



中华人民共和国国家标准

GB/T 15936.6—1996
idt ISO 8613-6:1989

信息处理 文本与办公系统 办公文件体系结构(ODA)和交换格式 第六部分:字符内容体系结构

Information processing—Text and office systems—
Office Document Architecture(ODA)and interchange format—
Part 6:Character content architectures

1996-04-10 发布

1996-12-01 实施

国家技术监督局 发布

目 次

前言	Ⅲ
ISO 前言	Ⅳ
1 范围	1
2 引用标准	1
3 定义	2
4 总则	2
4.1 类	2
4.2 内容	2
4.3 显现属性	2
4.4 控制功能	3
4.5 图形字符	3
4.6 间隔字符	3
4.7 内容信息的编码	3
4.8 内部结构	3
5 字符定位	4
5.1 基本概念	4
5.2 在行框内字符图像的定位	7
5.3 基本布局客体内行框的定位	13
6 字符成像	13
6.1 强调	13
6.2 字型选择	15
6.3 下标和上标	16
6.4 字符复合	16
7 字符显现属性的定义	16
7.1 共用显现属性	16
7.2 布局显现属性	21
7.3 逻辑显现属性	22
7.4 内容体系结构类属性	23
7.5 显现属性和布局导则之间的相互作用	23
8 字符内容部分属性	23
8.1 通用编码属性	23
8.2 其他编码属性	23
9 字符内容体系结构相关数据类型的正式定义	24
9.1 引言	24
9.2 显现属性的表示	24
9.3 编码属性的表示	27

9.4 非基本特性和非标准缺省的表示	27
10 图形字符	28
11 控制功能和字符 SPACE 的定义	28
11.1 共用控制功能	29
11.2 布局控制功能	33
11.3 逻辑控制功能	34
11.4 限定符	35
11.5 SP——间隔	35
12 内容布局过程	35
12.1 引言	35
12.2 对可处理内容的内容布局过程	36
12.3 格式化可处理内容的内容布局过程	40
12.4 格式化内容的内容布局过程	41
13 内容成像过程	41
13.1 引言	41
13.2 格式化内容的内容成像过程	41
13.3 格式化可处理内容的内容成像过程	41
14 显现属性和控制功能之间的相互作用	41
15 字符内容体系结构类的定义	42
15.1 格式化字符内容体系结构类	42
15.2 可处理字符内容体系结构类	42
15.3 格式化可处理字符内容体系结构类	43
附录 A(提示的附录) 内容体系结构类汇总	45
附录 B(提示的附录) 字符内容体系结构层	49
附录 C(提示的附录) 控制功能的编码表示	59
附录 D(提示的附录) 客体标识符汇总	60
附录 E(标准的附录) ODL 用字符内容特定属性的 SGML 表示	61

前 言

本标准等同采用国际标准 ISO 8613-6:1989《信息处理 文本和办公系统 办公文件体系结构 (ODA)和交换格式 第6部分:字符内容体系结构》。

通过制定这项国家标准,使信息处理文本和办公系统的文件便于交换。

本标准与 ISO 8613-6 的区别如下:

a) 正文和附录中引用其他标准时,用我国的标准编号代替相应的国际标准编号,其对应关系是:

GB 1988—1989	代替 ISO 646:1983
GB 2311—1990	代替 ISO 2022:1986
GB 5261—1985	代替 ISO 6429:1983
GB 8565—1988	代替 ISO 6937:1983
GB/T 14814—1993	代替 ISO 8879:1986
GB/T 15936	代替 ISO 8613:1989
GB/T 16212—1996	代替 ISO 8824:1990

b) 根据编写国家标准的基本规定,本标准保留了被采用国际标准的前言,同时增加“前言”。

GB/T 15936 在《信息处理 文本与办公系统 办公文件体系结构和交换格式》总标题下,目前包括以下八个部分:

- 第一部分(即 GB/T 15936.1):引言和导则;
- 第二部分(即 GB/T 15936.2):文件结构;
- 第四部分(即 GB/T 15936.4):文件轮廓;
- 第五部分(即 GB/T 15936.5):办公文件交换格式(ODIF);
- 第六部分(即 GB/T 15936.6):字符内容体系结构;
- 第七部分(即 GB/T 15936.7):光栅图形内容体系结构;
- 第八部分(即 GB/T 15936.8):几何图形内容体系结构;
- 第十部分(即 GB/T 15936.10):形式规范。

本标准的附录 E 是标准的附录。

本标准的附录 A、附录 B、附录 C 和附录 D 都是提示的附录。

本标准由中华人民共和国电子工业部提出。

本标准由电子工业部标准化研究所归口。

本标准起草单位:电子工业部标准化研究所。

本标准主要承办人:高健、冯惠、黄伟敏。

ISO 前言

ISO(国际标准化组织)是由各个国家标准化机构(ISO的成员体)联合组成的一个世界性组织。该组织通过其各个技术委员会进行国际标准的制定工作。凡是对于已设有技术委员会的某一专业感兴趣的每一个成员体,都有权参加该技术委员会。与ISO有联系的官方和非官方国际组织也可参与国际标准的制定工作。ISO与国际电工委员会(IEC)在电子技术标准化的所有方面都进行密切合作。

各个技术委员会提出的国际标准草案,须先分发给各成员体表决通过后,再由ISO理事会批准为国际标准。根据ISO工作导则,国际标准至少需要投票成员体的75%赞成。

国际标准ISO 8613-6是由ISO/TC 97“信息处理系统”技术委员会制定的。

目前,ISO 8613由七个部分组成。

- 第1部分:引言和总则;
- 第2部分:文件结构;
- 第4部分:文件轮廓;
- 第5部分:办公文件交换格式(ODIF);
- 第6部分:字符内容体系结构;
- 第7部分:光栅图形内容体系结构;
- 第8部分:几何图形内容体系结构。

注:目前没有第3部分。

本国际标准可增补新的部分。

本国际标准是与下述标准并行制定的:

- ECMA 101:1985:办公文件体系结构;
- CCITT T. 73:1984:用于远程信息处理服务的文件交换协议;
- CCITT T. 410 系列建议:1988:开放文件体系结构(ODA)和交换格式。

本标准包含五个附录:

- 附录A(提示的附录):内容体系结构类汇总;
- 附录B(提示的附录):字符内容体系结构级;
- 附录C(提示的附录):控制功能的编码表示;
- 附录D(提示的附录):客体标识符汇总;
- 附录E(标准的附录):ODL用字符内容特定属性的SGML表示。

中华人民共和国国家标准

信息处理 文本与办公系统 办公文件体系结构(ODA)和交换格式 第六部分:字符内容体系结构

GB/T 15936.6—1996
idt ISO 8613-6:1989

Information processing—Text and office systems—
Office Document Architecture(ODA)and interchange format—
Part 6:Character content architectures

1 范围

GB/T 15936 的目的是便于文件的交换。

在 GB/T 15936 中,文件被认为是可以包括图片和表格材料的诸如备忘录、信件、发票、表格和报告的一些项。文件中使用的内容元素可以包含图形字符、几何图形元素和光栅图形元素,也可以同时包含全部这些内容。

注: GB/T 15936 被设计为允许扩充的,以便在其中包括印刷特征、颜色、传单及其他诸如声音的内容类型。

GB/T 15936 适用于以数据通信方式或交换存储媒体方式进行的文件交换。

GB/T 15936 为下列两个目的或其中之一而提供文件的交换:

- 允许按始发者的意图进行显现;
- 允许进行诸如编辑和重新格式化处理。

交换中文件的组成可以有几种形式:

- 允许文件显现的格式化形式;
- 允许文件处理的可处理形式;
- 既允许显现又允许处理的格式化可处理形式。

GB/T 15936 还提供了用于处理被交换文件的 ODA 信息结构的交换。

此外,GB/T 15936 还允许包含一种或多种不同类型内容的文件交换,诸如字符文本、图像、图形和声音。

GB/T 15936 的这一部分

- 定义了可与 GB/T 15936.2 中规定的文件体系结构一起使用的字符内容体系结构;
- 定义了与这些字符内容体系结构一致的内容内部结构;
- 定义了适用于这些字符内容体系结构的显现在基本布局客体中的定位和成像;
- 定义了适用于这些字符内容体系结构的显现属性和控制功能;
- 描述了一个内容布局过程,它与 GB/T 15936.2 中所描述的文件布局过程一起,决定了基本布局客体中字符内容的布局 and 这些基本布局客体的尺寸。

2 引用标准

下列标准所包含的条文,通过在本标准中引用而构成为本标准的条文。本标准出版时,所示版本均为有效。所有标准都会被修订,使用本标准的各方应探讨使用下列标准最新版本的可能性。

GB 1988—1989 信息处理 信息交换用七位编码字符集(idt ISO 646:1983)

国家技术监督局1996-04-10批准

1996-12-01实施

GB 2311—1990 信息处理 七位和八位编码字符集 代码扩充技术(idt ISO 2022:1986)
GB 5261—1985 信息处理 文字和符号成形设备用的增补控制功能(idt ISO 6429:1983)
GB 8565—1988 信息处理 文本通信用编码字符集(idt ISO 6937:1983)
GB/T 14814—1993 信息处理 文本和办公系统 标准通用置标语言(SGML)(idt ISO 8879:1986)
GB/T 16262—1996 信息处理系统 开放系统互连 抽象语法记法一(ASN.1)规范(idt ISO 8824:1990)
ISO 7350:1984 文本通信 图形字符子表的登记
ISO 9541-5¹⁾ 信息处理 字型和字符信息交换 第5部分:字型属性和字符模式

3 定义

GB/T 15936 的这一部分所用的术语,采用 GB/T 15936.1 中给出的术语定义。

4 总则

4.1 类

字符内容体系结构可分为三类:

——用于格式化内容的一种字符内容体系结构,它可将文件内容按照始发者的意愿显现出来(如,打印或显示)。格式化内容可用于任一基本成分中;

——用于可处理内容的一种字符内容体系结构,允许对文件内容进行处理(如,编辑或格式化)。可处理内容能用于任一基本逻辑成分中;

——用于格式化可处理内容的一种字符内容体系结构,它既允许对文件的内容进行处理也允许按照始发者的意愿将其表示出来。格式化可处理内容能用于任一基本成分中。

4.2 内容

与字符内容体系结构一致的基本成分内容是一个字符串。该字符串是由将基本成分内容部分的各字符串链接在一起而形成的。

内容字符串由图形字符、控制功能和间隔字符的复合而组成。

4.3 显现属性

显现属性适用于基本逻辑和布局成分。它们所包含的信息规定了与这些基本成分内容所用的图形字符的布局、成像和选择有关的初始条件。这些条件中的某些条件可由包含在内容中控制功能进行修改。

显现属性分为如下几类:

——逻辑显现属性,能够与处理的和格式化可处理的字符内容相联系。这些属性在内容布局过程中起作用,但在内容成像过程中则被忽略;

——布局显现属性,能够与格式化的和格式化可处理的字符内容相联系。这些属性在内容成像过程时起作用。它们既可由内容布局过程生成也可由产生或编辑格式化或格式化可处理内容的处理过程生成;

——共享显现属性,能够与所有各类字符内容体系结构相连。这些属性在内容布局和成像处理的任一个或二个过程中均可起作用。

注:显现属性可以三种方式中的一种适用于基本成分的内容。可将它们直接规定为一个客体或客体类描述,或可通过与客体或客体类描述有关的显现式样对其进行规定。或者,也可利用缺省值表(见 GB/T 15936.2)将它们间

1) 正在印刷。

接地用于客体或客体类描述。

4.4 控制功能

带有 0 个或多个参数控制功能可以规定与后续图形字符的布局或成像有关的信息。控制功能也可用来扩充或取代正在使用的图形字符集。所有控制功能的应用范围都只限于它们所出现在的基本成分中。

控制功能的分类与显现属性的分类相似：

——逻辑控制功能，能够用于可处理的和格式化可处理的字符内容。这些控制功能在内容布局过程中起作用，但在内容成像过程中则被忽略；

——布局控制功能，能够用于格式化和格式化可处理的字符内容。这些控制功能在内容成像过程中起作用。它们由内容布局过程生成。或者，它们也可由产生或编辑这些格式化或格式化可处理内容的处理过程（不在 GB/T 15936 中叙述）插入；

——共享控制功能，能够用于所有各类字符内容体系结构。这些控制功能在内容布局及成像过程的任一个或二个过程中均可起作用。

此外，格式化可处理内容可以包含被称为限定符的控制功能。这些限定符用于指示已由内容布局过程插入的一个或多个图形字符和/或控制功能串（见第 12 章）。被限定的图形字符和/或控制功能只在内容成像过程中起作用。限定符则通过删除它们及所限定的字符序列而在内容布局处理过程中起作用。

4.5 图形字符

在基本成分内容中所用的图形字符集和它们的编码表示，由显现属性和代码扩充控制功能规定（见第 10 章和 11.1.17）。

任一个或多个图形字符集都可用于基本成分的内容中，但要受到与所用的特有内容体系结构相关的一些限制，并按 GB 2311 进行适当的指明和调用。

包含在图形字符集中的任何非进格字符均不可单独使用，而只能与进格字符复合使用。

4.6 间隔字符

字符“间隔(SP)”既可被当作一个逻辑控制功能也可被当作一个图形字符。作为图形字符，它具有一个无图形符号出现的图形表示。作为控制功能，它指示一个潜在的行间断点（见 12.2.1.3.2）。

注：NBSP（无间断间隔）及任一定宽的间隔字符，如“数字间隔”，“em 间隔”及“en 间隔”都被看作图形字符，即：不把它们当作行间断点。

4.7 内容信息的编码

在内容部分中的内容信息的编码表示与 GB 2311 中规定的规则一致。

控制功能的编码表示由 GB 5261 定义，并归纳于附录 C。

4.8 内部结构

4.8.1 格式化内容

格式化内容是已把所有的与这些内容的布局 and 成像有关的必要信息都做了规定的内容。这种形式的内容是要按规定来成像，而不要在编辑过程中修改或重新格式化。

与格式化字符内容体系结构一致的基本成分内容包括一行或多行字符。每对相邻的行由一个硬行终止符分开。最后一行（或仅有的一行）可用也可不用硬行终止符终止；基本成分内容的结束也就隐含地终止了最后一行。

4.8.2 可处理内容

可处理内容是尚未经过布局的内容。这种形式的内容适合于由编辑处理修改。

注：编辑处理与实现有关，且不在 GB/T 15936 中描述。

为了对这种形式的内容成像，就要将内容进行内容布局处理（见第 12 章），从而将可处理内容转换为格式化内容（见 4.8.1）或格式化可处理内容（见 4.8.3）。

为配合对可处理内容的处理（即编码或布局处理），已定义了一些逻辑显现属性和控制功能（见第 7

章和第1章)。此外,字符“间隔”既可用作图形字符也可用作控制功能,在对内容进行布局时,用于指示将于何处出现行间断。

与可处理字符内容体系结构一致的基本成分内容包括一个或多个字符序列。每对相邻的字符序列由硬行终止符控制功能分开。最后(或仅有的一个)字符序列可用也可不用硬行终止符终止。

若在基本逻辑成分内容的尾端省略了硬行终止符,而又有另一个基本逻辑成分与之相连(见GB/T 15936.2),则最后一个字符序列将连接到下一个基本逻辑成分的内容中。在所有其他情况下,基本逻辑成分内容的结束便隐含地终止了最后一个字符序列。

字符序列的划分表示了基本逻辑成分的可处理内容的内部结构。每个字符序列都是无名的,既没有名字也没有标识符与之相联系,并且在字符序列之间除序列本身外,不存在任何关系。

4.8.3 格式化可处理内容

格式化可处理内容是将格式化内容和可处理内容作为其子集而构造的内容。除了可能包含已由内容布局处理而加上的附加控制功能和图形字符之外,其结构与可处理内容的结构相同。同时,除了可能包含逻辑控制功能和限定符外,其结构也与格式化内容的结构相同。

因此,可通过删去(或忽略)所有的布局控制功能、所有的限定符以及所有出现在那些限定符之内的控制功能和字符而将格式化可处理内容转化为可处理内容。

或者,也可通过删去(或忽略)所有的逻辑控制功能和限定符但保留限定符内的控制功能和字符而将格式化可处理内容转化为格式化内容。

注:由格式化可处理内容到可处理内容的转换是可逆处理过程(所提供的相同布局约束条件也可适用于内容布局处理过程),但从格式化可处理内容到格式化内容的转换却是不可逆的。

与格式化可处理字符内容体系结构一致的基本成分的格式化形式包含有一行或多行字符,每对相邻的行由一个硬行终止符或一个软行终止符分开,最后一行(或仅有的一行)可用也可不用硬终止符终止,基本布局成分内容的结束也隐含着终止最后一行。

与格式化可处理字符内容体系结构一致的基本成分的可处理形式包含有一个或多个字符序列。每对相邻的字符序列由硬行终止符分开。最后的一个(或仅有的一个)字符序列可用也可不用硬行终止符终止。

若在基本逻辑成分内容的尾端省略了硬行终止符,而又有另一个基本逻辑成分与之链接(见GB/T 15936.2),则最后一个字符序列将连接到下一个基本逻辑成分的内容中。在所有其他情况下,基本逻辑成分内容的结束便隐含着终止最后一个字符序列。

软行终止符用作字符序列内行与行之间的分隔符。字符序列的划分表示了基本逻辑成分内容的内部结构。每个字符序列都是无名的,既没有名字也没有标识符与之相联系,在字符序列之间除序列本身外,不存在任何关系。

5 字符定位

本章规定了在基本布局客体内字符的定位方式。以助人们对与字符定位有关的显现属性及控制功能的理解。

本章给出了对按照ISO 9541-5所定义的任一种字型的定位。本章还为在同一基本布局客体内从属于不同字型的字符的定位提供了必要的条件。

5.1 基本概念

5.1.1 字符字型

GB/T 15936这一部分的上下文中,术语“图形字符”是一个抽象概念;该术语所代表的是用于表示信息的图形符号集中的一个成员。而“字符图像”则代表一个图形字符在显现媒体上的呈现。

字型是字符图像集,这些字符图像通常带有共同的设计和大小。字型属性集是与整套字型相联系的,而字符属性集则是与每个单独的字符相联系的。这些属性由ISO 9541-5规定。

字型属性主要用于为接收者标识始发者所用的字型种类,而且在还没有特定的字型可用时,字型和字符属性便成为接收者在其他已有的字型中寻找合适的替换字型的指南。

关于在基本客体中指明和调用不同字型的进一步的信息在第6章中给出。

5.1.2 方向

在 GB/T 15936 这一部分的上下文中,所有的方向都是由与某一特定的参考方向沿逆时针方向所旋转的角度(以°为单位)来表示的(图1中给出了一个示例)。

字符走向是指在一个行框(在5.1.7中定义)内连续的字符图像的前进方向,并由与布局客体的水平方向相对的方向来表示(见图4)。

行前进是指在基本布局客体内连续行框的前进方向,并由相对于字符走向的方向来表示(见图11)。

字符取向是指与字符走向相关的字符基线(在5.1.3中定义)的方向。

对字符走向、行前进和字符取向而言,只有一个值可规定为基本成分。

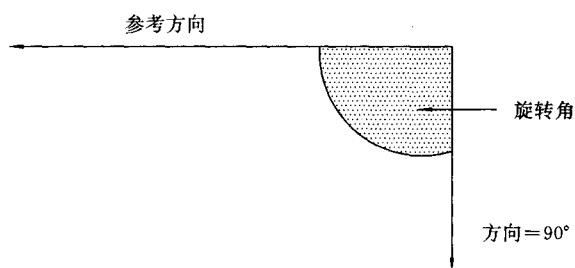


图1 方向的例子

5.1.3 字符图像模型

定位点是指与字符图像有关的参考点(见图2)。它用于在一个行框内字符图像的定位。转移点也是指与字符图像有关的参考点(见图2)。它用于为下一个字符图像定位。

字符基线是一条穿越字符图像的假设线,用于定义字符取向。当字符图像处于正视取向时,字符基线是一条水平直线(见图3)。

对于每一个所要使用的字符取向都要定义定位点和转移点(见图3);即:在字型描述中定义与所要求的字符取向相对应的“书写方式”或者,在文件应用轮廓中定义后退。

出界是指超出定位点和转移点的那部分字符图像(见图2)。

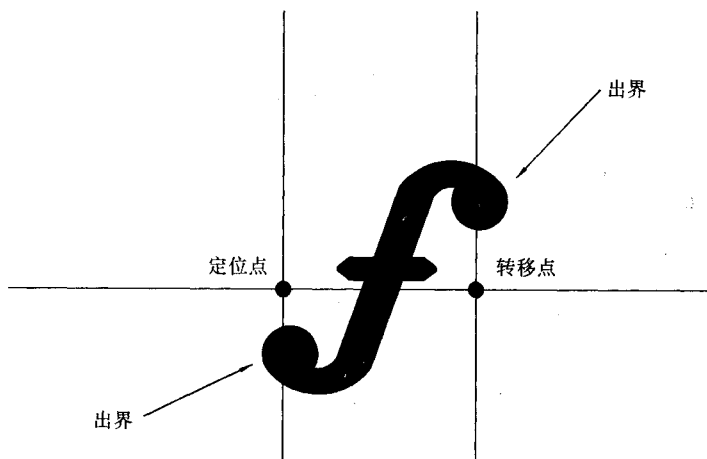


图2 出界的图示

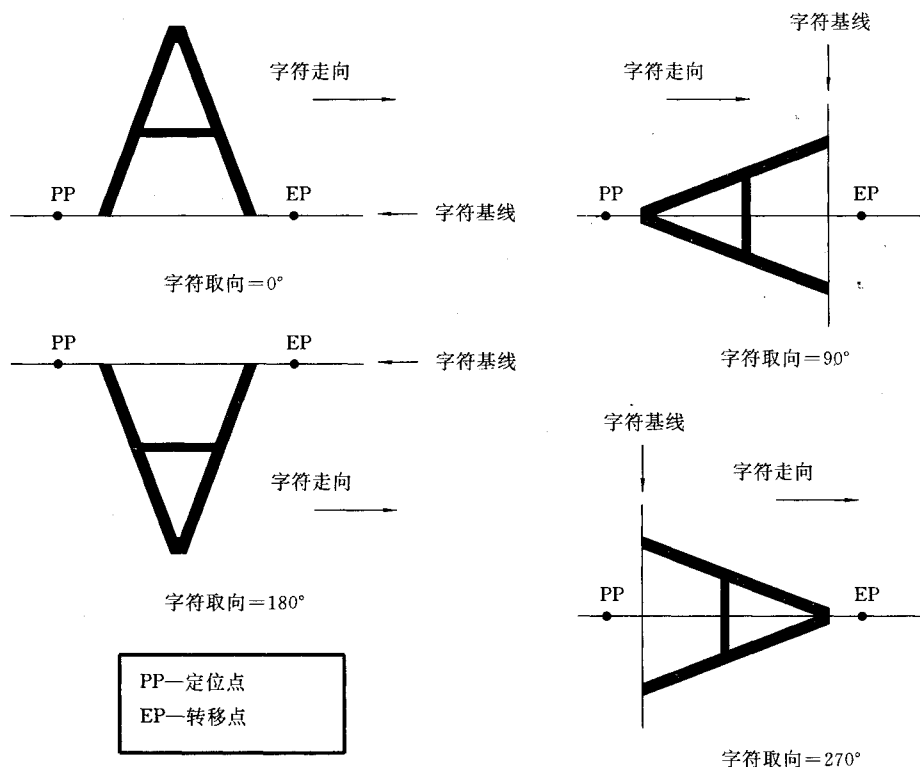


图3 字符图像定位的参考点

5.1.4 字符间距

字符间距的概念只适用于选择了固定间距的字型。字符间距(与字符间的间隙一起)用于决定在一个行距内字符图像之间的距离,如 5.2.1 中规定。

当字符与字符之间的间隙等于零时,字符间距等于连续的字符图像的定位点间的距离。

字符间距与字符图像的定位点和转移点之间的距离无关。

5.1.5 动作位置

动作位置是类似于光标那样的成像设备中的抽象概念。在假设按时序方法处理字符串的场合下,此概念用于对控制功能的定义(见第 11 章)。

动作位置指示了在基本布局客体的定位区内的一个点,在这个点上将进行下一个字符(图形字符或控制功能)规定的操作。

如果下一个字符是一个图形字符,其字符图像将定位在动作位置处的定位点,而动作位置则沿字符走向的方向前进,间距的值则由 5.2.1 规定。如果下一个字符是一个控制功能,则会导致动作位置移到定位区内的另一个点。

5.1.6 定位区

定位区是一个矩形的区域,整个地包含在一个基本布局客体内,定位点和转移点都在此区内进行定位(见图 4)。字符图像的出界允许延伸到定位区之外,但不允许超出基本布局客体的边缘。

定位区的两条边缘分别称为起始边缘和结尾边缘(见图 4)。起始边缘和结尾边缘的定义是这样决定的:从起始边缘至结尾边缘的方向要与字符走向的方向相同。

定位区的另外两条边缘分别称为上边缘和下边缘。上边缘和下边缘的定义是这样决定的:从上边缘至下边缘的方向要与行前进的方向相同。

起始和结尾边缘分别比基本布局客体的相应边缘缩进一段距离,此距离称为出界位移量(见图 4)。出界位移量的规定取决于要在定位区中定位的字符的字型。选择它的值时要使得当多个字符图像的序列在一个基本布局客体内定位时,任一字符图像及其出界部分都不会超出基本布局客体的边缘。

5.1.7 行框

在定位区内,一个字符图像序列在一个称为行框的区域内定位(见图 11)。每个行框均为一个从定位区的起始边缘至结尾边缘的矩形区域。

每个行框包含一个称为行原点位置的参考点(见图 4)。此点用于在基本布局客体内行框的定位。它也作为每一行的第一个图形字符或控制功能的动作位置。

每个行框包含一个称为参考线的假想线(见图 4)。参考线沿字符走向的方向穿越行原点位置。它在行框内从起始边缘延伸至结尾边缘,并用于字符图像的对齐。

行框的长度等于起始边缘至结尾边缘之间的距离;行框的宽度(或高度)则等于行框正向外延和反向外延的总和(见图 4)。行框反向外延是指从参考线沿着行前进方向的反方向至行框边缘之间的距离;而行框正向外延则是指从参考线沿着行前进方向至行框边缘之间的距离。

正向和反向外延的值取决于在行框内所用的字符字型的最大外延(垂直于参考行测量)。在决定正向和反向外延时要考虑到垂直于字符走向的字符图像的所有位移量的值,如:下标、上标和并行注释。

注:作为一个例子,取一种拉丁字型,其字符走向为 0° ,行前进方向 270° ,当在行框内使用单个字型时,其正向和反向外延分别等于其最大右向外延和最大左向外延,如 ISO 9541-5 中规定。

5.2 在行框内字符图像的定位

连续的字符图像在一个行框内定位时所取的方向是字符走向的方向。

除了某些成像为下标、上标或并行注释的字符,字符图像的定位点需与参考线对齐。

下列几个因素对沿参考线定位字符图像有影响:

- 字符间的间距;
- 对齐;
- 制表;
- 字符排序;
- 并行注释;
- 下标/上标;
- 双字紧排;
- 第一行位移;
- 项化。

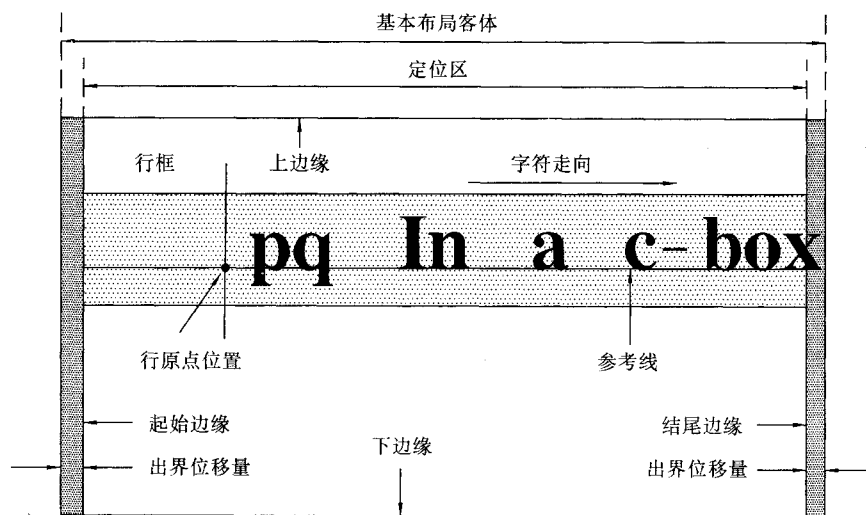


图 4 字符定位概念的图示

5.2.1 字符间距

字符间的间隙是相邻字符图像定位点之间间距沿字符走向方向的一个附加量(见 5.1.4)。负值则表示相邻字符图像间间距的一个缩小量。

相邻字符图像的定位点之间的距离既可以是一个定量也可以是一个变量,这取决于字型,具体如下:

——对于具有固定间距的字型,相邻字符图像定位点之间的距离独立于字符,并且是字符间距(由显现属性和控制功能规定)和字符之间的间隙之和(见图 5)。

——对于具有可变间距的字型,相邻字符图像定位点之间的距离与该字符有关,即:通常与该字符的定位点和转移点之间的距离有关,并且是该字符的净转移量(由字型决定)与字符间距的和(见图 6)。

间隔的宽度,即 SP 字符图像的宽度,可如下决定:

——若某一个 SP 字符,处于一行中第一个图形字符之前,并后随于软行终止符,或者处于一个软行终止符之前,且后随于一行中的最后一个图形字符,则该 SP 字符的宽度等于零;

——在固定间距字型中,其缺省宽度等于字符间距;

——在可变间距字型中,其宽度由字型隐含地定义;

——对于所有的字型,其宽度可由控制功能规定。

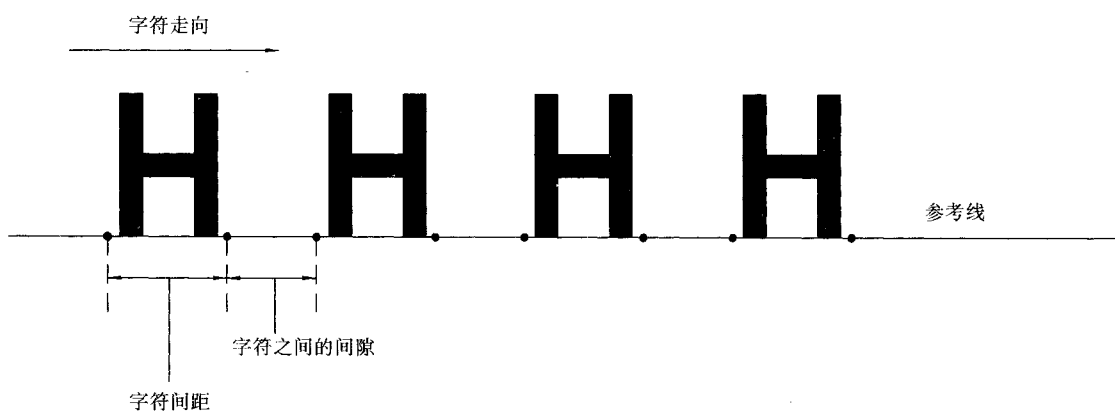


图 5 固定间距字型的间距

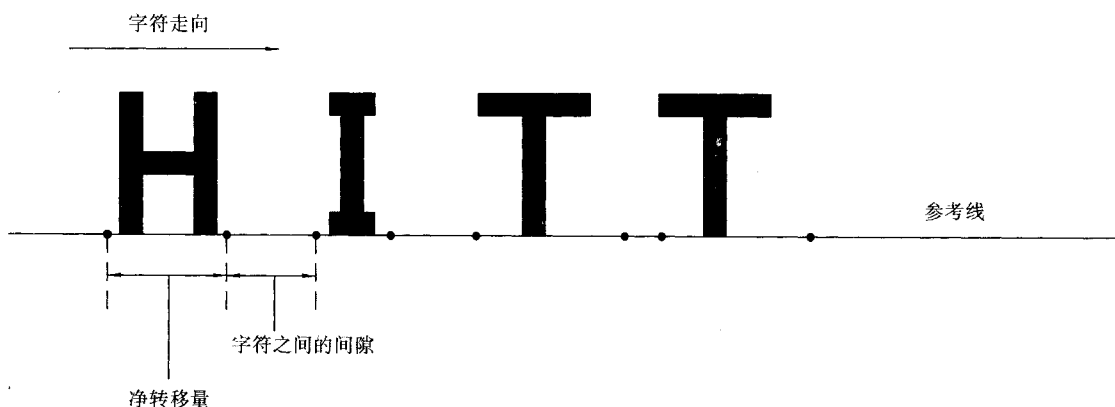


图 6 可变间距字型的间距

5.2.2 对齐

字符图像按照对齐属性以如下方式在行框内定位:

——起始对齐:将第一个字符图像的定位点置于行原点位置;

——结尾对齐:将最后一个字符的转移点置于定位区的结尾边缘;

——居中:使从行原点位置到第一个字符图像定位点的距离近似地等于从最后一个字符图像转移点到定位区结尾边缘的距离;

——齐行:通过适当设置间隔字符的宽度和/或字符间隙的大小,使得第一个字符图像的的定位点置

于行原点位置,并且使最后一个字符图像的转移点与定位区结尾边缘相重叠。

5.2.3 制表

字符图像在参考线上的位置可由一系列制表停止来控制。每个制表停止规定位于参考行上的一个相对于定位区起始边缘的点。

一个字符图像串可由夹在文本中的控制功能设置于一个制表停止处。这个串可以按如下规定进行起始对齐、结尾对齐、居中或环绕该串中的一个或多个特定字符对齐(见图 7):

——起始对齐:该串中第一个字符图像的定位点置于制表停止处;

——结尾对齐:该串中最后一个字符图像的转移点置于制表停止处;

——居中:将串如此放置,使得其第一个字符图像定位点和最后一个字符图像转移点到制表停止处的距离近似相等;

——环绕对齐:将该串中的第一次出现的一组特定的字符中的第一个字符图像的定位点定位于制表停止处。如果该组特定字符没有出现在与制表停止相关的文本中,则其缺省对齐方式为如上所规定的结尾对齐。

制表停止			
起始对齐	结尾对齐	居中	环绕对齐
This	This	This	124.56
is	is	is	23.5
start	end	centred	1254.76
aligned	aligned	text	32.465
text	text		342

图 7 制表

5.2.4 字符排序

在一个基本客体内容中,其交换顺序永远是所用的语言的阅读顺序。

在某些语言的情况下,如阿拉伯语和希伯来语,其字母数字文本是从右至左阅读,而其数字文本则从左到右阅读,故交换流必须在适当的地方指出显现方向的改变。

这样做很有必要,因为字符内容体系结构中的控制功能是根据它们在字符流中的位置顺序地进行操作的。

当某一串字符以相反的显现方向夹在取正常显现方向的文本中时,具有相反显现方向的这串字符的最后一个字符图像便被定位于与具有正常显现方向的前一串字符中最后一个字符图像相邻的位置(见图 8)。

注:在图 8 中,术语“第一”和“最后”的使用与交换顺序有关,而术语“正常”和“相反”则与字符走向的方向有关。

5.2.5 并行注释

可将两个连续的字符串限定为,使得第二个串用于指示第一个串的发音和解释。打算将其用于日语,以假名字符的形式为一个或多个日本汉字字符提供发音和解释信息(Ruby)。在格式化文本中,对于字符图像取向是垂直或平行于字符走向的情况,由假名字符分别在汉字字符的上方或右方居中地进行上述表示(见图 9)。

居中放置时,便会导致假名字符被定位到定位区之外,故对假名字符串宜采取这样的定位方式:使得它们与可用区的边缘进行起始对齐或结尾对齐。

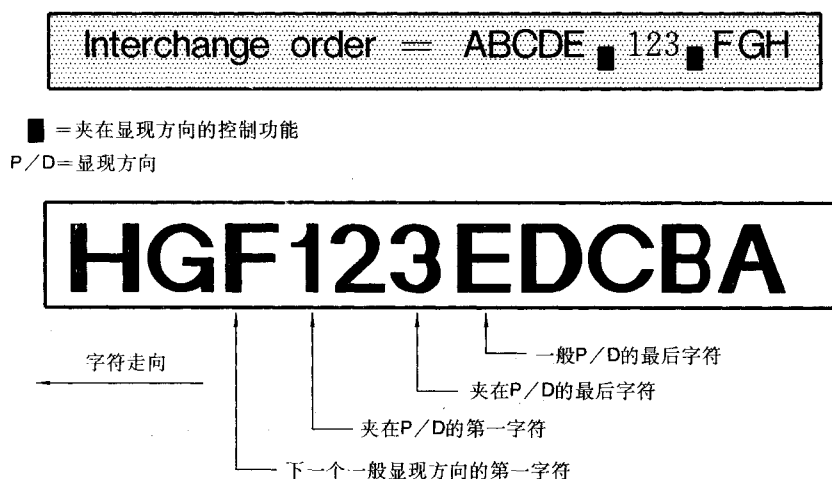


图8 字符排序

5.2.6 下标/上标

下标的呈现允许动作位置从参考线开始沿行前进方向进行移动。

上标的呈现允许动作位置从参考线开始沿行前进方向的反方向进行移动。

在一个行框内,所有的下标/上标呈现的组合结果应使动作位置在硬或软行终止符出现之前返回至参考线。

5.2.7 双字紧排

双字紧排允许由前一个字符所定义的动作位置的移动。其距离和方向则由正在成像的字符和前一个字符共同决定。

在固定间距字型的情况下,则略去双字紧排。

在可变间距字型的情况下,一个字符图像的转移点和下一个字符图像的定位点之间的距离由符合字型的字符属性所规定的紧排信息进行修改。

5.2.8 首行位移

首行位移可用于使基本成分首行的字符图像开始于一个不同于行原点位置的位置。

位移既可以与字符走向同向(产生首行缩进)也可与字符走向反向(产生伸出),如图10所示。

5.2.9 项化

项化用于使基本成分首行的项标识符成像于不直接受行原点位置和首行位移制约的位置上(见图10)。

项标识符是一个前置于基本成分首行其他内容并与其分开的字符串。

起始位移与结尾位移相对于与行原点位置未定义。这两个位移决定了项标识符的定位点及项标识符和行原点位置之间的间隔。

可以按如下方式,依项标识符对齐属性,将项标识符定位于首行:

——起始对齐:项标识符第一个字符图像的定位点置于起始位移;

——结尾对齐:项标识符最后一个字符图像的转移点置于结尾位移。

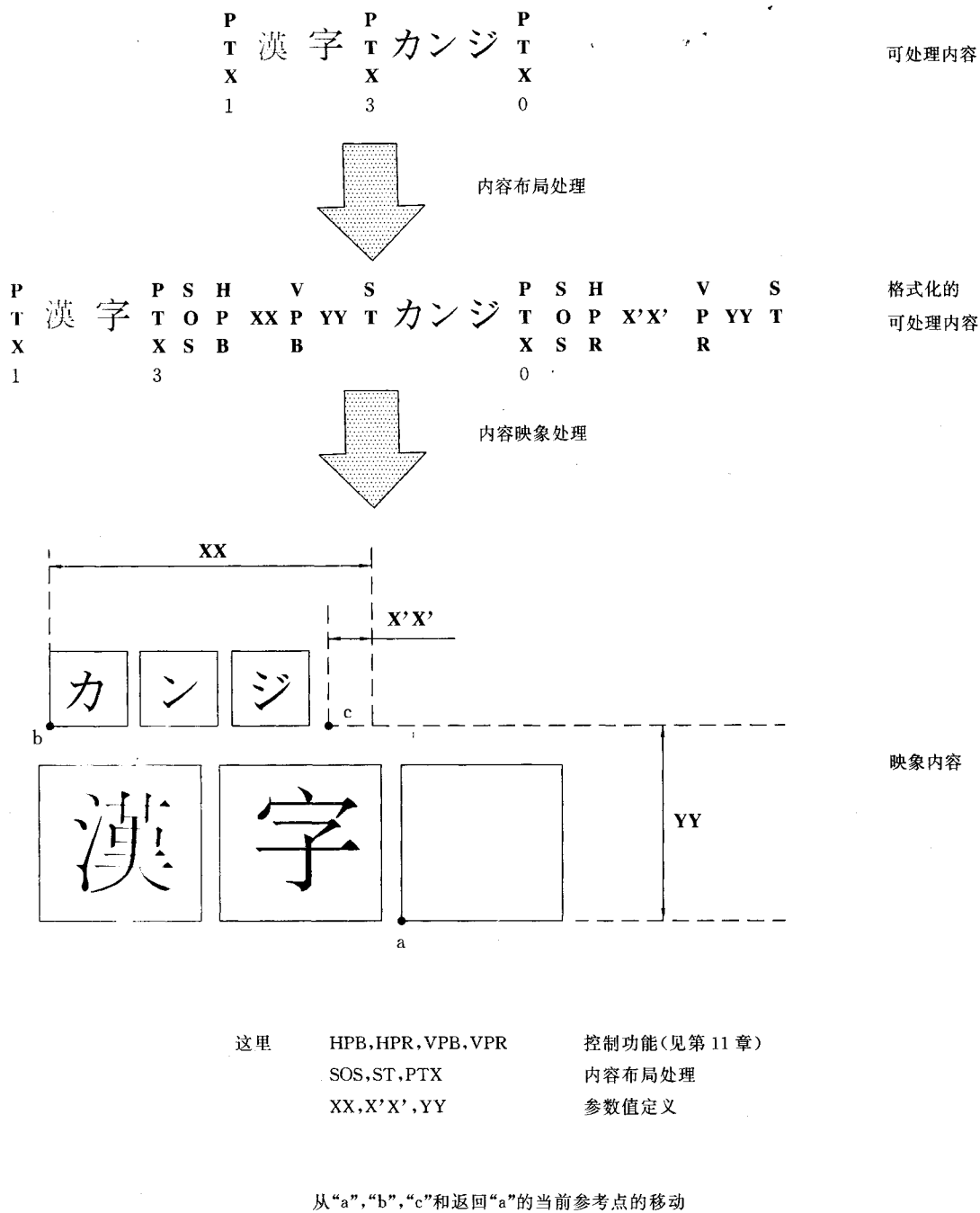


图 9 日文 Ruby 规范的一个例子