

J
15

中国国家标准汇编

195

GB 14814~14866

中国标准出版社

1 9 9 5

(京)新登字 023 号

T-652.1/2(195)

图书在版编目 (CIP) 数据

中国国家标准汇编 195: GB 14814~14866/中国标准
出版社总编室编. —北京: 中国标准出版社, 1995. 6
ISBN 7-5066-1110-4

I. 中… II. 中… III. 国家标准-中国-汇编 . T-652
. 1

中国版本图书馆 CIP 数据核字 (95) 第 05801 号

中国标准出版社出版

北京复兴门外三里河北街 16 号

邮政编码: 100045

电 话: 8522112

中国标准出版社秦皇岛印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

版权专有 不得翻印

*

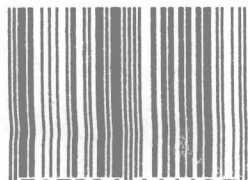
开本 880×1230 1/16 印张 49 字数 1 558 千字

1995 年 7 月第一版 1995 年 7 月第一次印刷

*

印数 1—4 000 定价 65.00 元

ISBN 7-5066-1110-4



9 787506 611107 >

出版说明

《中国国家标准汇编》是一部大型综合性工具书,自 1983 年起,以精装本、平装本两种装帧形式,分若干分册陆续出版。本汇编在一定程度上反映了我国建国以来标准化事业发展的基本情况和主要成就,是各级标准化管理机构及工矿企事业单位,农林牧副渔系统,科研、设计、教学等部门必不可少的工具书。

本汇编收入公开发行的全部现行国家标准,按国家标准号顺序编排。凡遇到顺序号短缺,除特殊注明外,均为作废标准号或空号。

本分册为第 195 分册,收入了国家标准 GB 14814~14866 的最新版本。由于标准不断修订,读者在使用和保存本汇编时,请注意及时更换修订过的标准。

中国标准出版社除出版《中国国家标准汇编》外,还出版国家标准、行业标准的单行本及各种专业标准汇编,以满足不同读者的需要。

中国标准出版社

1995 年 6 月

目 录

GB/T 14814—93 信息处理 文本和办公系统 标准通用置标语言(SGML)	(1)
GB/T 14815.1—93 信息处理 图片编码表示 第一部分:在七位或八位环境中图片表示的 编码原则	(160)
GB/T 14816—93 井用潜油三相异步电动机通用技术条件	(184)
GB/T 14817—93 永磁式直流伺服电动机通用技术条件	(203)
GB/T 14818—93 线绕盘式直流伺服电动机通用技术条件	(214)
GB/T 14819—93 电磁式直流伺服电动机通用技术条件	(230)
GB/T 14820.1—93 公路车辆用高压点火电线 第1部分 一般规定	(241)
GB/T 14820.2—93 公路车辆用高压点火电线 第2部分 铜芯高压点火电线	(249)
GB/T 14820.3—93 公路车辆用高压点火电线 第3部分 阻尼芯高压点火电线	(253)
GB 14821.1—93 建筑物的电气装置 电击防护	(257)
GB/T 14822—93 手持式电动工具的安全防护结构设计导则	(274)
GB/T 14823.1—93 电气安装用导管 特殊要求——金属导管	(307)
GB/T 14823.2—93 电气安装用导管 特殊要求——刚性绝缘材料平导管	(323)
GB/T 14823.4—93 电气安装用导管 特殊要求——可弯曲自恢复绝缘材料导管	(341)
GB/T 14824—93 发电机断路器通用技术条件	(358)
GB/T 14825—93 农药可湿性粉剂悬浮率测定方法	(376)
GB/T 14826—93 色漆涂层粉化程度的测定方法及评定	(380)
GB/T 14827—93 有机化工产品酸度、碱度的测定方法 容量法	(383)
GB/T 14828—93 农业轮胎牵引性能试验方法	(388)
GB/T 14829—93 农业机械用变速(半宽)V带	(391)
GB 14830—93 单个包装的避孕套技术要求	(394)
GB/T 14831—93 胶乳制品蒸汽老化试验方法	(398)
GB/T 14832—93 橡胶材料与液压流体的相容性试验	(400)
GB/T 14833—93 塑胶跑道	(411)
GB/T 14834—93 硫化橡胶与金属粘附性及对金属腐蚀作用的测定	(419)
GB/T 14835—93 硫化橡胶在玻璃下耐阳光曝露试验方法	(424)
GB/T 14836—93 硫化橡胶灰分的定性分析	(432)
GB/T 14837—93 橡胶及橡胶制品组分含量的测定 热重分析法	(435)
GB/T 14838—93 橡胶与橡胶制品 试验方法标准精密度的确定	(439)
GB/T 14839—93 地球化学勘查技术符号	(464)
GB/T 14840—93 石灰岩中游离二氧化硅化学分析方法	(469)
GB/T 14841—93 钽及钽合金棒材	(474)
GB/T 14842—93 铌棒材	(479)
GB/T 14843—93 铌酸锂单晶	(484)
GB/T 14844—93 半导体材料牌号表示方法	(492)

GB/T 14845—93	板式换热器用钛板	(496)
GB/T 14846—93	铝及铝合金挤压型材尺寸偏差	(500)
GB/T 14847—93	重掺杂衬底上轻掺杂硅外延层厚度的红外反射测量方法	(517)
GB/T 14848—93	地下水质量标准	(523)
GB/T 14849.1—93	工业硅化学分析方法 1,10-二氮杂菲分光光度法测定铁量	(527)
GB/T 14849.2—93	工业硅化学分析方法 铬天青-S 分光光度法测定铝量	(530)
GB/T 14849.3—93	工业硅化学分析方法 钙量的测定	(533)
GB/T 14850—93	气体分析词汇	(539)
GB/T 14851—93	电子工业用气体 磷化氢	(566)
GB/T 14852—93	气体中微量氧的测定 黄磷发光法	(577)
GB/T 14853.1—93	橡胶用造粒炭黑倾注密度的测定	(580)
GB/T 14853.2—93	橡胶用造粒炭黑细粉含量的测定	(582)
GB/T 14853.3—93	橡胶用造粒炭黑粒子磨损量的测定	(584)
GB/T 14853.4—93	橡胶用造粒炭黑堆积强度的测定	(586)
GB/T 14853.5—93	橡胶用造粒炭黑粒子尺寸分布的测定	(589)
GB/T 14853.6—93	橡胶用造粒炭黑单个粒子破碎强度的测定	(592)
GB/T 14853.7—94	橡胶用造粒炭黑技术条件	(595)
GB/T 14854.1—93	广播录像磁带测量方法	(598)
GB/T 14854.2—93	广播录像磁带性能要求	(606)
GB/T 14855—93	盒式音带制品质量总技术规范	(608)
GB/T 14856—93	海岸电台高频无线电话频道干扰的计算方法	(614)
GB/T 14857—93	演播室数字电视编码参数规范	(627)
GB/T 14858—93	黑白监视器通用技术条件	(636)
GB/T 14860—93	通信和电子设备用变压器和电感器总规范	(661)
GB 14861—93	应用电视设备安全要求及试验方法	(680)
GB/T 14862—93	半导体集成电路封装结到外壳热阻测试方法	(695)
GB/T 14863—93	用栅控和非栅控二极管的电压-电容关系测定硅外延层中净载流子浓度的 标准方法	(705)
GB/T 14864—93	实芯聚乙烯绝缘射频电缆	(715)
GB/T 14865—93	SMB 型射频同轴连接器	(743)
GB 14866—93	眼面护具通用技术条件	(769)

中华人民共和国国家标准

信息处理 文本和办公系统 标准通用置标语言(SGML)

GB/T 14814—93

Information processing—Text and office systems—
Standard generalized markup language(SGML)

本标准等效采用国际标准 ISO 8879—1986《信息处理 文本和办公系统 标准通用置标语言(SGML)》及修正 1—1988。

0 引言

本标准规定了一种用于文件表示的语言,称之为“标准通用置标语言”(SGML)。就其最广泛的定义而言,可以将 SGML 用于排版,其范围包括从传统的单一媒体数据的排版直到多媒体数据的排版。此外,也可以将 SGML 用于办公文件的处理,以满足人们阅读和在排版系统之间进行文件交换的需要。

0.1 背景

可以抽象地将文件看成是由多种类型的元素组成的一种结构。例如,作者可以将若干章组成一本书,每章中又含有段落和附有文字性注解的插图。又如,编辑人员可以将若干篇文章编成一本刊物,而每篇文章中又可以含有若干段,段中又可以含有文字,等等。

处理程序采用不同的方法来处理上述这些元素。例如,格式化程序就可用醒目的字体打印标题,在段首或各段之间留空,这就从视觉上将文件的结构和属性展示给了读者。而当为信息检索系统建立标题字典时,则可能对标题中的文字赋予特别的含义。

虽然在文件的属性与其处理之间的这种联系现在看来是清晰的,然而就早期的文本处理方法而言,这种联系是模糊的。在自动化排版出现之前,编辑人员先使用特殊的处理指令对底稿进行“置标”,然后再由排版工人按照这些指令排出所要的格式。指令和文件结构之间的所有联系完全存在于编辑人员的脑中。

早期计算机排版继承了上述方法,用与处理相关的“置标”加入计算机可读的文件卷中。虽然加入的置标依然由特定的处理指令组成,但这些处理指令采用的是格式化程序的语言,而不是排版工人使用的语言。然而,这种文卷如果不改变其中的全部置标,就很难再用于其他不同的目的或不同的计算机系统中。

随着用户水平日益提高及文本处理程序功能日益增强,人们开发出许多解决上述问题的方法。例如,使用“宏调用”(或“格式调用”)将文件中需要处理的地方标识出来,而将实际的处理指令放在文件外部的“过程”之中(或“宏定义”或“存储格式”之中),这样就能够比较容易地对它们进行修改。

虽然可以将宏调用放在文件中的任意位置上,但是用户开始逐步认识到大部分宏调用都是放在文件元素的开头或结尾处。于是,很自然地选择表示元素类型的“通用标识符”来命名这些宏调用,而不是用特定处理的名字命名(例如,用“heading”代替“format—17”,即用“标题”代替“格式 17”),这是实行“通用编码”(或“通用化的标记”)的开始。

通用编码使自动文本处理系统迈出重要的一步,它反映了文件属性与处理之间的自然关系。七十年代初,“通用置标语言”的出现为通用编码技术提供了一种形式化语言的基础,从而进一步推动了这种发

国家技术监督局 1993-12-24 批准

1994-08-01 实施

展趋势。通用置标语言主要遵循以下两条原则：

- a. 描述性置标起主要作用，并与处理指令相区别。

描述性置标包括通用标识符和文件元素的其他属性，它们可调用处理指令。处理指令可用任何语言表示，并且常被汇集在文件外部的过程中。

当为了发现置标而扫描源文卷并识别出各种元素后，处理系统执行与每个元素和属性相关的过程。若用其他处理系统，同样的元素及属性可与不同的过程相关联，而不必改变文件的置标。

当需要直接将处理指令放入文件中时，它的定界不同于描述性置标，这样可以容易地使不同的处理系统找到并修改。

- b. 对每种文件类型都形式化地定义其置标。

通用置标语言用加入“文件类型定义”的做法使文件置标形式化。类型定义包括哪些元素和属性可以在文件中出现并以何种顺序出现等方面的说明（类似于形式文法）。根据这些信息可以确定对某一文件的置标是否正确（即是否与类型定义相符合），还能提供被漏掉的置标，因为可从其他置标中准确地作出推断。

注：对通用编码和标准通用置标语言概念的更加详细的介绍参见附录 A（参考件）。

0.2 目标

标准通用置标语言对通用编码和通用置标概念的应用进行标准化。它为描述用户选定的内容提供了一种清晰的且无歧义的语法。该语言包括：

- 用于文件元素的描述性置标的“抽象语法”；
- 将抽象语法与特定的定界符字符和数量相结合的“基准具体语法”。用户可定义变异的具体语法以满足自己的需要；
- 置标声明，允许用户针对不同的文件类型定义通用标识符和属性的特定词汇；
- 对任意数据内容的规定。在通用置标中，“数据”是指未用置标语言定义的内容，包括特定的并且要求与一般文本作不同解释的“数据内容记法”，如：公式、图像、非拉丁字母、前面已提到的格式化文本或图形等；
- 实体引用：为了引用文件的主数据流之外的内容（如：分开编写的章节、临时补入的字符、照片等）而采取的一种非系统特定的技术；
- 将处理指令与描述性置标相互区别用的专用定界符。对于那些不能由过程来处理的情况，需要时可以插入处理指令，但当文件送到另一个处理系统时，又能很容易地找到并修改这些处理指令。

然而，为使通用置标语言成为一种可以接受的标准，还需要更多的功能。为了满足在复杂环境中使用该语言所需的各种约束条件，该语言还必须具有元语言的特征。现将主要约束条件，以及标准通用置标语言用以满足这些约束条件的方法总结如下：

- a. 该语言“置标”的文件必须能被广泛使用的文本处理系统和文字处理系统所接受。

具有全部可选特征的完整语言为高级系统提供了可以利用的通用性和灵活性；而功能较弱的系统可以不支持这些特征。为了便于在不同的系统之间进行文件交换，可用“SGML 声明”描述文件中使用的置标特征或具体语法的变异。

- b. 必须支持现有的大量正文输入设备。

用基准具体语法的 SGML 文件，没有机器帮助也能易于被人们键入和理解。因此：

- SGML 的使用并不需要等到开发出新一代硬件，只需要能在现有机器上处理文件的软件；
- 随着用户对 SGML 越来越熟悉，将 SGML 移植到新一代硬件（当它们出现时）上也将更加容易。

- c. 由于可在不同的设备上键入文件，因而该语言必须不依赖于任何字符集。

该语言不依赖于某一特定字符集。如果一个字符集中含有以位组形式表示的字母、数字、空格以及

定界符等字符,那么该字符集就是可以接受的。

d. 必须具有相对于处理程序、系统或设备的独立性。

由于通用置标主要是描述性的,因而它在本质上就具有这种独立性。又由于偶然出现的处理指令都被特别加以定界,所以当不同的处理程序要放弃与其不相关的指令或要进行文件交换时,就可以查找出这些指令,并对其进行转换。

对文件外部部分的引用是间接的。对实际系统存储的映射是由出现在文件开头部分中的“外部实体声明”实现的,故在文件交换时,能很容易对它们修改。

可以通过 SGML 声明来改变具体语法,以适应任何保留的系统字符。

e. 必须没有任何民族语言偏向。

可以用任何特定民族语言中的字符来扩充用于命名的字符。描述性置标中使用的通用标识符、属性名、以及其他名字都由用户在元素和实体声明中定义。

置标声明中使用的声明的名字和关键字也都可以改变。

就象在多语种文件中使用的那样,也支持多种字符表。

f. 该语言必须适应人们熟悉的有关打字机和文字处理程序方面的习惯。

用“便捷引用”和“数据标记”性能支持打字机文本输入方面的习惯。对于含有段和引文的常规文本可按照 SGML 解释,尽管没有键入看得见的置标。

g. 该语言必须不依赖于任何特定的数据流或物理文卷的组织。

置标语言采用的是虚拟存储模型,其中文件由一个或多个存储实体组成,而每个存储实体又是一个字符序列。对于所有实际文卷的访问都由处理系统完成。处理系统确定是否将字符序列看成是连续的,或确定字符序列是否表示物理记录的边界。

h. “已置标”的正文必须与其他数据共存。

只要能够确定出文本的开头和结尾,那么处理系统就能够使符合本标准的文本出现在带有其他内容的数据流之中。

类似地,系统可以允许未被 SGML 定义的数据内容出现在符合本标准的文件之中。为了便于交换,这类数据的出现由置标声明指明。

i. 置标对人和程序均是可用的。

标准通用置标语言的目的是不需要预处理程序就能成为键盘输入和交换的合适接口。为了适应用户在文本输入方面的习惯及经验,以及各种类型的键盘和显示设备的要求,允许对语言进行广泛的剪裁。

然而,人们认识到许多实现者都将试图利用通用置标语言在信息获取能力上的长处,以提供智能编辑或从文字处理前端环境中建立 SGML 文件的功能。通过提供下述功能将使 SGML 能够满足上述这些应用:

- 元素内容可以和置标分开存储;
- 控制字符可以作为定界符使用;
- 文件中允许出现数据表示的混合模式;
- 支持多重并发逻辑结构和版面结构。

0.3 组织

本标准的内容组织如下:

- a. 实体结构方式的 SGML 文件的物理组织,在第 6 章说明;
- b. 元素结构方式的 SGML 文件的逻辑组织和采用描述性置标的表示法,在第 7 章说明;
- c. 处理指令在第 8 章讨论;
- d. 公用置标构件,如字符、实体引用、以及处理指令在第 9 章说明;
- e. 具有通用性的置标说明(注释、实体、以及特殊标出节)在第 10 章说明;

f. 主要用于说明文件类型定义(文件类型、元素、记法、便捷引用映射、以及便捷引用使用)的置标声明在第 11 章定义;

g. 主要用于说明链接处理定义(链接类型、链接属性、链接集、以及链接集使用)的置标声明在第 12 章定义;

h. 用于说明文件字符集、容量集、具体语法、以及各种特征的 SGML 声明在第 13 章定义;

i. 基准具体语法在第 14 章定义;

j. 文件、应用程序的一致性在第 15 章定义。

最后还有一批附录,作为本标准的参考件。

注:本标准是一种计算机语言的形式化说明,对于其专长是产生文件而不是编译方面的人员来说,阅读本标准可能会有一定的困难。附录 A、B、C 以一种非形式化的教学方式讨论了其中的主要概念,这种方式对大部分读者来说应更容易接受。然而,读者也应知道,这些附录既没有概括 SGML 的所有结构,也没有概括这些结构的所有细节,而且为了能够清晰地表示问题,某些细微的区别常常被忽略。

1 主题内容

本标准:

a. 规定了一种称为标准通用置标语言(SGML)的抽象语法。该语言既描述了如何表示文件的结构和其他属性,也给出了对置标进行解释的其他信息;

b. 规定了将抽象语法与特定的字符和数量相结合的基准具体语法,给出了定义变异具体语法的准则;

c. 根据文件使用的语言成分定义了一致性文件;

d. 根据系统处理一致性文件以及识别一致性文件中的置标错误的能力,定义了一致性系统;

e. 规定了怎样才能将未被本标准定义的数据(如图像、图形和格式化的文本)包含在一个一致性文件中。

注:本标准:

a. 没有标出或规定“标准的”文件类型、文件结构或文本结构;

b. 没有规定一致性系统的实现、体系结构或置标错误的处理;

c. 没有规定怎样创建一致性文件;

d. 没有规定数据流、消息管理系统、文卷结构、存储或交换一致性文件的物理表示,或为此目的使一致性文件可相互转换的字符集或编码方案;

e. 没有规定数据内容的表示或包含在一致性文件中的图像、图形、格式化的文本等的表示法。

2 适用范围

标准通用置标语言可用于被任何文本处理或文字处理系统加工的文件。它特别适用于:

a. 在使用不同文本处理语言的系统之间交换的文件;

b. 即使使用同一种文本处理语言,也可用不同方式处理的文件。

只以最终图像形式存在的文件不在本标准适用范围内。

3 引用标准

GB 1988 信息处理 信息交换用七位编码字符集

GB 4880 语种名称代码

GB 13000.1 信息技术 通用多八位编码字符集(UCS) 第一部分:体系结构与基本多文种平面

ISO 9069 信息处理 SGML 支持工具 SGML 文件交换格式(SDIF)

ISO 9070 信息处理 SGML 支持工具 公用文本登记过程

下面的引用标准是与示例材料一起使用的:

- GB 2311 信息处理 七位和八位编码字符集代码扩充技术
- GB 2659 世界各国和地区名称代码
- GB 8565.1~8565.3 信息处理 文本通信用编码字符集
- GB 11383 信息处理 信息交换用八位代码结构和编码规则
- ISO 8632/2 信息处理系统 计算机图形 用于图像描述信息转换与存储的元文卷 第二部分: 字符编码
- ISO 8632/4 信息处理系统 计算机图形 用于图像描述信息转换与存储的元文卷 第四部分: 清晰的文本编码
- ISO 8825 信息技术——开放系统互连——抽象语法表示法 1(ASN.1)基本编码规则规范

4 定义

下列定义适用于本标准。

- 4.1 (SGML)抽象语法 abstract syntax(of SGML)
定义如何将置标加入文件数据中的一组规则,与表示置标的特定字符无关。
- 4.2 活动文件类型(声明) active document type(declaration)
已被系统标识为活动的文件类型。
注:如果一个 SGML 实体有对应的活动文件类型,则依照其活动文件类型对该实体进行语法分析,否则,就按其基文件类型以及任意活动链接类型对该实体进行语法分析。
- 4.3 活动链接类型(声明) active link type(declaration)
已被系统标识为活动的链接处理。
- 4.4 歧义性内容模型 ambiguous content model
文件实例中的元素或字符串能满足一个以上的初始内容记号的一种内容模型。
注:SGML 中禁止使用歧义性内容模型。
- 4.5 应用 application
文本处理应用程序。
- 4.6 应用约定 application convention
在 SGML 允许用户选择的范围内,管理文件正文的特定应用规则。
注:有两类应用约定:内容约定和置标约定。
- 4.7 特定应用信息 application-specific information
SGML 声明的参数,它指明应用和(或)其结构所需要的信息。
注:例如,该信息能够识别一个结构和(或)一个应用,或使一个系统能够确定它能否处理该文件。
- 4.8 相关元素类型 associated element type
用相关元素类型参数与置标声明对象相关联的元素类型。
- 4.8.1 相关记法(名) associated notation(name)
用相关记法名参数与置标声明对象相关联的记法名。
- 4.9 (元素的)属性 attribute(of an element)
既非类型也非内容的特征量。
- 4.10 属性定义 attribute definition
属性定义表的成员;定义属性名、允许值、及缺省值。
- 4.11 属性定义表 attribute definition list
由属性定义表声明中的属性定义表参数定义的一个或多个属性定义的集合。
- 4.12 属性(定义)表声明 attribute(definition)list declaration
将属性定义表与一个或多个元素类型相关联的置标声明。

4.13 属性表 attribute list

属性说明表。

4.14 属性表声明 attribute list declaration

属性定义表声明。

4.15 属性说明 attribute specification

属性说明表的成员,它说明单个属性的值。

4.16 属性(说明)表 attribute(specification)list

一个或多个属性说明的集合构成的置标。

注:属性说明表出现在开始标记和链接集合中。

4.17 属性值的字面值 attribute value literal

通过替换引用以及忽略或转换功能字符而被解释为属性值的已定界的字符串。

4.18 可用的公用文本 available public text

可为一般公众使用的公用文本,它的拥有者可以要求为此付款或遵守其他条件。

4.19 B 序列 B sequence

一个连续的大写字母“B”的序列;在一个作为便捷引用的串中,它表示最小长度与 B 序列的长度相同的空白序列。

4.20 基文件元素 base document element

其文件类型是基文件类型的文件元素。

4.21 基文件类型 base document type

在前导说明中,第一个文件类型声明所说明的文件类型。

4.22 基本的 SGML 文件 basic SGML document

符合 SGML 的文件,它使用基准具体语法和容量集,以及置标最简化特征 SHORTTAG 和 OMIT-TAG。

注:它还可以根据基准具体语法使用 SHORTREF 特征。

4.23 位 bit

二进制数字,即 0 或 1。

4.24 位组 bit combination

位的有序集合,可解释为二进制数。

4.25 空白序列 blank sequence

SPACE 和(或)分隔字符(SEPCHAR)的一个连续序列。

4.26 容量 capacity

有关文件的大小或复杂性等方面的命名的限制,用若干种单位表示,用以累计一类对象或所有对象。

注:容量集由抽象语法定义,但这些值是由单个文件和 SGML 系统赋给它们的。

4.27 容量集 capacity set

将数值赋给容量名的赋值集合。

注:在 SGML 声明中,容量集指出文件的最大容量要求(实际要求可能稍低)。容量集也可由应用程序定义,以便限制该应用程序实现时必须处理的文件容量要求。容量集还可由系统定义,以指明系统所能满足的容量要求。

4.28 CDATA

字符数据。

4.29 CDATA 实体 CDATA entity

字符数据实体。

4.30 (链接)处理的链 chain of(link)processes

相继执行并形成一条链的若干处理,在链中,第一个处理的源是一个基文件类型实例,并且除最后一个处理外,每个处理的结果都是下一处理的源。该链的任何部分都可以重复。

注:例如,一个复杂的分页置标应用程序系统可以包含三种文件类型:逻辑文件、长校样文件和分页文件,同时还包含两个链接处理:“调整版面”及“估计页数”。“调整版面”处理从一个逻辑文件实例创建一个长校样文件实例,而“估计页数”处理将长校样文件实例按顺序创建分页文件实例。由于“估计页数”处理中作出的决定可能要求进一步对长校样文件的大小进行调整,所以这两个处理可以重复。

4.31 字符 character

由字符表定义的具有单独意义的信息最小单位。

注:① 有两类字符:图形字符和控制字符。

② 一个字符在上下文中的意义是由置标或数据内容记法定义的,这样就取消或补充了该字符在字符表中的意义。

4.32 (字符)类 (character)class

在抽象语法中具有共同作用的字符集,如非 SGML 字符或分隔符字符。

注:有四种不同方法将一些特定的字符指定为字符类:

- a. 由抽象语法明确指定(如 Special, Digit, LC Letter 和 UC Letter);
- b. 由具体语法明确指定(如 LCNMSTRT FUNCHAR SEPCHAR 等);
- c. 作为明确指定定界角色或其他字符类的结果而隐含地指定(如 DELMCHAR 和 DATACHAR);
- d. 由文件字符集明确指定(如 NONSGML)。

4.33 字符数据 character data

零个或多个出现在一段不识别置标的上下文中的字符,但不包括结束该字符数据的定界符。这样的字符被归类为数据字符,因为它们就是被如此声明的。

4.34 字符数据实体 character data entity

当引用时其正文被看作字符数据的实体,并且其正文不依赖于特定的系统、设备或应用处理程序。

4.35 字符实体集 character entity set

由通用实体组成的公用实体集,它们本身是图形字符。

注:① 字符实体适用于在文件字符集中无编码表示的字符,或不能方便地从键盘上键入的字符,或者为了与设备无关,其位组不能在所有输出设备上产生正确显示的字符。

② 有两类字符实体集:定义的和显示的。

4.36 字符编号 character number

一个与字符编码表示等值的十进制整数,其值是将位组序列看作单个二进制整数而得到的。

4.37 字符引用 character reference

用单个字符作替换的引用。

注:有两类字符引用:命名字符引用和数字字符引用。

4.38 字符表 character repertoire

同时使用的字符集合。定义了每个字符的意义,还能定义多字符的控制序列。

注:当一个字符出现在控制序列中时,则该序列的含义替代了单个字符的含义。

4.39 字符集 character set

从字符表到代码集的映射,这就使每个字符都与其编码表示对应。

4.40 (字符)串 (character)string

字符的序列。

4.41 类 class

字符类。

4.42 代码扩充 code extension

使用单个代码表示多个字符,而不改变文件字符集。

注：当文件中出现多种民族语言时，图形表代码的扩充可能是有用的。

4.43 代码集 code set

长度相同的位组的集合，按其数值排序，这些数值必须是连续的。

注：例如，一个其位组为八位的代码集（一个“八位码”）可以构成 256 个位组，其值的范围从 00000000 到 11111111（十进制为 0 到 255），该代码集也可由这些位组的任何连续子集构成。

4.44 代码集位置 code set position

代码集中位组的数值。

4.45 代码表示 coded representation

一个字符用一个或多个长度相同的位组的序列表示。

4.46 注释 comment

置标声明的一部分，它含有辅助用户使用该文件的解释或评注信息。

4.47 注释声明 comment declaration

只含有注释的置标声明。

4.48 (SGML 的)具体语法 concrete syntax(of SGML)

抽象语法与特定的定界符、数量、置标声明名字等的结合。

4.49 具体语法参数 concrete syntax parameter

SGML 声明的参数，它标识文件元素以及（通常情况下）前导说明部分中使用的具体语法。

注：该参数由标识基准语法字符集、功能字符、回避字符、命名规则、定界符使用、保留名使用，以及数量特征等参数组成。

4.50 符合于 SGML 的应用 conforming SGML application

要求文件是符合 SGML 的文件，且编制的文档满足本标准要求的 SGML 应用。

4.51 符合于 SGML 的文件 conforming SGML document

符合本标准全部规定的 SGML 文件。

4.52 包含元素 containing element

在其中出现子元素的元素。

4.53 内容 content

文件实例中出现在开始标记与结束标记之间的字符，它们可解释为数据、真子元素、包含的子元素、其他置标，或上述内容的混合。

注：如果元素有显式的内容引用，或其声明的内容为“EMPTY”，则其内容都为空。在这种情况下，数据可能由应用本身产生，并与内容数据一样处理。

4.54 内容约定 content convention

控制数据内容的应用程序约定，如长度限制、允许的字符或大小写字母的使用。

注：内容约定本质上是非正式的数据内容记法，通常限于单个元素类型。

4.55 (内容)模型 (content)model

元素声明中的参数，它说明了模型组和例外参数，这两者定义了元素中允许的内容。

4.56 内容模型嵌套层次 content model nesting level

在没有对应 grpc 或 dtgc 定界符的内容模型中连续出现的 grpo 或 dtgo 定界符的最大数目。

4.57 内容引用(属性) content reference(attribute)

一个可隐含的属性，其值由应用程序引用以产生内容数据。

注：当元素有显式的内容引用时，则文件实例中该元素的内容为空。

4.58 上下文序列 contextual sequence

一个或多个置标字符构成的序列，它必须跟在同一个实体中的定界符串之后，从而使该串被识别为定界符。

4.59 上下文中的可选元素 contextually optional element

即这样的元素：

- a. 只因它是包含元素才能够出现；或
- b. 它在当前适用的模型组中的内容记号是上下文中的可选记号。

4.60 上下文中的可选记号 contextually optional token

即这样的内容记号：

- a. 它是一个固有的可选记号；或
- b. 它有一个 plus 出现指示符并且已被满足；或
- c. 它在一个模型组中，该模型组本身是一个上下文中的可选记号，且没有已满足的记号。

4.61 上下文中必要的元素 contextually required element

一个并非上下文中可选的元素，且：

- a. 它的通用标识符是文件类型名；或
- b. 它的当前适用模型记号是上下文中必要的记号。

注：一个元素可以既不是上下文中必要的，同时也不是上下文中可选的。例如：一个元素其当前适用的模型记号出现在一个没有固有的任选记号的 or 组中。

4.62 上下文中必要的记号 contextually required token

即这样的内容记号：

- a. 它是其模型组中仅有的一个内容记号；或
- b. 它出现在一个 seq 组中，并且：
 - i)——该 seq 组本身就是一个上下文中必要的记号；或
——含有一个已经满足的记号；并且
 - ii)它前面的所有记号
——已经满足；或
——是上下文中可选的。

4.63 控制字符 control character

对其后的字符的解释、表示或其他处理进行控制的字符；例如，制表字符。

4.64 控制序列 control sequence

以控制字符为开头的字符序列，它控制其后的字符的解释、表示或其他处理；例如，转义序列。

4.65 核心具体语法 core concrete syntax

基准具体语法的一种变异，它没有便捷引用定界符。

4.66 (内容记号的)对应的内容 corresponding content(of a content token)

文件实例中对应内容记号的元素和(或)数据。

4.67 当前属性 current attribute

其当前(即最近指定的)值成为其缺省值的属性。

注：当带有当前属性的元素第一次出现时，不能省略开始标记。

4.68 当前元素 current element

其开始标记(或由于置标简化而省略)最近出现的开元素。

4.69 当前链接集 current link set

由元素内容中的链接集使用声明，或链接处理定义而与当前元素相对应的链接集。如果该当前元素没有相关联的链接集，则以前的当前链接集仍作为当前链接集。

4.70 当前映射 current map

由元素内容中的便捷引用使用声明或者文件类型定义而与当前元素相对应的便捷引用映射。如果当前元素没有相对应的映射，则以前的当前映射将继续作为当前映射。

4.71 当前层 current rank

用以得到通用标识符时,附加在标记中层词干之后的一个数字。对于开始标记,它是最近元素的层后缀,该元素带有相同层词干或相同分层组中的层词干。对于结束标记,它是带有相同层词干的最近开元素的层后缀。

4.72 数据 data

文件中表示其固有信息内容的字符;这些字符不被识别为置标。

4.72.1 数据属性 data attribute

符合特定数据内容记法的数据的属性。

注:在大多数情况下,数据属性的值必须在数据可据记法进行解释之前知道。

4.73 数据字符 data character

或者因为被说明为数据,或者因为没有被识别为置标而在其所出现的上下文中被解释为数据的 SGML 字符。

4.74 数据内容 data content

元素内容的一部分,它是数据而不是置标或子元素。

4.75 数据内容记法 data content notation

对元素的数据内容或数据实体的特定应用的解释,它通常补充或不同于文件字符集的常规含义。

注:数据内容记法对于元素内容是由记法属性说明的,对于数据实体,则由实体声明中的记法名参数说明。

4.75.1 数据实体 data entity

被声明为数据的实体,因而引用时不作语法分析。

注:① 有三类数据实体:字符数据实体、特定的字符数据实体以及非 SGML 数据实体。

② 一个数据实体的解释,由数据内容记法决定,此记法由其他标准定义。

4.76 数据标记 data tag

与一个开元素的数据标记模式相匹配的字符串,它既是该开元素的结束标记,同时又是包含它的元素中的字符数据。

4.77 数据标记组 data tag group

联系数据标记模式与目标元素类型的模型组记号。

注:在目标元素实例中,对数据内容及子元素的数据内容扫描,以找出与该模式(数据标记)相匹配的字符串。

4.78 数据标记模式 data tag pattern

定义了这样一些串的数据标记组记号,即如果这些串出现在适当的上下文中将构成一个数据标记。

4.79 声明 declaration

置标声明。

4.80 声明子集 declaration subset

置标声明的已定界部分,在其中可出现其他声明。

注:声明子集只能出现在文件类型、链接类型以及特殊标出节声明中。

4.81 已声明的具体语法 declared concrete syntax

由 SGML 声明中的具体语法参数描述的具体语法。

4.82 专用数据字符 dedicated data characters

由每个不可能具有置标含义的 SGML 字符构成的字符类;即它的成员只能被作为数据字符处理。

4.83 缺省实体 default entity

通用实体引用所引用的实体,此通用实体引用使用了未声明的名字。

4.84 缺省值 default value

属性定义的一部分,没有属性说明时使用的属性值。

4.85 定义性的(字符)实体集 definitional(character)entity set

字符实体集,其目的是为图形字符定义实体名,但并不实际显示它们。其公用标识符不包含公用文

本显示版本。

注：在处理过程中，系统为适合输出设备，用对应的显示字符实体集替代定义性实体集。

4.86 定界符字符 delimiter characters

除名字字符或功能字符外的其他 SGML 字符构成的字符类，它出现在由具体语法作为定界符使用的串中。

4.87 上下文中的定界符 delimiter-in-context

在同一实体中，一个定界符串之后紧跟一个上下文序列而构成的字符串。

4.88 定界符角色 delimiter role

由抽象语法定义，并由具体语法指定的字符串所担任的角色，它用于标出置标和(或)从数据中区分置标。

4.89 定界符集 delimiter set

对应抽象语法定界符角色的一组定界符串。

4.90 定界符集参数 delimiter set parameter

SGML 声明中的参数，它标识已声明的具体语法中所使用的定界符集。

4.91 定界符(串) delimiter(string)

由具体语法赋予定界符角色的字符串。

4.92 描述性置标 descriptive markup

按照非系统特定方式对文件结构及其他属性进行描述的置标，与对它进行的任何处理无关。特别是，它用标记表达元素的结构。

4.93 (公用文本的)依赖于设备的版本 device-dependent version(of public text)

是公用文本，它只通过增加公用文本显示版本而使它的形式公用标识符与其他公用文本不同，该版本标识出所支持的显示设备或所用的编码方案。

4.94 数字 digits

由“0”到“9”共 10 个阿拉伯数字组成的字符类。

4.95 显示(字符)实体集 display(character)entity set

与对应的定义性字符实体集具有相同实体名的实体集，但它用于字符的显示。它是对应定义性实体集的依赖设备的版本。

4.96 文件 document

作为一个单位处理的信息集。文件将按特定文件类型分类。

注：在本标准中，该术语总是指 SGML 文件(这并不失其准确性)。

4.97 文件的体系结构 document architecture

文本处理应用程序的形式化规则。

注：例如，文件的体系结构可定义为：

- a. 在各类元素定义中使用的属性语义；
- b. 基于元素所具有的属性的元素分类；
- c. 根据元素分类而定义文件类型的结构规则；
- d. 链接处理，以及它们如何被属性值影响；与(或)
- e. 在文件交换中，与文件相伴的信息(“文件外形”)。

4.98 文件字符集 document character set

用于 SGML 文件中的所有置标，以及(至少)最初用于数据的字符集。

注：当在系统之间交换文件时，其字符集将转换成接受系统的字符集。

4.99 文件元素 document element

文件类型实例的最外层元素，亦即，该元素的通用标识符为文件类型号。

4.100 文件实例 document instance