

四角号码检字法

chaishushan@gmail.com

四角号码有用吗？

四角号码有用吗？

答：非常有用！

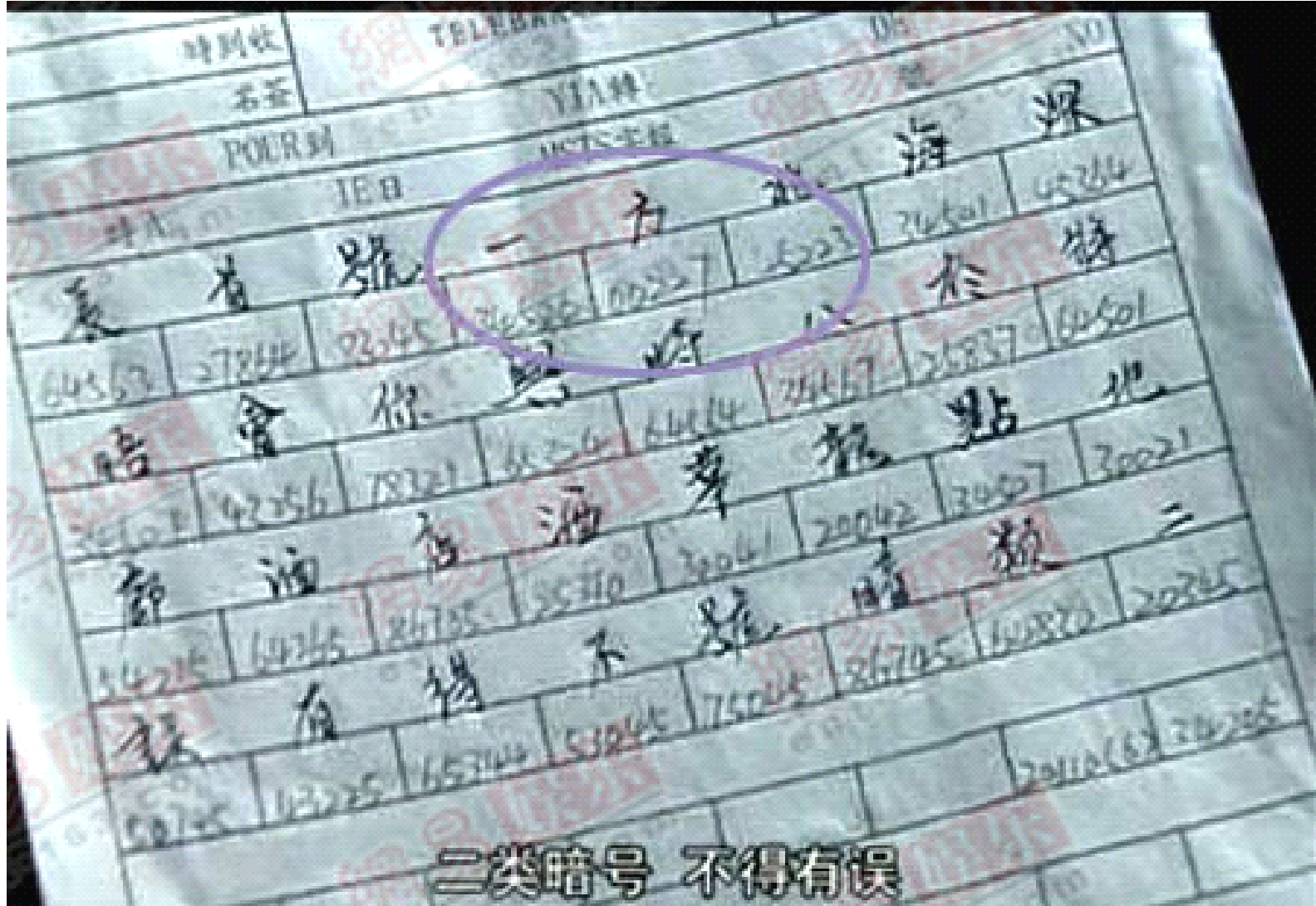
余则成在接收情报...



陆桥山截获情报

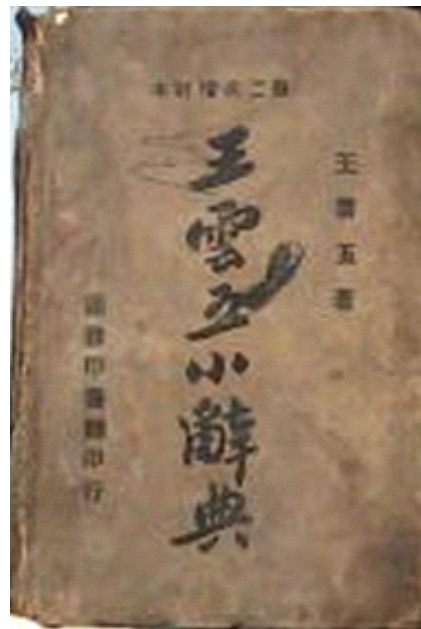
- 陆桥山办公室内...
- 一个特务拿着电文进来：处长，刚截获的。
- 陆桥山接过来看：谁的？
- 特务：不清楚，但是落款跟以前呼叫秋掌柜的电文一样。特务又拿出过去的一张电文。
- 陆桥山看着，新电文的最后一组数字是**20110**，老电文的第一个数字也是**20110**。
- 陆桥山一脸不解：姓秋的又回来了？叫汤四毛。
- 汤四毛老老实实地看着电文：这是以前规定的不定期联络密码，从来没有用过。这里，**他用手指着电文最后20110（6）后面括号里的6：这指的是译码手册。**
- 陆桥山：你能记得是哪个手册吗？
- 汤四毛回忆：教练手册上的教练码，6指的是...，**给我找一本民国20年的《四角号码》。**
- 汤四毛对着一个破旧的《四角号码》翻译着，陆桥山等在边上。
- 汤四毛翻译完了：**“深海”，“北方一号”首长将于明晚八时与你会晤，地点，龙华酒店酒廊，二类暗号，不得有误。**
- 陆桥山如获至宝，**拿过电文：深海...，北方一号首长...**

李涯伪造的电文



检字法的三大发明

- 一、东汉许慎的部首检字法。
- 二、明朝/清朝部首加笔画检字法。
- 三、1925年王云五的四角号码检字法。



汉语拼音怎么来的？

- 古代汉字多采用“**两字相切**”的读音方法，如“**东**”(**都笼切**)
- 1626年，**法国传教士金尼格莱**写的《**西中儒耳目资**》首次用**拉丁拼音字母**记录汉字的读音
- 中国**大陆**于**1955年—1957年**文字改革研究制定汉语拼音方案(**汉字拉丁化**)

部首检字法(有难又慢)

47 漢字部首查檢法

先	熟悉部首	諸書部首各異	《說文》540部，《康熙字典》214部，現代字典多另行調整。	
		部首筆畫有序	各部首按筆畫為序排列，如《康熙字典》從一部到龠部共十七畫。	
		瞭解部首變形	如口作巳、心作忄、火作灬、走作辵，見本書《漢字部首異形表》。	
		確定所屬部首	參考諸書檢字說明及本表，決定每字所屬的準確部首。	
		翻查各部頁碼	部首確定後，按字典的目錄，翻到該部首所在之頁碼。	
定	部	瞭解各字的部首位置(形聲字多為形旁)	上	今(人)字(宀)尖(小)花(艸)笙(竹)霖(雨)等，及「上形下聲」一類字。
			下	啓(口)基(土)婆(女)忠(心)盲(目)背(肉)等，及「下形上聲」一類字。
			左	惜(心)湖(水)矜(矛)蝗(虫)銅(金)軀(身)等，及「左形右聲」一類字。
			右	我(戈)期(月)歌(欠)皺(皮)視(見)頂(頁)等，及「右形左聲」一類字。
			外	亞(二)園(口)班(玉)衛(行)閣(門)闕(門)等，及「外形內聲」一類字。
			內	丹(冫)問(口)嬖(女)愛(心)辯(彳)讎(言)等，及「內形外聲」一類字。
			中坐	事(亅)兆(儿)坐(土)奭(大)東(木)裏(衣)等。
			左上	嗣(口)顛(竹)皝(白)聖(耳)荊(艸)虧(虍)等。
			右上	匏(匚)匙(匕)弑(弋)整(攴)望(月)號(虍)等。
			左下	彀(弓)歸(止)栽(木)毓(母)穎(禾)糴(米)等。
右下	乾(乙)務(力)叢(又)幹(干)懿(心)疑(疋)等。			
再數筆畫	通常字典各部首下列的字也按筆畫(不包括部首本身)排序，在選好部首、翻到該部後，計算該字部首以外的筆畫，即可查到該字。			
難字查表	① 有些字難以確定部首，要憑經驗或牢記。			
	② 查檢難字，可參考本書的《難檢部首字例》表。			
	③ 通常字典都有難字表以供查閱，其排列以全字筆畫多寡為序。			

四角检字法(这么简单?)

横一垂二三点捺

叉四插五方框六

七角八八九是小

点下有横变零头

四角号码为何不流行？

- 曾经非常流行(那是民国时期), 但是...
 - 现在老师不教(可能也不会)
 - 现在教育界不推广(课本中就是不说)
 - 现在出版界不支持(四角字典少, 且价格贵)
 - 或许这是国家机密...
-
- <康熙字典><辞海>有四角号码索引
 - 商务印书馆<新华字典>有四角号码索引, 价格贵
 - <古汉语常用字字典>没有四角号码索引
 - 新版<现代汉语词典>取消了四角号码索引
 - 居然有学者建议<辞海>取消四角索引

四角号码简介

- **四角号码**：把笔划分为十类，用最多五个数字对汉字排序
- 王云五发明**四角号码**检字法，高梦旦发明“**附角**”
- 1925年05月《**号码检字法**》由商务印书馆出版
- 1926年09月《**四角号码检字法**》出版，**蔡元培/胡适**等为序
- **1933年**12月《**四角号码检字法 附检字表**》出版（收录在万有文库第一集），加上高梦旦序

- **胡适**曾作歌诀赠**王云五**，名为《**笔画号码歌**》
- **1949年后**，**王云五**受到**批判(第十五号战犯)**，四角号码**被边缘化**
- **1949年后**，采用**黄维荣**修改版本的口诀

- 注：*商务印书馆成立于1897年*

笔画号码歌

胡适原作

一横二垂三点捺，
点下带横变零头；
叉四插五方块六，
七角八八小是九。

黄维荣改

横一垂二三点捺，
叉四插五方框六；
七角八八九是小，
点下有横变零头。

使用范例

四角号码便于人脑理解



QR码便于电脑理解



笔画规则 - 单笔

复笔	头	0	丶	主病广言
单 笔	横	1	一	天士
			㇇ ㇆ ㇅ ㇄	活培织兄风
	垂	2	丨	旧山
			㇇ ㇆ ㇅ ㇄	千顺力则
	点	3	丶	宝社罕外去亦
			㇇ ㇆	造瓜

笔画规则 - 复笔

复 笔	叉	4	十	古草
			𠄎 𠄎 𠄎 𠄎	对式皮猪
	插	5	丰	青本
			扌 戈 丈 彡 丰	打戈史泰申
	方	6	口 口	另扣国甲由曲
			冂 冂	目四
	角	7	丿 丿 ㇇ ㇇	刀写亡表
			冫 冫 冫	阳兵雪
	八	8	八 八	分共
			人 入 人 ㇇ ㇇	余余央羊午
	小	9	小	尖宗
			小 小 ㇇ ㇇ ㇇	快木录当兴组

附加规则

- 取笔时应以复笔优先
- “口/门/鬥”组成的字下方取内部笔画
- “口/门/鬥”上下左右有其他笔时除外
- 每笔用过后再充他角作0

范例讲解

四角號碼
6021₀ 2722₇ 6121₇ 1162₇

檢字法
4898₆ 3040₇ 3413₁

生字练习

毛

脑

冀

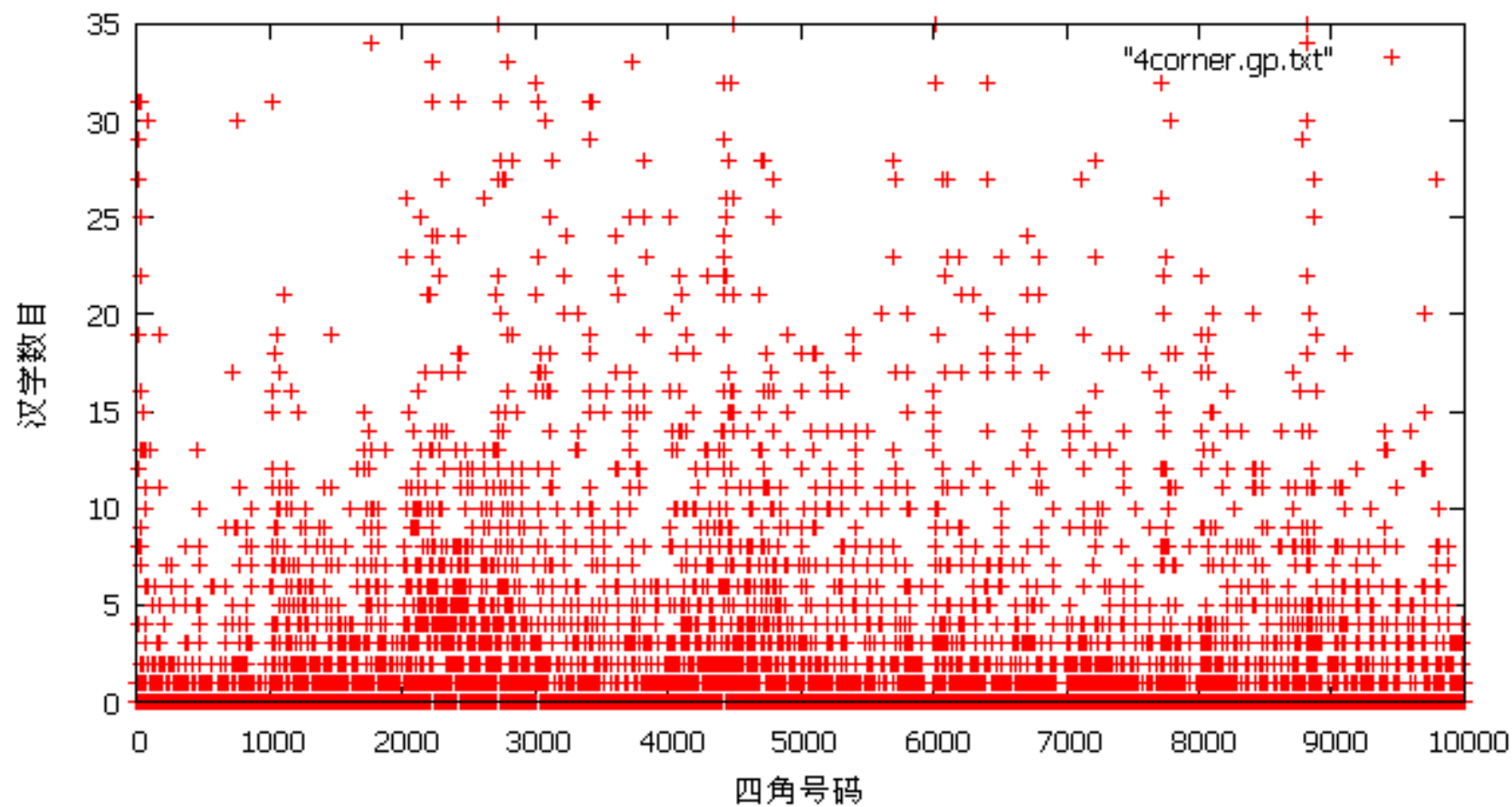
撞

性能分析

- <通用规范汉字表>收录汉字**8300**个
- 常用字典/词典收录汉字在**11000+**个左右
- 拼音声母**23**个, 韵母**24**个, 组合约**400**个
- 主部首**201**个, 附部首**99**个, 共**300**个
- 常用汉字笔画最多约**40**个
- 四角号码有**10000**个
- 对8300个标准汉字统计分析图(gnuplot)

四角号码分布图

四角号码分布图 - 20902个汉字



检字法本质

- **1键对N字**, 是一种**HASH**函数
- `std::multimap<key, values>` dictionary;
- HASH函数要求: **计算方便**; **分布均匀**

- 各个算法的**重码率**(**假设均匀分布**)
- 拼音: $11000/400 = 27.5$
- 部首: $11000/300 = 36.7$
- 笔画: $11000/40 = 275$
- 四角: $11000/10000 = 1.1$

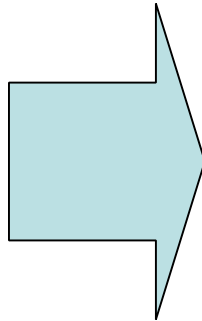
检字法对比

- 拼音: 重码较多(慢), 不认识/读不准 的字
 - 部首: 重码较多(慢)
 - 笔画: 重码很多(慢)
 - 四角: 重码很少(快), 不知道字形
-
- 汉字的结构比较稳定
 - 汉字的读音不够稳定(方言/ch/sh/zh/l/n/...)

更多的生僻字(输入/检索)

- 卅 𠂇 𠂉 𠂊 𠂋 𠂌 𠂍 𠂎 𠂏 𠂐 𠂑 𠂒 𠂓 𠂔 𠂕
- 𠂖 𠂗 𠂘 𠂙 𠂚 𠂛 𠂜 𠂝 𠂞 𠂟 𠂠 𠂡 𠂢
- 𠂣 𠂤 𠂥 𠂦 𠂧 𠂨 𠂩 𠂪 𠂫 𠂬 𠂭 𠂮 𠂯 𠂰
- 𠂱 𠂲 𠂳 𠂴 𠂵 𠂶 𠂷 𠂸 𠂹 𠂺 𠂻 𠂼 𠂽
- 𠂾 𠂿 ...

- 怎么输入电脑?
- 怎么发音?
- 什么含义?



本质都是查找汉字,
重点是: 音/形转换!

电脑输入法

- 通用字输入:
- 一般根据拼音输入
- 采用词组和拼音算法可极大减少重码

- 生僻字输入:
- 四角/五笔/笔画/其他

Google拼音输入法(推荐)

- 动态词序/自定义词组/帐号同步
 - i模式: 扩展命令(Lua脚本)
 - u模式: 笔画输入/部件输入
 - v模式: 英文输入
-
- 搜狗拼音对比:
 - 优点: 专业字库
 - 缺点: 扩展不方便/缺少笔画输入

Google输入法扩展(Lua)

```
ime.register_command(  
    command_name, lua_function_name,  
    description [, leading [, help]]  
)  
ime.register_trigger(  
    lua_function_name, description,  
    input_trigger_strings, candidate_trigger_strings  
)  
ime.register_converter(  
    lua_function_name, description  
)
```

注册扩展命令(四角)

-- 查询四角号码表

```
function SearchTableBy4C(args)
  if isArgsValid(args) then
    return { "第一个", "第二个", "第三个", }
  else
    return "-- 有效格式: 108;1080;1080?;... --"
  end
end
```

-- 注册命令

```
ime.register_command(
  "sj", "SearchTableBy4C", "四角号码", "alpha",
  "四角号码输入法, 例如isj1080[?] -> 厶"
)
```

查表函数实现(四角)

-- 查询四角号码表

```
function SearchTableBy4C(id)
    return im4c_table[tonumber(id)]
end
```

-- 初始化四角号码表

```
local im4c_table = {}
im4c_table[0000]={"一"}
im4c_table[0002]={"弯"}
im4c_table[0003]={"变"}
im4c_table[0010]={"主","亠","亠","塹","壅","壘",...}
...
im4c_table[9999]={"糝"}
```

Google输入法扩展调试

```
C:\api>GooglePinyinApiConsole.exe base.lua
Google Pinyin IME Extension API Console V2.2.13.76
Type "help" for more information.
> ?
i          - lists all commands
i [COMMAND] - evaluates command without argument
i [COMMAND] [ARGUMENT] - evaluates command with argument
g [TRIGGER_STRING] - tests a trigger string, fires trigger if hit
c          - lists all converters
c [FUNCTION] [STRING] - tests a converter function
quit      - quits the shell
help      - shows this message
> i rq 2009-3-3
a. 2009-03-03 b. 2009年3月3日 c. 二〇〇九年三月三日
> i
gz.打印稿纸 > hh.画画 > js.计算模式 > rq.输入日期 > sj.输入时间 > sz.掷骰子
> xz.查询星座 > zf.打印字符 >
> _
```

四角号码扩展 - 使用

isj108| 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 天 b. 𠄎 c. 页 d. 灭 e. 贡

isj1080?| 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 天 [1080, tian1] » b. 𠄎 [1080, tian1] » c. 页 [1080, ye4] » d. 灭 [1080, mie4] » e. 贡 [1080, gong4] »

isj108??| 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 天 [1080, tian1] b. 𠄎 [1080, tian1] c. 页 [1080, ye4] d. 灭 [1080, mie4] e. 贡 [1080, gong4]

isj1?80| 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 天 b. 𠄎 c. 页 d. 灭 e. 贡

四角号码扩展 - 帮助

isj? 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 王云五 发明四角号码, b. 高梦旦 发明“附角”。 c. 1926年《四角号码检字法》出版, d. 蔡元培、胡适等为序。

isj?? 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 胡适创作歌诀: b. 一横二垂三点捺, c. 点下带横变零头, d. 叉四插五方块六, e. 七角八八小是九。

isj??? 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 黄维荣修改版: b. 横一垂二三点捺, c. 叉四插五方框六; d. 七角八八九是小, e. 点下有横变零头。

isj???? 四角号码输入法, 例如isj1080[?] -> 天 Google
a. 扩展作者: b. ChaiShushan c. <chai-shushan@gmail.com>

isj????? 四角号码输入法, 例如isj1080[?] -> 天 Google
a. <http://chai2010.bitbucket.org/im4corner.html>

Vim - 编辑操作系统

- **Vim是个编辑操作系统!**
 - 有编辑接口/命令接口/编程接口
 - 针对不同格式/对象可处理各种事件
- Vim是编辑**效率最高的编辑器!**
 - 不仅仅是列模式, Vim可定制任意模式
 - 支持**多标签页/拆分窗口/缓冲区**
- 基于Vim可以做什么?
 - 可以高效编辑文章, 特别是写代码(自动补全)
 - 可以自定义代码的缩进/高亮/模板等
 - 可以**开发游戏/绘图板**
 - 可以**开发输入法**

Vim - IDE

The screenshot shows the Vim IDE interface. At the top, there is a menu bar with options: 文件(F), 编辑(E), 工具(T), 语法(S), 缓冲区(B), 窗口(W), C/C++, 帮助(H). Below the menu bar is a toolbar with various icons for file operations and editing. The main window is divided into three panes:

- File Explorer (Left):** Shows a directory tree for a project. The root is `[1:main.c]*[7:term.c]`. The tree includes: `-MiniBufExplorer-`, `4 "= /home/wooin/vim71/src/`, `5 ../`, `6 auto/`, `7 objects/`, `8 proto/`, `9 testdir/`, `10 xxd/`, `11 INSTALL`, `12 Makefile`, `13 README.txt`, and `14 arabic.c`.
- Code Editor (Center):** Displays C code. The current file is `main.c` at `Line:36/72, Column:5` (50% zoom). The code includes: `969 # endif`, `970 . got_int = TRUE;`, `971 . need_wait_return = FALSE;`, `972 . global_busy = FALSE;`, `973 . exmode_active = 0;`, `974 . skip_redraw = FALSE;`, `975 . RedrawingDisabled = 0;`, `976 . no_wait_return = 0;`, `977 # ifdef FEAT_EVAL`, `978 . msg_skip = 0;`, `979 # endif`, `980 . msg_off = 0;`, `981 # ifdef FEAT_MOUSE`, `982 . setmouse();` (highlighted with a red box), `983 # endif`, `984 . settmode(TMODE_RAW);`, `985 . starttermcap();`, `986 . scroll_start();`, `987 . redraw_later_clear();`, `988 }`, `989 #endif`, `990`, `991 clear_oparg(&oa);`, `992 while (!cmdwin`, `993 #ifdef FEAT_CMDWIN`, `994 . || cmdwin_result == 0`, `995 #endif`.
- Terminal (Bottom):** Shows the current file path: `[-/vim71/src/main.c][c]` at `Line:982/3841, Column:2` (25% zoom).

Vim - 输入法接口

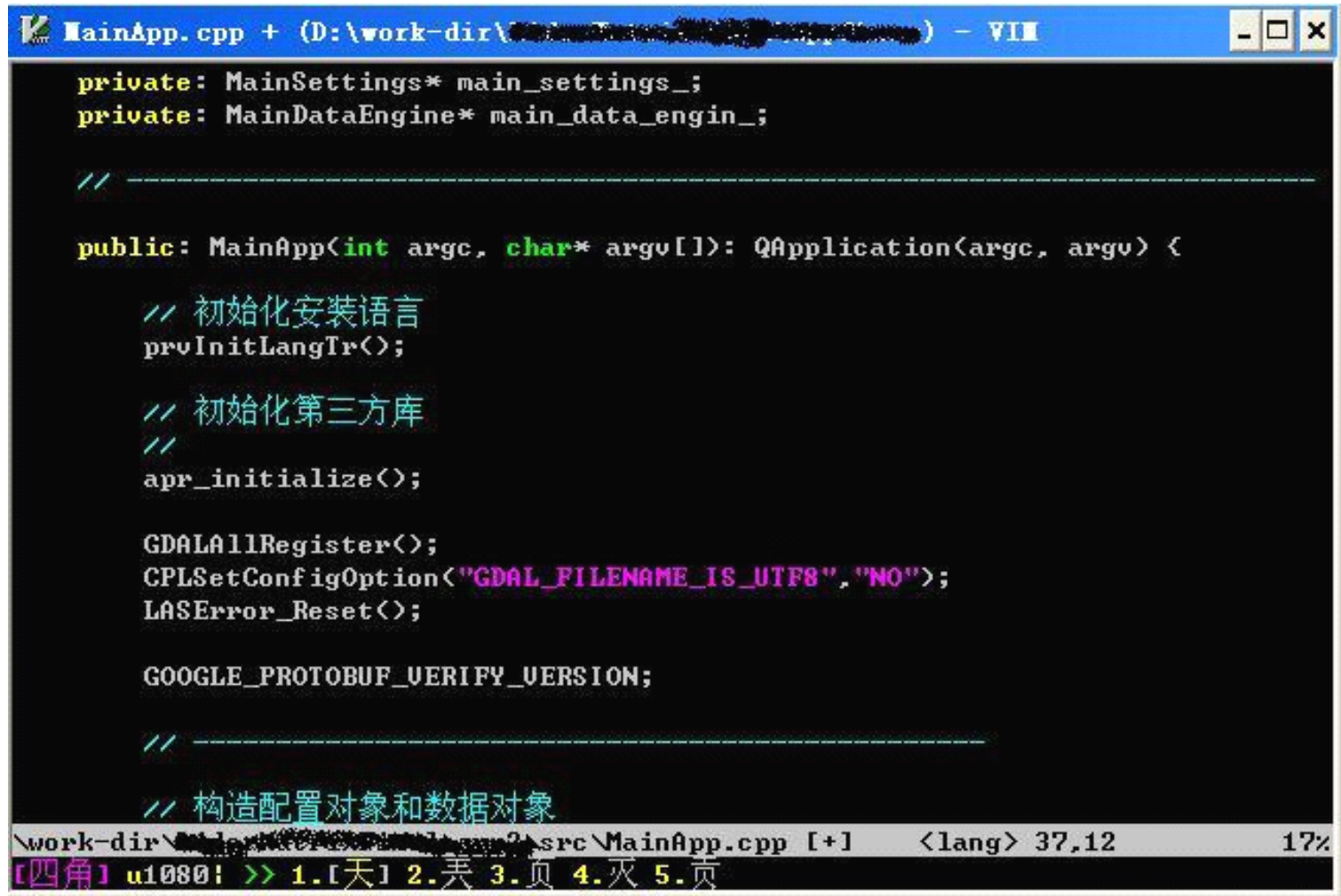
```
map.cnx - (帮助) - VIM
:ln[oremap] <lhs> <rhs>          mapmode-l          :ln  lnoremap
:cno[remap] <lhs> <rhs>          mapmode-c          :cno :cnoremap
在映射命令作用的模式中把键序列 <lhs> 映射为 <rhs>。禁止对映射后的 <rhs> 进行映射扫描。这个特性可以避免映射的嵌套和递归。通常用于重定义一个命令。<Vi 无此功能>

:unm[ap] <lhs>                    mapmode-nvo        :unm :unmap
:nun[map] <lhs>                    mapmode-n          :nun :nunmap
:vu[nmap] <lhs>                    mapmode-v          :vu  :vunmap
:xu[nmap] <lhs>                    mapmode-x          :xu  :xunmap
:sunm[ap] <lhs>                    mapmode-s          :sunm :sunmap
:ou[nmap] <lhs>                    mapmode-o          :ou  :ounmap
:unm[ap]! <lhs>                    mapmode-ic         :unm! :unmap!
:iu[nmap] <lhs>                    mapmode-i          :iu  :iunmap
:lu[nmap] <lhs>                    mapmode-l          :lu  :lunmap
:cu[nmap] <lhs>                    mapmode-c          :cu  :cunmap
在映射命令作用的模式中删除 <lhs> 的映射。该映射仍然可以在其它模式中保留其定义。
备注: <lhs> 包含末尾的空格。该映射取消操作_不会_生效:
      :map @e foo
      :unmap @e | print

:mapc[lear]                        mapmode-nvo        :mapc :mapclear
map.cnx [帮助][RO]                  70,46-64          5%
```

execute '**Inoremap** <buffer> <expr> ' . key . ' **VimPinyin_KeyFilter**(" . key . "')'

VimPinyin - 四角输入



```
private: MainSettings* main_settings_;
private: MainDataEngine* main_data_engine_;

// -----

public: MainApp(int argc, char* argv[]): QApplication(argc, argv) {

    // 初始化安装语言
    prvInitLangTr(<);

    // 初始化第三方库
    //
    apr_initialize(<);

    GDALAllRegister(<);
    CPLSetConfigOption("GDAL_FILENAME_IS_UTF8", "NO");
    LASError_Reset(<);

    GOOGLE_PROTOBUF_VERIFY_VERSION;

    // -----

    // 构造配置对象和数据对象

```

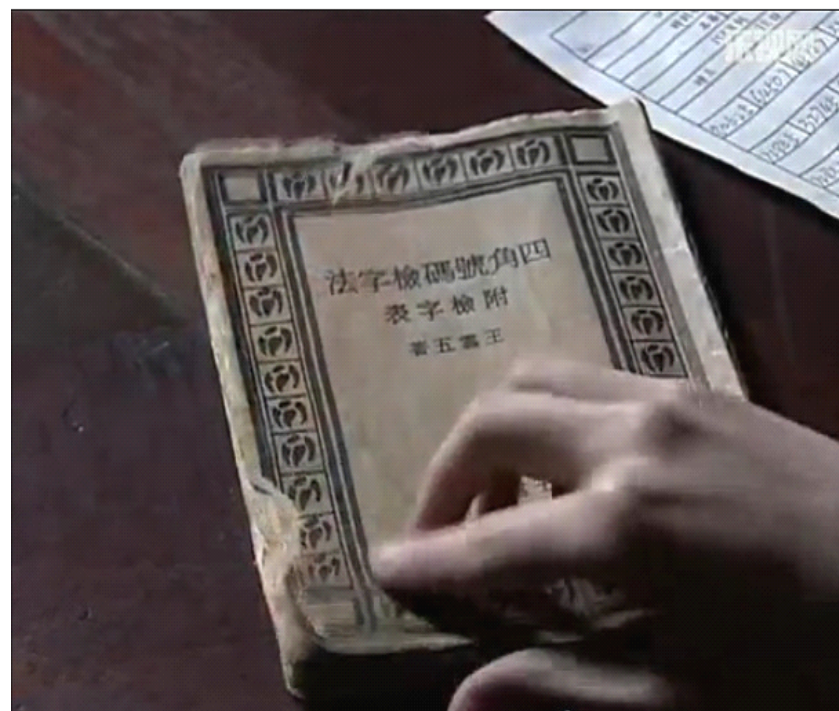
\work-dir\src\MainApp.cpp [+] <lang> 37,12 17%

[四角] u1080! >> 1.[天] 2.弄 3.页 4.灭 5.页

陆桥山是怎么被骗的？

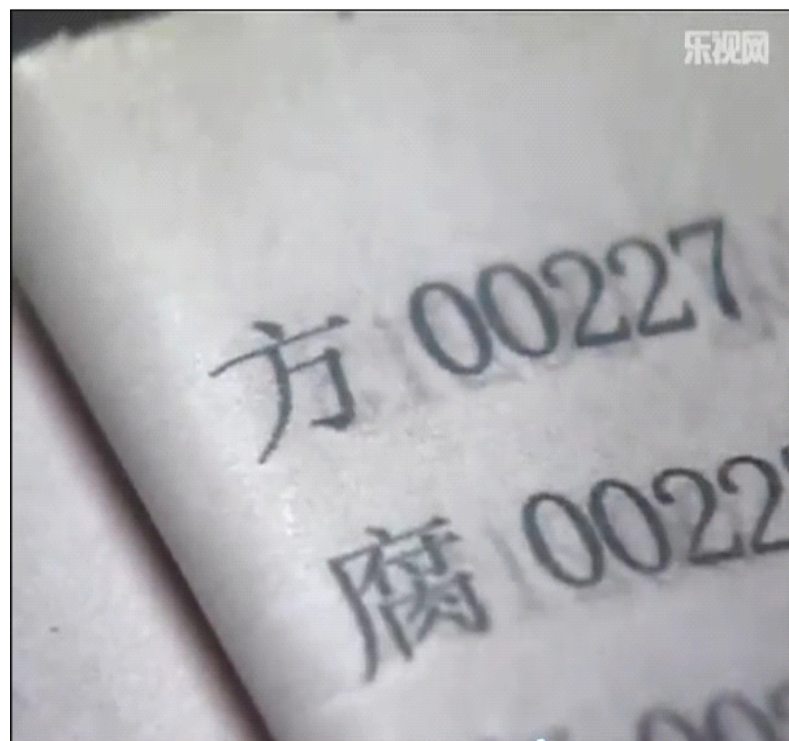
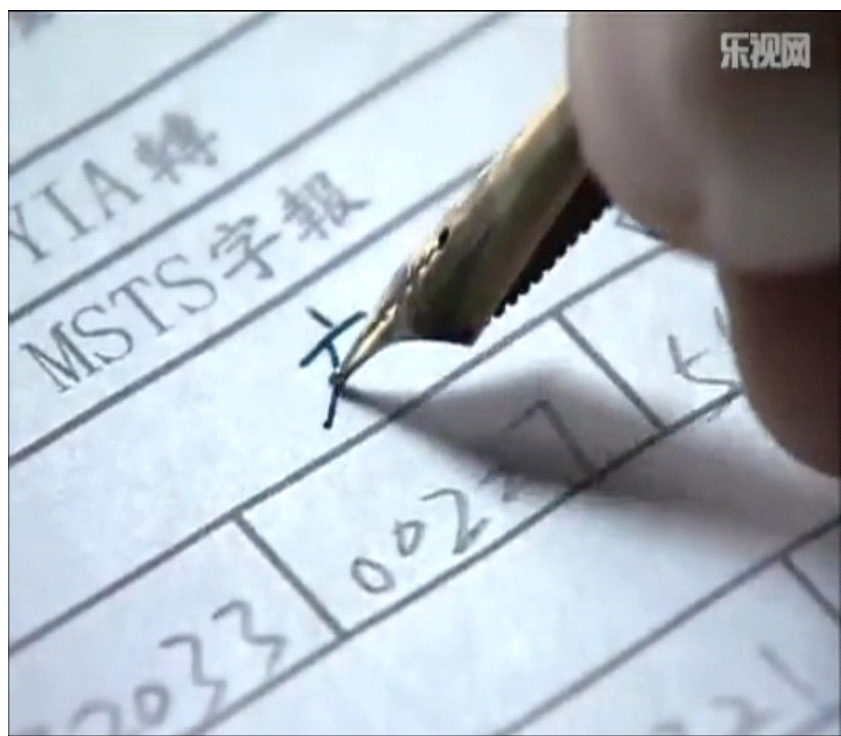


汤四毛的破绽



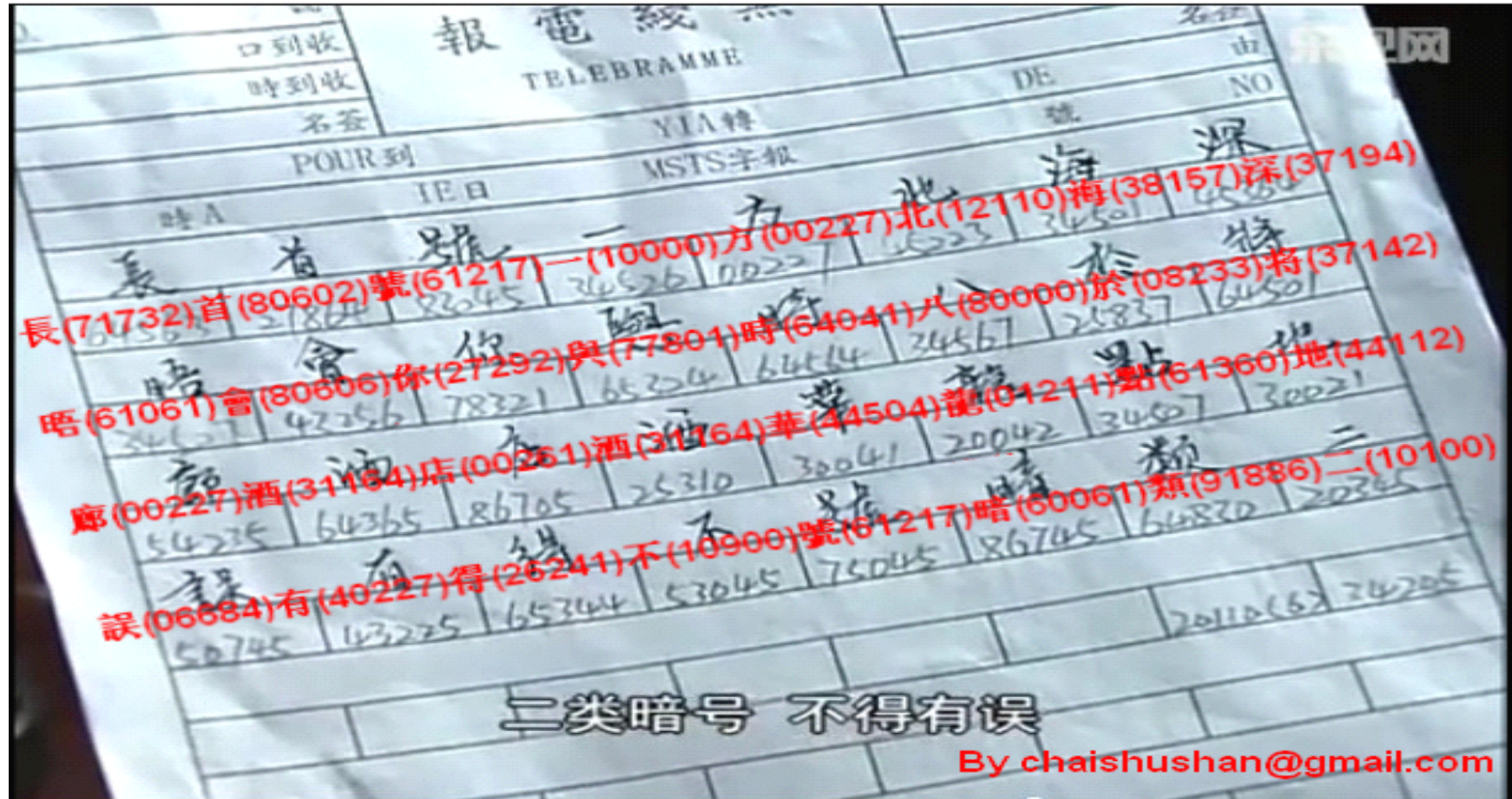
- <四角号码检字法-附检字表>是**1933年12月**出版
- 是**民国22年**，不是**民国20年**

译码就是查四角号码?!



从“方”的译码方式,可推测电文就是原文对应的四角号码,可是....

电文原来是个忽悠



- "方"字的左边右边电码忽然变了
- 译文除了"方"字之外居然没一个是对的

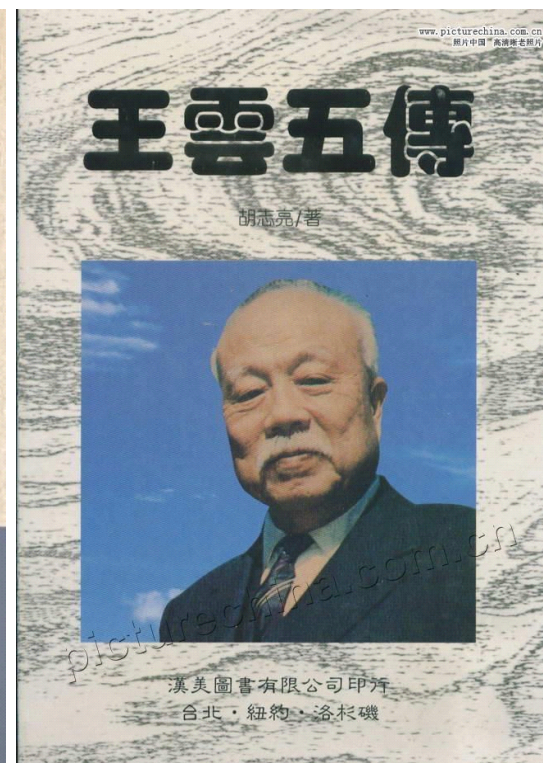
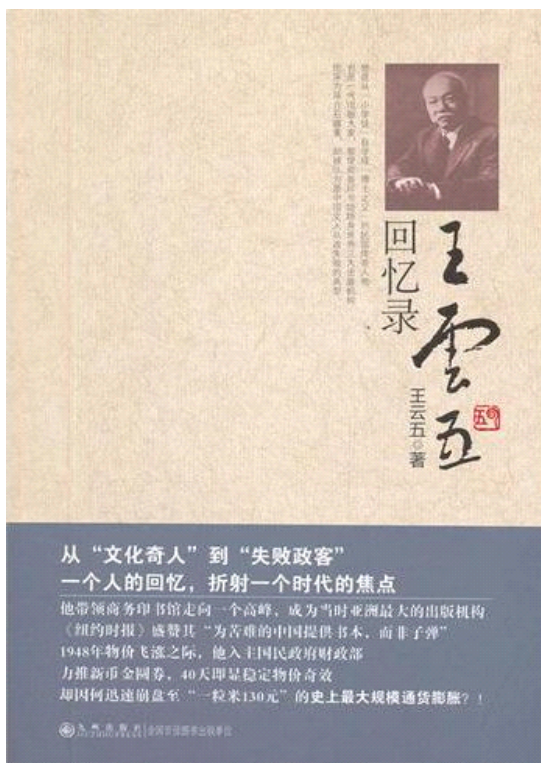
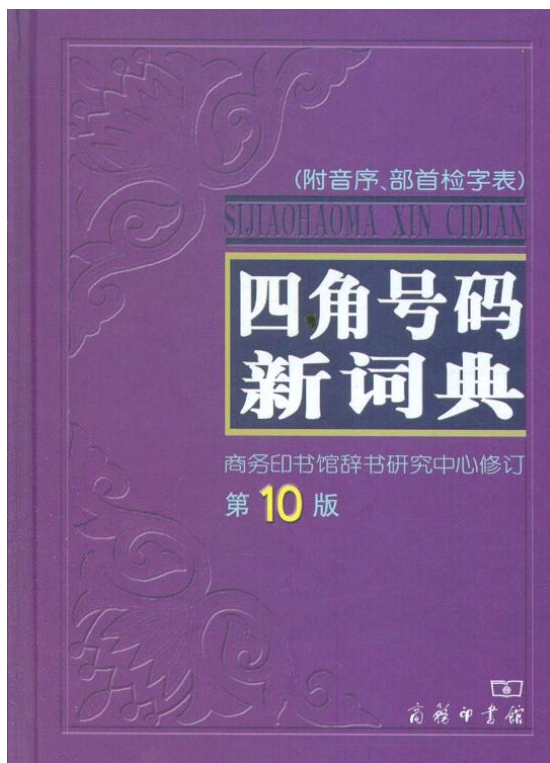
深入考证

- 四角号码七十年--钱伯城
- 王云五对我国图书馆事业的贡献及其启示
- 四角号码检字法亟待重视推广
- 四角号码新辞典

- 《王云五回忆录》2012年大陆出版
- 《王云五专》(台湾 胡志亮)
- 《学者的悲哀》民国名人 王云五 章节
- 《国民党四十三位：战犯的最后结局》 王云五 章节

- 王云五的成就总结：四百万!

相关图书



相关资源

- 四角输入 : <https://bitbucket.org/chai2010/im4corner>
- 四角号码 : <http://zh.wikipedia.org/zh-cn/四角号码>
- 四角号码表: <http://vimim-data.googlecode.com/svn-history/r128/trunk/data/vimim.4corner.txt>
- 汉语拼音表: <http://vimim-data.googlecode.com/svn-history/r128/trunk/data/vimim.pinyin.txt>
- 输入法扩展: <http://www.google.com/ime/pinyin/api.html>
- 调试器代码: <https://github.com/phuang/ibus-pinyin/blob/master/lua/lua-ext-console.c>
- VimPinyin: <http://code.google.com/p/vimpinyin/>

Thanks!