- 1 曹国钧 王健 编著 Windows 95/98 中文版注册表、配置与应用实例 清华大学出版社
- 2 Richard C. Leinecker & Tom Archer 著 张艳 王文学 张谦 尹岩青等译 Visual C++ 6 宝典 电子工业出版社, 99 年 4 月第一版
- 3 朱友芹 主编 张争平审校 新编 Windows API 参考大全 电子工业出版社 2000 年 3 月 第一版