小企鹅输入法(FCITX)3.4.1(GirlDog)

Yuking

一、小企鹅中文输入法(FCITX)是什么?

小企鹅中文输入法 (Free Chinese Input Toy for X, FCITX) 是一个以 GPL 方式发布的 XIM 简体中文输入法,编写它的目的就是为了能够提供一个轻量级、基本不依赖其它库的简体中文输入方案。

本程序使用了IMDKIT 输入接口,可为支持 XIM 的 X 应用程序提供简体中文输入服务 FCITX 提供了区位和全/简/双拼,并支持基于码表的输入法(程序自带了五笔、二笔和仓颉)。您可以从 www.fcitx.org 下载最新版。

二、FCITX的安装

FCITX 通常以源码的方式发布,您需要编译安装后才能使用。编译 FCITX 需要以下包:

- GCC-2.95.3 或更新
- XFree86-4.3.0/xorg 或更新(及其开发包)

首先下载 FCITX 的源码包,文件名为 fcitx-3.4.1.tar.bz2,存放在某个临时的目录中(如/temp)。执行:

cd /temp

tar jxvf fcitx-3.4.1.tar.bz2

这样会建立目录 fcitx-3.4.1, 所有的源码都在该目录中。 执行以下命令编译安装:

cd fcitx-3.4.1

./configure -prefix=<安装目录> make && make install

如果您的系统安装 XFT 并配置正确,上述命令将编译安装 XFT 版。如果您不想使用 XFT,请执行以下命令:

cd fcitx-3.4.1

./configure -prefix=<安装目录> --disable-xft make && make install

如果您的系统没有安装 XFT,安装程序会自动关闭 XFT 支持。

应用程序通过环境变量 XMODIFIERS 与 XIM 沟通,设置 XMODIFIERS 的方法如下(以 bash 为例):

export XMODIFIERS="@im=YOUR XIM NAME"

上面的 YOUR_XIM_NAME 为 XIM 程序在系统注册的名字。应用程序启动时会根据该变量查找相应的 XIM 服务器。因此,即便系统中同时运行了若干个 XIM 程序,一个应用程序在某个时刻也只能使用一个 XIM 输入法。有时候您会遇到这种情况,明明 XIM 已经启动了,但应用程序就是无法使用。这往往是因为 XMODIFIERS 设置不正确所致,如您配置系统时在不同的脚本文件中修改了 XMODIFIERS,从而导致 X 启动时与启动完成后该变量不一致。fcitx 缺省注册的 XIM 名为 fcitx,但如果 fcitx 启动时 XMODIFIERS 已经设置好,fcitx 会自动以系统的设置来注册合适的名字。

另外, XIM 还依赖于系统 locale 的设置,如果 XIM 启动的 locale 与应用程序启动时的 locale 不一致,也可能无法使用。需要注意的是,locale 设置为 en US 将无法输

入中文。如果您想使用英文环境,请使用 en US.UTF8。

对于新安装的 Mandrake 和 RedHat7/8/9,最简单的方法是以 root 身份执行以下命令(请注意备份 chinput):

cd /usr/bin
ln -s fcitx chinput

如果您使用的是其它系统,则需要手工设置环境变量。一般可以在~/.bashrc增加行:

export XMODIFIERS="@im=fcitx"

重新登录当前用户,启动 X,然后在 X 下手工运行 fcitx(在某些环境中,这样启动 FCITX,可能只能为之后启动的程序输入中文)。您也可以将 fcitx 加到窗口管理器的自动启动中,具体方法请参考相应窗口管理器的说明。

四、FCITX的配置文件

fcitx的配置文件为~/.fcitx/config(首次运行时fcitx会自动创建该目录),该文件由中文写成(编码为gb2312),从中可以了解几乎所有的功能。下面是配置文件的内容:

[程序]

显示字体(中)=* 显示字体(英)=Courier New 显示字体大小=12 主窗口字体大小=11 字体区域=zh_CN.UTF-8 是否使用 AA 字体=1

[输出]

数字后跟半角符号=1 Enter 键行为=2 分号键行为=2 大写字母输入英文=1 转换英文中的标点=1 联想方式禁止翻页=1

「界面1

候选词个数=5

主窗口是否使用 3D 界面=0

输入条使用 3D 界面=2

主窗口隐藏模式=0

显示虚拟键盘=0

是否自动隐藏输入条=0

输入条是否居中=1

首次显示输入条=1

#输入条固定宽度(仅适用于码表输入法),0表示不固定宽度

输入条固定宽度=0

输入条偏移量 X=0

输入条偏移量 Y=12

序号后加点=1

显示打字速度=1

光标色=92 210 131

主窗口背景色=240 255 240

主窗口线条色=150 220 150

```
主窗口输入法名称色=170 170 170 150 200 150 0 0 255
输入窗背景色=240 255 240
输入窗提示色=255 0 0
输入窗用户输入色=0 0 255
输入窗序号色=200 0 0
输入窗第一个候选字色=0 150 100
#该颜色值只用于拼音中的用户自造词
输入窗用户词组色=0 0 255
输入窗提示编码色=100 100 255
#五笔、拼音的单字/系统词组均使用该颜色
输入窗其它文本色=0 0 0
输入窗线条色=90 160 90
输入窗箭头色=255 0 0
虚拟键盘窗背景色=220 220 220
虚拟键盘窗字母色=80 0 0
虚拟键盘窗符号色=0 0 0
#除了"中英文快速切换键"外,其它的热键均可设置为两个,中间用空格分隔
[热键]
打开/关闭输入法=CTRL SPACE
#中英文快速切换键 可以设置为 L CTRL R CTRL L SHIFT R SHIFT
中英文快速切换键=L CTRL
双击中英文切换=0
击键时间间隔=250
光标跟随=CTRL K
GBK 支持=CTRL M
GBK 繁体切换键=CTRL ALT F
联想支持=CTRL L
反查拼音=CTRL ALT E
全半角=SHIFT SPACE
中文标点=ALT SPACE
上一页=- ,
下一页==
第二三候选词选择键=SHIFT
[输入法]
使用拼音=1
使用双拼=1
默认双拼方案=自然码
使用区位=1
使用码表=1
提示词库中的词组=1
[拼音]
使用全拼=0
拼音自动组词=1
保存自动组词=0
增加拼音常用字=CTRL 8
删除拼音常用字=CTRL 7
删除拼音用户词组=CTRL DELETE
#拼音以词定字键,等号后面紧接键,不要有空格
拼音以词定字键=[]
```

#重码调整方式说明: 0-->不调整 1-->快速调整 2-->按频率调整 拼音单字重码调整方式=2 拼音词组重码调整方式=0 是否模糊 an 和 ang=0 是否模糊 en 和 eng=0 是否模糊 in 和 ing=0 是否模糊 ou 和 u=0 是否模糊 uan 和 uang=0 是否模糊 c 和 ch=0 是否模糊 f 和 h=0 是否模糊 s 和 sh=0 是否模糊 z 和 zh=0

下面对其中部分设置加以说明:

- **字体区域**: 当系统 locale 设置为非 zh_CN*时,可能会导致中文显示为乱码,指 定该值可解决这个问题(如 zh_CN.GBK)。
- **联想方式禁止翻页**:由于默认的,/.为翻页键,因此联想方式下会导致标点符号输入不方便,打开该选项会使得在联想方式下不翻页。
- **候选词个数**:选择范围为 1~10,如果大于 10,则为 10。
- **输入条使用 3D 界面**: 0-->不使用 3D 界面; 1-->输入条使用凸出效果; 2-->输入 条使用凹进效果。
- 是否自动隐藏输入条:设置为1时,只有当输入条有提示信息前才会显示。
- **主窗口隐藏模式**: 0-->一直显示在顶层; 1-->只有在中文输入状态下才显示; 2-->不显示。
- Enter 键行为: 当有用户输入时, 0-->不处理; 1-->清除用户输入; 2-->用户输入的字母上屏。
- 分号键行为: 0-->只作为标点处理(如果标点文件中包含了分号,否则就要半角字符处理); 1-->进入英文输入状态; 2-->进入快速输入状态(稍后有详细说明)。
- 双击中英文切换:有些程序应用程序不支持按键的释放事件,此时快速中英文键会 失效,打开该选项可以利用双击来进行中英文切换。
- 反查拼音:反查汉字读音,只对码表输入法有效。
- 使用拼音:是否使用全拼。
- 使用双拼:是否使用双拼。
- 使用区位:是否使用区位输入。
- **使用码表**:是否使用码表输入法(码表输入法还依赖于 table.conf 的设置,请参看"码表输入法"部分)。
- 提示词库中的词组:该选项针对码表输入法,打开的话会提示您系统已有的词组,然后根据屏幕提示操作即可。
- 默认双拼方案: 设置默认拼音方案,但如果用户目录的.fcitx中有 sp.dat,并 在其中设置了默认的双拼方案,则该设置无效。
- · 使用全拼:打开该选项的话,拼音将以全拼的方式工作,此时不进行简拼的匹配。
- 拼音自动组词:该功能指定是否进行拼音的短语输入。
- 保存自动组词:是否保存自动组成的拼音短句。

颜色是按照 'R G B'的方式设置,您可以参照作图软件(如GIMP)的颜色表来设置。

五、输入法

1、拼音输入

- 支持全/简/双拼,双拼可自定义键盘方案
- 常用字表
- 特殊符号输入
- 完善的预编辑
- 支持以词定字
- 模糊拼音

* 常用字表:

设置常用字表的目的是为了方便录入某些常用字。

在候选字中,处于该表中的字总是排在其它字的前面。刚刚使用 fcitx 时,常用字表是空的。您可以用左 ctrl+8/7 来添加/删除常用字。例如,您希望录入"d"时,"的"总放在第一个,可以按以下操作:

- 进入拼音状态
- 输入"d",用翻页键-=/,.查找"的"字,直到它显示在候选字表中按 ctrl+8, 然后按"的"前面的数字序号即可
- 按上述操作可以为某个拼音编码设置多个常用字。您可以设置常用字表内的字频调节方式。

* 双拼输入:

如果在配置文件打开双拼选项,即可进行双拼输入。您可以定义双拼键盘方案,如果您没有定义,程序将采用自然码的双拼方案。如果您需要自定义,请将存放双拼方案的文件放在~/.fcitx下,名为sp.dat。程序源码包中的data目录下有一个名为sp.dat的文件,该文件已经定义了一些双拼方案,修改"默认方案"即可以使用用户指定的方案。注意,模糊拼音的设置也会在双拼中生效。

* 以词定字

fcitx 支持"以词定字",以方便录入某些排序比较靠后的字。利用 SHIFT_<词组序号>,按以词定字键(默认为'['/']')选择该词组的第一个或最后一个字。

比如,在默认设置下,如果需要录入"亥"字,可以输入"jh",按翻页键让词组"己亥"显示在输入条上,然后按SHIFT_<它前面的序号>,再按']'即可(如果按'['则是输出"己"字)。

* 特殊符号输入:

为了录入特殊符号, 您必须将特殊符号按下列格式放在一个文件中:

编码 符号

每个符号占一行。编码部分必须是英文小写字母,且经拼音解析后的长度不应超过 10 个汉字(如 zzz 是 3 个汉字长度,而 zhangzhangzhang 也是 3 个汉字长度), fcitx的源码包中有一个例子(文件 pySym.mb)。

编辑号特殊符号后,将其保存为文件 pySym.mb,并放在~/.fcitx中,重新启动fcitx即可。

- *建议: 特殊符号的编码应该足够长, 免得影响正常的汉字录入。
- *注意:如果定义的编码与常用字表的编码相同,该特殊符号将无法使用。
- *如果要录入"西安"这样的词组,请用单引号作为分隔符(xi'an)。

*制作拼音库

在 tools 目录下提供了一个制作拼音库的工具 create PYMB。用法如下:

./createPYMB <拼音单字库> <拼音词组库>

该工具将在当前目录下生成 pybase.mb 和 pyphrase.mb,将它复制到您的 fcitx 安装目录下的 share/fcitx/data 中替换原有文件即可。

在 data 目录下包括了单字库和词组库,文件名为 gbkpy.org 和 pyPhrase.org。它们都是文本文件。一般说来,单字库无需更改,您可以将自己的词组加入到

pyPhrase.org 中,从而生成自己的词库。如果您使用的不是 IA32 架构的计算机,请您用该工具生成拼音码表,并替换安装目录中的相应文件,否则将无法工作。

2、区位输入

fcitx支持0~9的区位输入。

3、码表输入法

fcitx 支持用户自定义码表输入法。只要用户在安装目录的 tables.conf 或 ~/.fcitx/tables.conf 中提供相应的码表信息,就可以使用自己的码表输入法。下面是该文件的一个例子:

[码表] 名称=五笔 码表=wbx.mb 调频=0 拼音=1 拼音键=z 自动上屏=1 模糊=1 模糊键=z 自动词组=1 自动词组长度=4 词组参与自动造词=1 保存自动词组=0 精确匹配=0 提示编码=0 符号=zzzz [码表] 名称=二笔 码表=erbi.mb 调频=0 拼音=1 拼音键=「 自动上屏=1 模糊=1 模糊键=? 自动词组=1 自动词组长度=4 词组参与自动造词=0 保存自动词组=0 精确匹配=0 提示编码=0 [码表] 名称=仓颉 码表=cj.mb 调频=0 拼音=1 拼音键=[自动上屏=1 模糊=1 模糊键=?

自动词组=0 精确匹配=0 提示编码=0

该文件用"[码表]"来区分不同的码表。其它设置说明如下:

- 名称:显示在程序主窗口的输入法名称。
- **码表**:该码表输入法的码表文件名,程序会先查找~/.fcitx,然后在安装目录中找。
- 调频:与拼音中的相应设置意义一样。
- 拼音:是否使用临时拼音输入。
- 拼音键:如果"拼音"选项打开,则以该字母起头的输入按全拼处理。
- 自动上屏: 当输入达到最大码表且只有一个候选词时是否自动上屏。
- 模糊:是否使用模糊(通配符)输入。
- 模糊键:模糊键(通配符)。
- 自动词组:是否使用自动组词功能(后面有详细说明)。
- 自动词组长度:自动组词长度。
- 词组参与自动造词: 指定录入的词组是否参与自动组词。
- **保存自动词组**:设定自动生成的词组被选择多少次后才被保存。0表示不保存。
- 精确匹配:是否只在候选字表中显示精确匹配的结果。
- **提示编码**:是否提示录入字/词的编码。
- 符号:设置该选项则开启特殊符号输入功能。
- 符号文件: 特殊符号所在的文件。

*注意: fcitx以"[码表]"开始一个新的码表输入法,因此,即便只配置一种码表输入法,也应该有该行。

码表输入法提供了两种在线造词方法(词组最长为10个汉字):

- 1) 在中文输入方式下按 CTRL_8,则利用将刚刚输入的内容造词,默认为最近输入法两个字,可以用左右方向键的增加或减少词组中的字数。
- 2) 自动组词:将需要造的词按单字连续输入后,再按它的组词规则连续输入编码,程序会提示用户这个新词。如果此时按空格或它前面的序号则将这个新词输入到用户程序中,您可以设置这个新词是否进入词库。如果不想录入该词,继续进行下一次输入即可(fcitx只能记录最近2048个输入的字)。

如果想删除词库中的词,先让该词显示中输入条上,按 CTRL_7,并按提示操作即可;或是当程序提示有该词组时,按 CTRL DEL 删除。

如果想调整词库中词的顺序,按 CTRL_6,并按提示操作即可。 (如果调频设置为 1 或 2, 您可能看不到有什么变化)。

*制作码表

data 目录下包括制造码表的工具: txt2mb 和 mb2txt。前者是将码表源文件转换为码表输入法所需的格式;后者是将码表文件转换为文本文件。码表源文件格式如下:

```
键码=abcdefghijklmnopqrstuvwxyz;,./
码长=4
规避字符=;iuv
[组词规则]
e2=p11+p12+p21+p22
e3=p11+p21+p31+p32
a4=p11+p21+p31+n11
[数据]
a 工
```

aa 式 aaa 工 aaaa 工

说明如下:

- 键码:表示该码表输入法需要用的键;
- 码长: 指该码表输入法最长码长:
- **规避字符**:在象二笔这样的码表输入法中,某些字符如果出现在编码的第一个,表示特殊用途,虽然可以组成单字,但却不参与组词。如果有这样的字符,可以列在这个地方,以免自动造词出错。
- [组词规则]:表示自动组词的规则,如果没该标志,则不能在线造词;
- **[数据]**: 该标志以下为该码表输入法的数据,按"编码<空格>对应的汉字"。这部分无需排序。

组词规则:

- 组词规则的数目应该比码长小1。
- 等号左边为词组的汉字个数。其中 e 表示"等于", a 表示"大于等于"(只能有这两个设置)。
- 等号右边为规则,其中 p 指正序, n 指逆序。如 p11 指第一个字的第一个编码, n11 则是指最后一个字的第一个编码。

按上述格式制作好编码后,利用 txt2mb 将其转换为 fcitx 需要的格式:

./txt2mb <源文件> <目标文件>

然后将目标文件复制到~/.fcitx中,并在~./fcitx/tables.conf中设置好,然后在中文输入状态下按CTRL 5重新读入配置文件即可。

如果希望将已经制作好的码表文件转换为文本文件,可以:

./mb2txt <码表文件> > <文本文件>

生成的文本文件将形如上述码表源文件的格式。

在码表输入法模式下,只输入单个"拼音键"指定的字符可以重复上次的录入。

*反查拼音

如果不知道某个汉字的读音,可以先用码表输入法录入这个字,然后按反查拼音的热键 (默认为 CTRL ALT E),就可以查到该字的读音。

六、中文标点

fcitx的中文标点由一个文本文件设置,您可以自己的需要修改。该文件即<FCITX的安装目录>/share/fcitx/punc.mb(或用户目录下的 fcitx/punc.mb)。格式如下:

对应的英文符号 中文标点 <中文标点>

其中英文符号应该是类似 "&*()"这样的符号,而中文标点最长为两个汉字,最多有两组,中间由空格隔开。如:

-,

七、快速输入

当配置文件中的,"分号键行为"设置为2时,按下分号键则进入快速输入模式。在这种模式下,您可以设置某些常用短语或符号的快速录入规则。

为了使用该功能,您需要将常用短语和符号按如下格式编辑

<字符组合> <短语>

并保存在~/.fcitx/QuickPhrase.mb(或fcitx的安装目录下的share/data)中,一个短语一行。如

zg 中华人民共和国 h http

八、热键

您可在 FCITX 的配置文件中使用以下热键:

LCTRL LSHIFT RCTRL LALT RSHIFT RALT INSERT HOME PGUP END PGDN CTRL CTRL CTRL_RSHIFT CTRL_RALT
SHIFT_LALT SHIFT_RCTRL CTRL_LALT CTRL LSHIFT SHIFT LCTRL SHIFT SHIFT SHIFT RALT ALT LCTRL ALT LSHIFT ALT ALT ALT RCTRL ALT RSHIFT

其中,L前缀表示左键,R前缀表示右键。如果需要CTRL/ALT+<字母/数字>的组合键,直接写入数字或字母即可,如CTRL K等。

九、其它功能

- 可以任意拖动 fcitx 显示的图标和输入条。
- 可以用鼠标左键单击输入条的相应部分切换输入法/改变标点状态/改变全半角设置/改变 GBK 状态/改变联想状态。
- · 支持 UTF8 方式下的中文输入。
- 在主窗口上按鼠标右键可以切换主窗口的显示模式。有两种:简洁模式和完整模式。
- 按 CTRL_5 可以重新读取配置文件。向 fcitx 发送 SIGHUP 也可以重新读取配置文件。
- 在联想方式下,按第二选择键输出半角空格,按第三选择键输出全角空格。
- 单击企鹅图标可以改变输入法的状态。
- 支持输出繁体汉字

十、其它说明

fcitx(3.1x以前的版本)使用了来自于cjkvinput项目的五笔86码表,请参看源码中的doc/cjkvinput.txt。

fcitx-3.2 使用了拼音佳佳 4.0 的拼音词库,并根据该词库制作了五笔 86 的词库。详情请参看源码包中的 doc/pinyin.txt。

fcitx-3.3.X以后使用了kardinal@linuxsir制作的五笔86词库(稍做修改)。

十一、致谢

本程序参考了rfinput-2.x的源码,在此向其作者 Zhao yang(zyang@redflag-linux.com) 致谢。

感谢 <u>www.linuxsir.org</u>的北南南北为 fcitx 提供一个家。同时,他与小凡为gWuBi/fcitx的推广和测试做了很多工作。

感谢 xsim 的作者楚狂、scim 的 james su 提供了宝贵的技术支持;

感谢二笔输入法 (ebf) 的作者 tram, 他为 fcitx 做了很多工作。

感谢 winix 多次为 qWuBi/fcitx 提出宝贵的意见和建议。

感谢 windrose 精心测试了双拼功能。

感谢 Wang Yin 为拼音输入提出了宝贵的意见。

感谢 Xie Yanbo 制作了安装配置脚本。

同时感谢所有其它测试 gWuBi/fcitx 的朋友,非常感谢。

常见问题(FAQs)

1. 编译时说找不到文件 X11/Xlib.h

这个信息是说编译程序没有找到X的头文件,一般是因为您没有安装X的开发包(在RH/MDK中名为XF*-devel-*或XOFG-devel-*)。

2. 编译安装完成后,按 ctrl space 无法打开 fcitx 的输入条啊?

这个的原因有很多,比如fcitx没有启动,或是环境变量设置不正确。请参看说明文档的"环境配置"部分。如果还有问题,就给我写信或到linuxsir上去说吧。

3. 界面的颜色怎么那么难看?

就是,我也这么认为。但我不知道什么样的颜色搭配才好,不过,您可以根据说明文档自己设置颜色,更改字体的大小。

4. 我将fcitx加到gnome的自动启动后,gnome怎么启动那么慢?

把~/.gnome2/session-manual 里面 0, Priority=31的31换成50即可。

5. 安装好了fcitx,可是为什么没有五笔输入法啊?

呵呵, 您一定没有好好看说明文档了……

6. 在我的 FC3 上为什么不能用使用?

FC3 使用了 IIIMF 作为默认的输入接口,它与 XIM 不兼容。您可以利用它的配置工具将 IIIMF 停用,然后按配置 XIM 的方式配置 FCITX 即可。

7. 为什么我只按了一个键,输入条上却出现两个相同的字符?

这种情况一般出现在 KDE 中。应该是 QT 的 XIM 支持出现问题。出现这个现象时,请在窗口管理器的标题栏按鼠标右键弹出系统菜单,然后按 ctrl_space 关闭输入法,重新回到应用程序就可以了。

8. 当我使用 en US.UTF-8 时,为什么无法激活 fcitx?

这种情况应该是 GTK2 的应用程序。当 locale 为 en_US.UTF-8 时。GTK2 默认的输入模块不是 XIM,因此无法激活 fcitx。此时在文字输入框中按鼠标右键,把 Input Method 设置为 X input method,然后就可以使用 fcitx 了。比较好的解决方法是,在 ~/.bash profile 中加入以下两行:

export GTK_IM_MODULE=xim

export XMODIFIERS="@im=fcitx"

9. 为什么我修改了配置文件却无法生效?

FCITX的配置文件的编码是GB2312,如果保存为UTF8,FCITX将会采用默认配置。

10. 我发现了 fcitx 的 BUG 怎么办?

您可以将该 BUG 的情况贴在 www.linuxsir.org 的输入法版块或是直接写信告诉我 (yuking net@sohu.com), 我将尽快修正。

11. 我过去说过很多建议或意见,为什么新的版本还是老的样子?

这个……不好意思,肯定是我忘记了,以后我一定会将意见归纳好的。请您给我写信再说说吧,yuking_net@sohu.com等着您呢!