

# 关于“信息处理用 GB 13000.1 字符集汉字部件规范”中 若干问题的讨论

肖金卯 武云翠

## 摘要:

本文首先肯定了“信息处理用 GB 13000.1 字符集汉字部件规范”制定的重要意义和取得的积极作用,同时对部件的定义、基础部件的设置、部件中的一些错误提出了个人的看法,并对语言文字的规范发表了个人意见。

**关键词:** 部件 规范

## 1 引言

国家语言文字工作委员会于 1997-12-01 发布,1998-05-01 实施了 GF 3001-1997 “信息处理用 GB 13000.1 字符集汉字部件规范”(以下简称“部件规范”),对于中文信息处理,特别是对字形汉字键盘输入方法的研究具有规范作用,对于汉字部件的选取、部件的拆分有了遵循的依据,对制止编码混乱现象起到控制作用。其积极作用应当肯定,部件规范的推出在社会上也引起一些反响,本文仅就有关部件和规范的制定发表一些意见,不对之处可以讨论。

## 2 关于部件方面的意见

### 2.1 关于部件的定义

根据本规范术语部件的定义:部件是由笔画组成的具有组配汉字功能的构字单位。由此可见,组配汉字是部件的主要特征。分析本“部件规范”中的 560 个部件可以看出,有一些部件不具有组配汉字的功能,只组成它们自己,例如“书”、“飞”、“凸”等等,它们都是独体汉字,不和其他任何部件组成汉字,在 560 个部件中共有 25 个,这种独体汉字是否称之为部件值得研究。另外,在规范部件中的 560 个基础部件中有些基础部件也不符合基础部件的定义,例如,为什么“烏”和“马”可以拆成两部分,而部件“黑”、“熏”就可以不拆呢?实践证明部件“黑”、“熏”不拆比拆成两部分好得多。这些部件不拆既符合构字规律,也符合人们的认知习惯。

### 2.2 关于基础部件

在研究由规范部件组成的 GB 13000.1 字符集中的 20902 汉字时可以发现,560 个基础部件中有不少部件具有固定组合现象,即:一些基础部件中只和某个部件组合才具有组字功能,除去这个部件外不和其他任何部件构成构字单位,例如部件“𠂔”,GB 13000.1 字符集 20902 个汉字中 47 个汉字和部件“𠂔”有关,但都是和部件“十”组成“卑”构成一个构字单位。用“卑”组配汉字比用“𠂔”组配显得完整,不零乱,而且“卑”有音、形、义的概念,而“𠂔”只有形,无音、无义,纯属人为新造的一个部件。文字作为人们记录语言的符号,经长期约定成俗,越稳定越好,一般情况下最好不轻易制造新的文字,以汉字为例,除去少数汉字是后人所造外,绝大多数都是数千年约定成俗,而“部件规范”中的 560 个基础部件很大一部分为新造部件,对于人们学习使用带来很大困难。如果用“卑”代替“𠂔”,人们接受起来容易得多。类似部件“𠂔”的还有不少,例

如可以用“其”代替“其”；用“可”代替“丁”；用“至”代替“云”等等。如何科学定义基础部件值得讨论。

### 2.3 关于一些部件的合法性

这里合法性仅指计算机通用操作系统是否允许其输入、显示、传输及交换，规范部件确定为信息处理用，也就是用计算机进行的处理，能够进行计算机处理的前题条件是计算机里必须有其标准的代码，计算机才能完成各种操作，没有这个标准的代码只能在本机运行，无法上网通信。在 560 个规范部件中有一百多个是新造部件，例如“止”、“力”、“勿”等等，没有包括在 GB 13000.1 字符集里，这一百多个部件没有标准的代码，计算机视这些符号为非法字符，从而不支持这些部件进行网上交换。这部分部件应认真研究，充分论证。

### 2.4 漏拆的部件

经对 GB 13000.1 字符集中的 20902 个汉字拆分发现有的汉字无法进行拆分，因为 560 个部件中没有这样的部件，例如汉字“𠂇”（72eb）、“𠂈”（6f40）、“𠂉”（5b90）等等。“𠂇”没有右边部件，“𠂈”、“𠂉”没有下边部件，只能拆成笔画，怀疑是否拆错？因为在规范部件第 250 号部件“𠂇”的例字中列有汉字“𠂇”，“𠂇”右边“𠂇”是两横，而部件“𠂇”只有一横，不是一回事。还是漏掉？

### 2.5 笔画搭挂

“部件规范”第 4 页，《汉字基础部件表》使用规则中讲到：交重不拆（可拆成笔画）。少数不影响结构和笔数的笔画搭挂，按相接处理。什么是笔画搭挂，在术语中没有论述，在举例“幾”、“孝”的拆分写到（“人”搭挂在“戈”上）；（“子”搭挂在“乚”上）。不清楚所说的笔画搭挂和交重有什么本质区别？和“幾”、“孝”有关的汉字有几十个，字数不算少。另外，汉字“在”第二笔和第三笔算交重呢？还是算搭挂呢？“在”是一个使用频度很高的汉字，但是“在”拆成力、丨、土三个部件，这又如何解译呢？规范作为明文规定的标准，理应严格科学，规范部件中的每一条规定对每一汉字都应平等，不能存在任何特殊的汉字，这样才能起到规范化的目的，制止信息处理领域中汉字乱拆分现象。

### 2.6 认同部件

为了减少部件数量，“部件规范”规定因在字中所处的部位不同而产生了笔画变形或比例变化的部件，可与本部件表的相应部件认同使用。例举了四对认同使用的部件，这无可非议，我认为对于认同使用的部件应该有穷，全部例出，方便使用。否则不能算是一个完善的规范，因为一些字多数情况下当部件使用时笔画发生变化，例如牛字旁、车字旁、雨字头等等，如不例全，对初学者容易造成错误。

### 2.7 只组一个汉字的部件

在 560 个基础部件中有一批部件只组成一个汉字，且使用频率极低，例如：“心”、“𠂇”、“乚”等等，每个部件只组成一个汉字，分别为：厯、𠂇、乚。有无必要只为一个汉字再造一个部件，值得研究。

### 2.8 部件的检索

为了便于查找，“部件规范”附表中给出笔画排序汉字基础部件检索表，每个部件按首笔笔画排序，该检索表中有的部件首笔与笔画数和国家语言文字工作委员会语言文字规范《GB 13000.1 字符集汉字笔顺规范》（以下称“笔顺规范”）有出入，如，第 14 页 23 号部件“𠂇”，“部件规范”首笔为撇“丿”，而“笔顺规范”为竖笔画；同页 508 号部件“𠂇”首笔为横“一”，而“笔顺规范”为竖笔画；第 15 页 258 号部件“𠂇”“部件规范”首笔为横“一”，而“笔顺规范”

为折笔画；第 16 页 120 号部件“門”，“部件规范”首笔为折“乚”，而“笔顺规范”为竖笔画；第 16 页 521 号部件“龜”，“部件规范”为 16 画，而“笔顺规范”为 17 画。部件的笔顺、笔画很重要，因为一个部件可能涉及一批汉字，以部件“門”为例，与“門”有关的汉字就有 155 个之多。所以，部件的笔画、笔顺一定要准确。

## 2.9 其他

“部件规范”文本印刷中也有错误地方，例如第 178 号，应当是“氷”，首笔为“丨”，不是“丿”；再如，前面曾经提到的第 250 号部件“冊”的例字中列有汉字“𪛗”，右边不是“𪛗”；第 367 号例字中“逵”，第 6 笔是竖还是折？在 GB 13000.1 字符集笔顺规范里第 6 笔是竖；另外，在 G 列以外汉字专用的部件，均“\*”标识，在《汉字基础部件表》及附表中标识的有不一致的地方，如，第 360 号部件“尗”在《汉字基础部件表》漏标。

## 3 关于语言文字规范

文字的规范是国家统一的象征，社会文明程度取决文化的发达程度，而文化的发达取决规范的文字。历朝历代都注重语言文字的规范，规范的语言文字也促进了社会的发展。在中国历史上比较典型的如秦和唐两个朝代，前者推出“书同文”，规范了当时诸国的混乱用字现象，秦朝为后人留下重多文化遗产，这和文字的统一有一定的联系。唐朝是我国历史上鼎盛时期之一，唐是国土大统一，国力大发展，汉字大规范的时代，汉字楷书得到规范，唐朝的文明不仅对我国乃至世界的影响都是很大的。新中国成立后，非常重视语言文字规范，建国初期，中央为使人民大众迅速提高文化，由毛主席和周总理亲自领导研究制定汉语拼音方案，实践证明，《汉语拼音方案》在当代文化领域和信息处理技术中发挥了巨大作用，由于《汉语拼音方案》的科学性使其成为国际标准，被全世界广为采用。近年来我国有关部门制定了一系列语言文字及信息处理方面的标准和规范，全国人大还通过了《中华人民共和国通用语言文字法》。回顾我国语言文字规范的道路，有成功的经验，也有走弯路的地方。

规范或标准是体现一个国家的主权的象征，成功的规范应该具有权威性、群众性、准确性、完整性和连续性。仍以《汉语拼音方案》为例：在《汉语拼音方案》制定过程中不仅由国家领导人亲自领导，听取汇报，而且由全国著名权威专家组成委员会，多次公开征求意见，收到国内外群众来信 4300 多件，收到群众创制方案 3300 多个。虽然至今仍有人提出一些意见，但《汉语拼音方案》仍不愧为一个最佳方案。对比“部件规范”的制定过程，则深入广泛讨论不够。文字是经过长期约定俗成而形成的，不可能由某个人或几个人生造出来的，汉字的部件也是如此，每一个部件都有它的意义，制定一个新的部件都要经过反复论证，仔细推敲，在 560 个“部件规范”中有相当一部分部件是否具有规范性值得研究，突然间增加了一百多个人们默生的部件，作为规范实施，群众不容易接受。不能简单地说组配汉字功能的构字单位就认为是汉字的部件。一个成熟的规范应当相对稳定，不要经常改变，在汉字部件规范制定过程中，一个板本换一批人，互不兼容，不具备连续性。97 年重排本《语言文字规范手册》有一批简化偏旁，如“荣”（榮）去掉下边“木”；恋、（戀）去掉下边“心”等，1997-12-01 发布的“部件规范”中没有收入这两个部件，认为它们不是规范基础部件，即使它们组合成偏旁，由于它们不是汉字，也不允许当成部件使用。但在 2000-03-17 发布的中华人民共和国标准 GB 18030 双字节 4 区中又将这两部件收入。另外，

在该标准中还收入“部件规范”560个部件以外的一些部件，如“眷”字的上半部、“朝”字的左半部等等。短短几年里，都称之为规范或标准，究境谁对？使用者不知所措。

“书同文”指字样相同或统一，随着时代发展，“书同文”的概念也在发展，首先语言和文字有着密切关系，不仅要求“书同文”，而且还要求“语同音”，在当今信息化社会时代，对“书同文”又有了更加丰富的概念，如利用形码电脑输入汉字时，对部件的规范提出要求，在人名排序时，对汉字的笔画、笔顺规范提出了要求，特别是互联网络的发展，必须要求汉字的统一的标准代码。因此，对于一个完整的规范还应考虑“笔同画”、“画同顺”、“检同部”、“讲同义”、“拆同序”、“出同源”、“码同位”。

本文承蒙原国家语言文字委员会副主任傅永和研究员的指点，并作了重要修改，在此深表感谢。