



中华人民共和国国家标准

GB/T 16684—1996
idt ISO/IEC 8211:1994

信息技术 信息交换用数据描述 文卷规范

Information technology—Specification for a data
descriptive file for information interchange

1996-12-18 发布

1997-07-01 实施

国家技术监督局 发布

中华人民共和国
国家标准
信息技术 信息交换用数据描述
文卷规范
GB/T 16684—1996

*

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:68522112

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 4½ 字数 136 千字

1997年10月第一版 1997年10月第一次印刷

印数 1—600

*

书号: 155066·1-14109 定价 27.00 元

*

标目 319—58

前 言

本标准等同采用国际标准 ISO/IEC 8211:1994《信息技术 信息交换用数据描述文卷规范》。

本标准与 ISO/IEC 8211:1994 版在引用标准中有一些小的差异,除此之外,本标准无论在技术内容上,还是在编排方式上均与国际标准保持一致。

本标准规定了在计算机系统之间进行信息交换用的且与媒体和系统无关的文卷和数据的记录格式。

本标准的附录 A 是标准的附录,附录 B 到附录 H 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。

本标准主要起草人:冯惠、李宁、王宝艾、金益民、徐祖渊、段祥。

ISO/IEC 前言

ISO(国际标准化组织)和 IEC(国际电工委员会)是世界性的标准化专门机构。国家成员体(它们都是 ISO 或 IEC 的成员国)通过国际组织建立的各个技术委员会参与制定针对特定技术范围内的国际标准。ISO 和 IEC 的各技术委员会在共同感兴趣的领域内进行合作。与 ISO 和 IEC 有联系的其他官方和非官方国际组织也可参与国际标准的制定工作。

对于信息技术,ISO 和 IEC 建立了一个联合技术委员会,即 ISO/IEC JTC1。由联合技术委员会提出的国际标准草案需分发给国家成员体进行表决。发布一项国际标准,至少需要 75% 的参与表决的国家成员体投票赞成。

国际标准 ISO/IEC 8211 是由 ISO/IEC JTC1“信息技术”联合技术委员会,SC21“开放系统互连、数据管理和开放分步式处理”分技术委员会制定的。

此第二版将取消并代替已作了技术性修改的第一版(ISO 8211:1985)。

ISO/IEC 8211 的这一版的实质性变化如下:

1. 用于数字值的二进制形式。
2. 二进制头标和目录。
3. 支持 ISO/IEC 10646。
4. FTAM 非结构化和结构化文件类型的定义。
5. 并置正则结构。
6. 递归树描述。
7. 人工可读的,数据字段描述的可替换形式。

本国际标准第二版与第一版向下兼容。

附录 A 构成了本国际标准的组成部分。附录 B 到附录 H 仅提供参考信息。

引 言

本标准是根据使数据结构易于从一个计算机系统传送到另一个系统所提出的机制的标识需要而制定的,这种机制与体系结构无关。需要交换的数据结构在复杂程度和大小上能有效地变化,并且需要完成这些交换的一般方法。还需要媒体(例如通信线路、磁带、磁盘组、软磁盘等)能用于物理交换,以及在目标系统中能成功地重新产生数据结构所必需的全部信息应包含在媒体上传输的信息中。

为了满足上述需要,本标准规定在计算机系统之间进行信息交换用的且与媒体和系统无关的文卷和数据的记录格式。本标准期望与物理记录媒体以及通信媒体一起使用。用户数据结构的内容由国际上可识别的字符集和代码组成,并以透明方式交换。传递信息的中间结构是为交换目的设计的,但也用于某种形式的一般处理,并适于在大容量的直接访问交换文卷媒体上使用直接访问方法。

本标准是具体语法和编码的标准,提供一种描述包含用户数据的文卷的工具,但并不规定用户数据字段或用户数据记录的内容或次序。本标准为这类记录和字段规定了一种综合的类属形式,以适应规范的用户需求,既适用于简单用户数据,也适用于复杂用户数据。应用必须设计他自己的符合标准的一个交换文卷实例,符合标准的文卷,数据和数据描述应能由同一软件处理,直到字段或子字段一级。当然,用户必须完成与自身应用系统的接口。

所使用的方法是定义一种交换格式,大部分信息结构及其内容可转换为该格式而不丢失信息,并且从该格式可恢复原始的结构和内容。该交换格式既适用于在物理媒体上记录,也适用于经通信系统传送。

该交换格式支持的数据结构有:基本数据、向量、数组和层次。可转换为该交换格式的文卷结构包括顺序、层次和关系。不直接支持网状结构,对网状结构为了保存逻辑链接,对网状结构需要辅助的预处理和后处理。

本标准与媒体无关,本标准假定支持标准的传输系统至少应能处理定长八位位组字符串。它需要能把用户文卷或数据库数据映射成交换文卷的计算机处理能力。此种映射功能必须提供必要的数据和结构转换。将选择和转换这些数据项和结构定义为本标准规定的格式所需的参数超出本标准的范围。本标准要求在控制字段中使用基于 GB/T 15237.1 和 GB 5261 的基本字符集,并允许在用户数据字段中使用扩充字符集。本标准提供三种交换级,用户可根据其数据结构的复杂程度从中选择。1 级交换级支持包含简单的、非结构字符串的多个字段。2 级交换级支持 1 级并处理包含由各种数据类型组成的结构用户数据的多个字段。3 级交换级支持 2 级和层次数据结构。

本版本中引入的改动增强了本标准的通用性,提供更有效的交换能力。还有一些改动是为了改善清晰性和用户可接受性。在支持本标准的标准化交换中的技术改动以及对于本标准的组织责任的改动导致了其他方面的扩展。本版标准为用户提供了一种改进了的交换工具,以适应用户不断增长的需求,并且集成在 OSI 环境中。

检索档案文卷时可能用到与原始档案系统不同的计算机系统。此种操作问题与在不同地点的计算机间传送文卷时所涉及的问题是相同的。本标准为这类应用提供了机制。

本标准基于 GB/T 2901,与 GB/T 2901 具有相同的记录结构,但数据描述成分不同。用于文卷传送和随机文卷访问的基于 GB/T 2901 的系统至少自 1970 年就已开始使用了,目前已被广泛应用。本标准所用的术语与其前驱 GB/T 2901 一致。其层次逻辑构造是文卷、可变量记录、可变量字段和子字段。在若干种编程语言中,与本标准的字段等价的称为记录,特定的应用要将语言中的记录作为本标准中的字段传送,相关的语言记录聚集为本标准的中记录。

GB/T 16684—1996

本标准的组织

本标准的内容组织如下：

1) 第5章描述所有逻辑记录公用的,且输入逻辑记录和完整字段时必须的头标、目录和字段区内内容的规范。

2) 第6章描述在子字段级输入用户数据时必须的数据描述规范。6.1进一步规定了包含与数据描述相关信息的DDR头标中的子字段。

3) 第7章描述扩展字符集的用法。

4) 附录A描述FTAM登记。

5) 附录B描述本标准文卷设计和数据描述的方法论。

6) 附录C到附录H对上述方法论提供指导信息。读者可在学习第5章前先阅读附录C,在学习第6章前先阅读附录D。

目 次

| | |
|--|-----|
| 前言 | III |
| ISO/IEC 前言 | IV |
| 引言 | V |
| 1 范围 | 1 |
| 2 引用标准 | 1 |
| 3 一致性 | 2 |
| 4 定义和缩略语 | 2 |
| 4.1 定义 | 2 |
| 4.2 缩略语 | 5 |
| 5 交换文卷和逻辑记录结构 | 6 |
| 5.1 文卷和逻辑记录结构 | 7 |
| 5.2 逻辑记录——头标和目录 | 8 |
| 5.3 逻辑记录字段区 | 11 |
| 6 用户数据类型和结构的描述 | 11 |
| 6.1 与数据描述相关的 DDR 头标字段 | 11 |
| 6.2 特殊字段标记(标记=0...0 到 0...9) | 12 |
| 6.3 1级文卷中的数据描述字段 | 14 |
| 6.4 2级和3级文卷中的数据描述字段 | 14 |
| 7 编码字符集的用法 | 20 |
| 7.1 编码字符集扩充的宣布 | 21 |
| 7.2 GB 2311 编码字符集扩充 | 21 |
| 7.3 GB 13000.1 编码字符集 | 22 |
| 附录 A(标准的附录) ASN.1 和 FTAM 登记 | 24 |
| 附录 B(提示的附录) GB/T 16684 应用规范 | 28 |
| 附录 C(提示的附录) 对 GB/T 16684 的非形式化简介 | 38 |
| 附录 D(提示的附录) GB/T 16684 数据描述简介 | 44 |
| 附录 E(提示的附录) 数据描述示例 | 48 |
| 附录 F(提示的附录) DDF 层次和网络数据结构 | 53 |
| 附录 G(提示的附录) 数据库数据传送 | 56 |
| 附录 H(提示的附录) 与其他的 OSI 工作的关系 | 58 |

中华人民共和国国家标准

信息技术 信息交换用数据描述 文卷规范

GB/T 16684—1996
idt ISO/IEC 8211:1994

Information technology—Specification for a data
descriptive file for information interchange

1 范围

本标准规定一种交换格式,以便在计算机系统之间传送包含数据记录的文卷或部分文卷。该交换格式并不对任何特定系统的固有文卷规定记录格式,但也可用于此目的。本标准定义了一种一般结构,该结构可被用于在系统之间传送多种数据类型和结构的文卷或记录。本标准规定了描述数据记录内容的方法,但不规定其应用语义,尽管在传送部分可包含应用语义。该交换格式也可用于传送单独的记录、单独的数据字段或单独的子字段及其描述。

本标准规定:

- 信息交换用的与媒体无关的文卷和数据记录描述;
- 包含字符串、位串和数值形式的数据元素、向量、数组和层次的描述;
- 由数据描述记录和伴随数据记录组成的数据描述文卷,它使得信息交换可以最小的特定外部描述出现;
 - 数据描述记录,它描述伴随数据记录中的每个数据字段特征;
 - 文卷和记录结构复杂程度的三种级别;
 - FTAM 非结构化和结构文件类型。

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—89 信息处理 信息交换用七位编码字符集(eqv ISO 646:1983)
- GB 2311—90 信息处理 七位和八位编码字符集 代码扩充技术(eqv ISO 2022:1986)
- GB 2901—92 书目信息 交换用磁带格式(neq ISO 2709:1981)
- GB 5261—94 信息技术 编码字符集的控制功能(idt ISO/IEC 6429:1988)
- GB 13000.1—93 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(idt ISO/IEC 10646-1:1993)
- GB/T 15273.1—94 信息处理 八位单字节编码图形字符集 第一部分:拉丁字母一
- GB/T 16505.1—1996 信息处理系统 开放系统互连 文卷传送、访问和管理 第1部分:概述(idt ISO 8571-1:1988)
- GB/T 16262—1996 信息技术 开放系统互连 抽象语法记法一(ASN.1)规范(idt ISO/IEC 8824:1990)
- GB/T 16263—1996 信息技术 开放系统互连 抽象语法记法一(ASN.1)基本编码规则规范

国家技术监督局 1996-12-18 批准

1997-07-01 实施

(idt ISO/IEC 8825:1990)

ISO 6093:1985 信息处理 信息交换用以字符串形式表示数值的方法

ISO/IEC 9834-2:1993 信息技术 开放系统互连 OSI 登记机构的工作规程 第 2 部分:OSI 文件类型的登记

IEC 559:1989 微处理器系统的二进制浮点运算(ANSI/IEEE 754:1985(R1991))

本标准还与下列文件有关:

与转义序列一起使用的字符集的 ISO 国际登记册。

3 一致性

当所有的数据描述记录和数据记录符合本标准规范时,则交换文卷与本标准一致。一致性语句应规定与文卷内容相一致的版本号 and 交换级。

本标准不规定处理和实现的要求,因此处理和实现可能与本标准一致。

4 定义和缩略语

4.1 定义

本标准采用下列定义。

4.1.1 字母数字字符 alphanumeric character

基本字符集的第 2 列到第 7 列(位),或 GB 13000.1 中表 1 行 00:基本拉丁文中对应的字符。这些字符对应

注:本标准中规定的字符用编码字符集表中的字符位置(列/行)来表示,或者用其缩略语或名称,例如 ESC、SPACE 和数字零或“0”来表示。在控制字段中显式规定的字母数字字符括在双引号内,例如“1”。本标准特定的定界符单元终结符(1/15)和字段终结符(1/14)除外。UT

4.1.2 数组 array

二维或多维的数据结构。

4.1.3 数组描述符 array descriptor

提供数组维数和间隔描述的笛卡儿坐标号或数值数组描述符

4.1.4 数据基地址 base address of

一数据元素,其值等于从数据目录的终结符后面的第一个数据字段的八位位组数,其中规定头标的第一个八位位组为起始序(C)。

4.1.5 基本字符集;BCS basic character set

包含 SPACE 的 GB/T 15273.1 作为 G0 和 G1 集以及 GB 5261 作为 C0 和 C1 集所组成的字符集。

注:此集与 GB 13000.1 基本多文种平面的行 00 兼容,可大大降低对各国变体字符集的依赖性。然而,当前 GB 5261 和 GB 13000.1 之间在 C1 集允许的法上还有差别。

4.1.6 二进制形式 binary form

二进制形式是一个或多个具有特定格式和含义的八位位组的序列。二进制形式中的八位位组可按最高有效八位位组在先(MSOF),其余八位位组按降序次序记录,也可按最低有效八位位组在先(LSOF),其余八位位组按升序次序记录。在一个八位位组中的位总是按降序排列。

4.1.7 二叉树 binary tree

每个结点具有零棵子树或最多两棵子树(称为左子树和右子树)的有根树。当不存在左子树或右子树时,余下的子树保留其左子树或右子树的标识(见附录 F)。

4.1.8 位字段 bit field

一数据字段,它仅由二进制数字组成,当需要时,从右边用二进制零填充成完整八位位组(见字符式位串)。

4.1.9 笛卡儿标号 Cartesian label

由二个(或多个)向量标号元素的笛卡儿乘积形成的标识符数组。数组元素的顺序与直接乘积元素的顺序相同,这样,如果 a 和 b 是向量标号, $a=[a(1), \dots, a(n)]$ 和 $b=[b(1), \dots, b(m)]$, 那么笛卡儿标号 $a * b=[a(1)b(1), a(1)b(2), \dots, a(1)b(m), \dots, a(n)b(m)]$, 其中 $a(i)b(j)$ 是 $a(i)$ 和 $b(j)$ 的并置, 它形成了相应数据数组的 j, i 元素的标识符(即所有的“ a ”是行标号, 所有的“ b ”是列标号)。笛卡儿标号是最一般形式的标号, 特例情况下包含一个向量标号, 或包含单个标号。高阶笛卡儿标号的展开自左至右, 即 $a * b * c=(a * b) * c, a * b$ 先展开。

4.1.10 字符式位串 character mode bit string

表示二进制数字串的字母字符(0或1)序列(见位字段)。

4.1.11 复合数据字段 compound data field

由一个或多个子字段组成的字段, 每个子字段包含一个基本数据元素。

4.1.12 对应二叉树 corresponding binary

表示一棵有序有根树结构的二叉树。该树的最左边的子结点成为该二叉树左子树的根, 该有序有根树中结点的其余子结点通过右链成为有序有根树的先根遍历顺序与二叉树的相同。

4.1.13 数据描述字段 data descriptive

一有标记的字段, 位于 DDR 中, 包含同一标记的文卷中所有数据字段的的标准数据描述。

4.1.14 数据描述文卷 DDF data desc

包含数据描述记录及其伴随数据记录。

4.1.15 数据描述记录 DR data descriptive record; DDR

包含解释伴随数据记录所需的控制参数和数据定义的逻辑记录。数据描述记录是文卷的第一个逻辑记录。

4.1.16 数据字段 data field

位于 DR 中, 包含用户数据的标记字。

4.1.17 数据记录 DR data record; DR

包含用户数据的逻辑记录。

4.1.18 定界结构 delimited structure

由定界符分隔开的数据元素汇集组成定界结构。

4.1.19 定界符 delimiter

分隔数据元素和数据子字段的单个字符(见定界结构的用法)。

4.1.20 目录 directory

包含一个标识符与引用对应数据项的数组的字段, 它由字段终结符终止。

4.1.21 目录项 directory entry

目录中的一个固定长字段, 它包含有关给定记录特定字段的标记、长度和位置的信息。

4.1.22 基本的 elementary

具有既不失意义又不可分割的性质。

4.1.23 项映射 entry map

头标中用来指示目录中项结构的字段。

4.1.24 转义字符; ESC escape character; ESC

用于代码扩充的一个控制字符, 它改变紧跟随其后有限位组合的含义。在 GB 2311 中规定了这个字符的用法。

4.1.25 外部文卷标题 external file title

为交换文卷提供可显示的描述性标题的字符串。不一定需与宿主系统的文卷名相同。

4.1.26 字段 field

用于描述一系列连续的包含一个或多个值的八位位组的类属术语。字段的成分通常称为子字段(见子字段)。

4.1.27 字段标记 field tag

目录项中用于标识数据字段或相关的数据描述字段的字符串。

4.1.28 字段终结符;FT field terminator;FT

用来终止记录中的可变长字段的字符(1/14)。

4.1.29 文卷 file

作为一个单位来处理的有关记录的集合。

4.1.30 森林 forest

不相交的有序有根树的有序集合。

4.1.31 交换格式 interchange format

与本地处理不同的用于交换的数据格式。

4.1.32 交换级;级 interchange level;level

本标准要求的子集的指明。

4.1.33 头标 leader

出现在每个记录的开始,并给记录的处理提供参数的固定长字段。

4.1.34 最低有效八位位组在先;LSOF least significant octet first;LSOF

八位位组的序列,其中最低有效八位位组置于最接近于文卷的始端。

4.1.35 定位;位置 location;position

字段第一个八位位组位置的八位位组计数。头标和目录中的定位与头标的第一个八位位组(0)有关,而字段的定位与数据基地址有关。

4.1.36 映射 to map

在两种不同数据结构元素之间建立对应关系。

4.1.37 媒体记录 media record

按媒体特性的物理记录,由单个系统级写语句写到媒体上,例如一个块,一个扇段或一个传输包(见记录)。

4.1.38 最高有效八位位组在先;MSOF most significant octet first;MSOF

八位位组的序列,其中最高有效八位位组置于最接近于文卷的始端。

4.1.39 n 元组 n-tuple

每个实例具有 n 个有序元素的向量。

4.1.40 空 null

指实体(通常指为数据元素、串或集合)不存在的情况。

4.1.41 数值数组描述符 numeric array descriptor

规定数组维数和每一维维展的数字序列。

4.1.42 有序有根树;树;层次 orderd rooted tree;tree;hierarchy

一种数据结构,它由直接链接的一系列结点所构成,而如此的一个结点是根结点,它没有输入链,其余结点只有一个输入链,结点的多个输出链的次序是有意义的。结点称为父结点和子结点,链的方向是从父结点到子结点。

4.1.43 位置 position

见定位。

4.1.44 先根遍历顺序 preorder traversal sequence

由下列递归算法产生的有序有根树结点的顺序:

- a) 从根结点进入树;
- b) 遍历以前从未遍历的最左子树,若没有如此的子树存在,则停止;
- c) 如果 b)不可能,则返回到该子树的上级结点并转向 b)。

4.1.45 记录;逻辑记录 record;logic record

交换文卷中的连续的八位位组串,前五个八位位组是记录长度(见媒体记录)。

4.1.46 记录长度 record length

其值等于该记录以八位位组为单位的长度的数据元素。

4.1.47 递归树 recurve tree

当树的结点又出现在其自己子树中,则该树为递归树。

4.1.48 相对位置;RP relative position;RP

相对于记录、字段或子字段开始并用十进制整数表示的八位位组位置。第一个八位位组的相对位置编号为“0”。

4.1.49 有符号二进制整数 signed binary integer

补码二进制数。

4.1.50 子字段 subfield

- a) 字段中的一个连续的八位位组串,其位置、长度和数据类型在字段数据描述中描述。
- b) 通常指一个字段的成分。

4.1.51 子字段标号 subfield label

在数据描述中用于标识子字段及其内容的字符串。

4.1.52 标记字段 tagged field

记录中的一个连续的八位位组串,由字段标记标识,其长度和相对位置在目录项中规定。标记字段可出现在 DDR 或 DR 中。

4.1.53 标记 tag

见字段标记。

4.1.54 单元终结符;UT unit terminator;UT

在 DDR 和 DR 两者可变长字段内给几种类型子字段定界所使用的字符(1/15)。

4.1.55 可变长字段 variable-length field

每次出现时其长度可能变化的字段。

4.1.56 向量 vector

一个一维数据结构,其每一实例的元素数量可能不相等(见 n 元组)。

4.1.57 向量标号 vector label

其元素是用来标识数据元素向量中每个元素标号(即“列”标或“行”标)的向量。向量标号是笛卡儿标号的特例。

4.2 缩略语

本标准采用下列缩略语:

| | |
|-------|----------------|
| BCS | 基本字符集 |
| DDF | 数据描述文卷 |
| DDR | 数据描述记录 |
| DFD | 数据字段描述, 见 B. 2 |
| DR | 数据记录 |
| ESC | 转义字符 |
| FT | 字段终结符 |
| FM | 文卷传递与访问管理 |
| LR | 逻辑记录 |
| LSOF | 最低有效八位位组在先 |
| MSOF | 最高有效八位位组在先 |
| RP | 相对位置 |
| SPACE | 间隔字符 |
| TAB | 横向制表字符 |
| UT | 单元终结符 |
| level | 交换级 |
| tag | 字段标记 |
| 0..n | 有一个或多个前导 0 的标记 |

5 交换文卷和逻辑记录结构

本章规定对交换文卷的一般要求, 和对所有逻辑记录公共的要求。这些要求规定了字段以外的逻辑记录结构。对字段内容的要求以及数据字段内容的描述在第 6 章给出。本章构造的非形式化描述见附录 C。

GB/T 16684 的文卷, 逻辑记录及其成分示意图见图 1。

| | | | |
|------|-------|--------|-------------|
| 文卷结构 | LR总结构 | LR详细示意 | 项映射和目录项详细示意 |
|------|-------|--------|-------------|

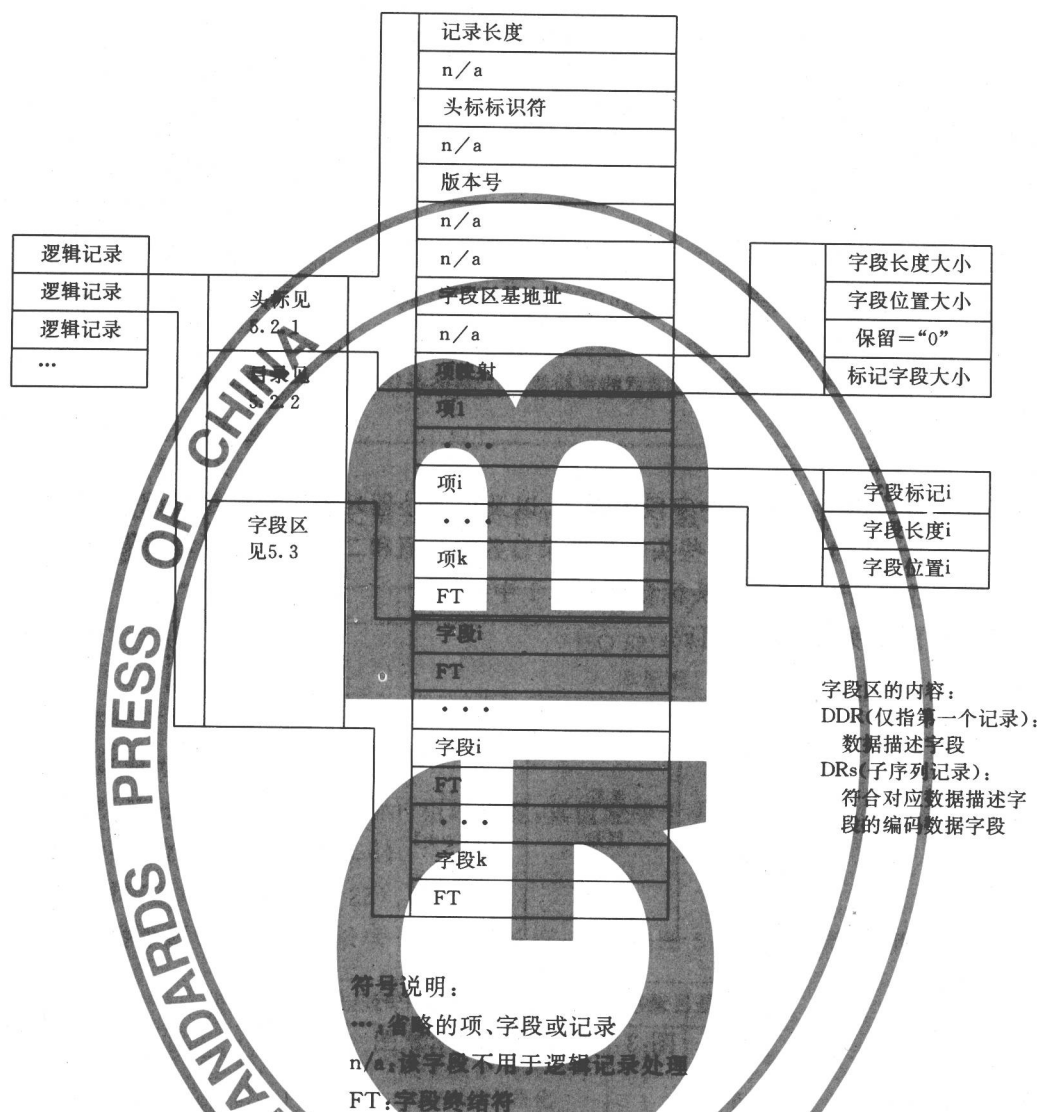


图1 文卷和逻辑记录的示意图

所有记录的头标包含读入记录和将目录分解为项所必须的参数。附加的 DDR 头标包含一些适用于整个文卷的数据描述参数。

所有记录的目录包含标识和定位字段区内每一字段所必须的参数。

仅在第一个记录 DDR 中,字段区包含数据描述字段,每个数据描述字段包含对后续 DR 的字段区中的用户数据译码所必须的信息。

5.1 文卷和逻辑记录结构

本标准规定数据描述文卷(DDF),每个 DDF 由若干逻辑记录(LR)组成。

该文卷可能被记录在物理上不可安装的媒体上,或经通信系统传输。当被记录于卷和文卷结构具有国际标准的媒体上时,对该媒体,每一文卷都应具有所需的国际标准交换文卷标号或头标。当被记录于适当的媒体上时,对该文卷可顺序或随机访问。

标准交换文卷集应由一个或多个图 2 所示的项目集合的实例组成。

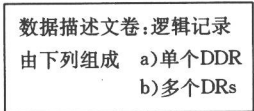


图 2 文卷的示意图

DDR 是 DDF 中的第一个记录。

5.1.1 交换逻辑记录

本标准中规定的逻辑记录应被分块写出,不作进一步区分地跨越定长的媒体物理记录(块或扇段)或传输包。

5.1.2 记录和媒体块的填充

无头标或目录的定长逻辑记录中未用到的部分,以及媒体块或其等价物中未用到的部分,都应用 CIRCUMFLEX 字符(5/14)填充。

注:此规范给出了在所有已知情况下令人满意的可检测的文卷终条件。处理数据终和文卷终条件的逻辑在 C1.6 中讨论。

5.2 逻辑记录——头标和目录

本条包含逻辑记录和逻辑记录中字段的规范,以及记录处理时需涉及的逻辑记录头标和目录的参数的规范。从属于八位位组计数的某些头标和目录参数的数值和二进制编码两者均得到支持并在 5.2.1.6 中描述。涉及数据描述的 DDR 头标参数在 6.1 中描述。

逻辑记录应由图 3 所示的区域组成。

| 区域名 | 长度 |
|-----|------------------|
| 头标 | 24 |
| 目录 | $k \times p + 1$ |
| 字段区 | 可变长的 |

其中:k 是目录项数,p 是记录目录项中的八位位组数

图 3 逻辑记录的示意图

5.2.1 逻辑记录头标

每个逻辑记录头标均应由图 4 所示的字段组成,并将在 5.2.1.1 到 5.2.1.5 和 6.1 中进一步规定。

图 4 中“用途”列指明作为记录处理控制的每种字段的用法。“全部”表示对所有记录该字段都存在一个有意义的值,“DDR”表示仅对 DDR 该字段存在一个有意义的值。

| 相对位置 | 长度 | 用途 | 字段名 | 逻辑记录头标内容 | | 相关条号 |
|------|----|-----|----------|-------------|---------|---------|
| | | | | DDR | DR | |
| 0 | 5 | 全部 | 记录长度 | 八位位组 | 八位位组 | 5.2.1.1 |
| 5 | 1 | DDR | 交换级 | “1” “2” “3” | 间隔字符 | 6.1.1 |
| 6 | 1 | 全部 | 头标标识符 | “L” | “D” “R” | 5.2.1.2 |
| 7 | 1 | DDR | 内码扩展指示符 | 字符 | 间隔字符 | 6.1.2 |
| 8 | 1 | DDR | 版本号 | 间隔字符 “1” | 间隔字符 | 5.2.1.3 |
| 9 | 1 | DDR | 应用指示符 | 字符 | 间隔字符 | 6.1.3 |
| 10 | 2 | DDR | 字段控制长度 | 数字 | 若干间隔字符 | 6.1.4 |
| 12 | 5 | 全部 | 字段区基地址 | 八位位组 | 八位位组 | 5.2.1.4 |
| 17 | 3 | DDR | 扩展字符集指示符 | 字符 | 若干间隔字符 | 6.1.5 |
| 20 | 4 | 全部 | 项映射 | 数字 | 数字 | 5.2.1.5 |

图 4 逻辑记录头标的示意图

5.2.1.1 记录长度字段(LR RP 0 到 4)

本字段应规定以八位位组为单位的 LR 总长度,包括本字段的 5 个八位位组在内(语法见 5.2.1.6)。

5.2.1.2 头标标识符字段(LR RP 6)

本字段应标识逻辑记录的类型并规定其头标和目录的性质。

若 DDF 中所有其余数据记录都具有相同的头标和目录,则允许在其头标标识符字段放置一个“R”,以便在后续记录中省略头标和目录。

头标标识符允许的取值及其含义是:

| | |
|---|---|
| L | 表示该记录是数据描述记录 |
| D | 表示该记录是一个数据记录,且下一个数据记录具有头标和目录 |
| R | 表示该记录是一个数据记录,且在任何后续逻辑记录中都不可能发现头标和目录。当前逻辑记录的头标和目录适用于后续的每一个逻辑记录 |

5.2.1.3 标准版本号(LR RP 8)

在 DDR 中,本字段应包含标识交换文卷所符合的版本号。版本号对记录和数据处理设置条件和限制。允许的取值及其含义是:

| 值 | 含 义 |
|------|---------------|
| 间隔字符 | ISO 8211:1985 |
| “1” | 本标准 |

在其余的逻辑记录和 DR 中不用此字段,应包含间隔字符。

5.2.1.4 字段区基地址(LR RP12 到 16)

本字段应规定字段区中第一个字段的相对位置,其值应等于头标和目录(包括目录结尾处的字段终结符在内)总的八位位组数(见 5.2.1.6)。

5.2.1.5 项映射字段(LR RP20 到 23)

本字段规定每个逻辑记录中目录项子字段的长度,并应由图 5 所示的子字段组成(见 5.2.1.6)。

| 相对位置 | 子字段名 | 长度 | 内容 |
|------|-----------|----|-----|
| 20 | 字段长度字段的大小 | 1 | 数字 |
| 21 | 字段位置字段的大小 | 1 | 数字 |
| 22 | 留待将来标准化 | 1 | 数字0 |
| 23 | 字段标记字段的大小 | 1 | 数字 |

图 5 逻辑记录项映射示意图

5.2.1.5.1 字段长度字段的大小(LR RP 20)

本子字段应规定目录项中以八位位组为单位的字段长度字段的大小,并且是“1”到“9”中的某个数字。

5.2.1.5.2 字段位置字段大小(LR RP 21)

本子字段应规定目录项中以八位位组为单位的字段位置字段的大小,并且是“1”到“9”中的某个数字。

5.2.1.5.3 留待将来标准化(LR RP 22)

本子字段留待将来作扩充项映射的标准化,其值应是数字“0”。

5.2.1.5.4 字段标记字段的大小(LR RP 23)

本子字段应规定目录项中以八位位组为单位的字段标记字段的大小,并且是“1”到“7”中的某个数字。所有逻辑记录中此子字段的值均应相同。

5.2.1.6 计数和字段位置的替代形式

本标准对八位位组计数和位置都可提供数值和二进制编码两种形式。

DDR 记录长度中的第 5 个八位位组(DDR RP4)应标明使用哪一种替代形式。所允许的值应具有下列含义:

| 值 | 表示的含义 |
|---------|--------------------|
| “0”~“9” | 使用字符形式(标志是值的最低有效位) |
| “B” | 使用最高有效八位位组在前的二进制形式 |
| “b” | 使用最低有效八位位组在前的二进制形式 |

形式应是互斥的,且在整个文卷的头标或目录中使用同一种形式。

对记录长度字段、字段区的基地址字段、字段长度字段和字段位置字段的内容,逻辑记录的头标和目录应采用下列形式:

a) 由十进制数字组成的字符形式。采用这种形式时,上述字段应包含一个右对齐的十进制整数,左边用 0 字符填充。记录长度字段为 0 时表示记录长度超过 99999。

b) 仅在 DDR 记录长度字段中,二进制形式由一个有符号二进制整数和一个形式标记组成。记录长度字段和字段区的基地址字段中的前 4 个八位位组应用来存放一个 32 位的有符号二进制整数。每项映射实例应限于取下列表中之一个值集:1,1,0,2;2,2,0,2;2,2,0,4 或 4,4,0,4。取决于项映射中所规定的字段大小,字段长度和字段位置的值应是一个适当的 8 位、16 位或 32 位的有符号二进制整数。DDF 记录成分的最大值受能被置于对应的控制字段中的最大值限制。未用的位置 DR RP 4 和 DR RP 16 中应包含“0”。

5.2.2 逻辑记录目录

逻辑记录目录应由重复的逻辑记录目录项组成,其字段长度应在项映射字段(见 5.2.1.5)中规定。逻辑记录目录应对每个字段区中每个字段包含一个逻辑记录目录项,并用字段终结符终止。逻辑记录目录项规定字段区中对应字段的定位和长度,并由图 6 所示的字段组成。每一项都应按顺序包含一个字段标记,一个字段长度和一个字段位置,它们由 $m+n+t$ 个八位位组组成,其中:

- m 是字段长度字段的大小;
- n 是字段位置字段的大小;
- t 是字段标记字段的大小。

逻辑记录目录项应与字段一一对应,次序相同。

每个目录项的定位、长度和内容在图 6 中规定,它们的含义在 5.2.2.1 至 5.2.2.3 中说明。

| 相对位置 | 字段名 | 长度 | 内容 |
|--------------|------|----|------|
| $p(i-1)$ | 字段标记 | t | 字母数字 |
| $p(i-1)+t$ | 字段长度 | m | 数字 |
| $p(i-1)+t+m$ | 字段位置 | n | 数字 |

其中: $p=t+m+n, i=$ 目录项的序号

图 6 逻辑记录目录项示意图

5.2.2.1 字段标记字段

本字段应包含标识对应字段的字段标记,并由在项映射(5.2.1.5.4)中规定的 1 到 7 个字母数字字符组成。

5.2.2.2 字段长度字段

本字段以八位位组为单位规定与其对应的字段(见 5.2.1.6)的长度。字段的长度应包括字段终结符。

5.2.2.3 字段位置字段

本字段规定对应字段(见 5.2.1.6)中第一个八位位组的相对位置。它应包含一个值,将该值与字段