数据部之二

——数据部的较高技巧

一、 数据在计算机内的表示形式

1. 计算机内存的组织形式：

   计算机是以二进制形式存放数据的，内存的最小单位是二进制位。8 位构成一个字节。

2. 字符型数据在内存中的存储形式

   不论是字母还是数字，都按一个字节存放一个字符存放。

3. 数值型数据在内存中的存储数据

   （1） 外部十进制：一个数字在内存中占一个字节

   （2） 外部浮点数形式：

   一般形式： **数符 数值部分 E 阶码符 阶码**

   例： 77 A PIC +9.99999E+99 表示 +1.23456E+59

   注意：外部符点形式的数据不能用 VALUE 语句赋初值。

   （3） 内部十进制数：只能存放 0—9 十个数字，每两个数字占一个字节。

   （4） 定点二进制形式：将十进制数转换成定点二进制数，存入内存。

4. 数据描述与存储形式的关系

   （1） 字母型、字符型、编辑型、外部十进制数据和以外部浮点形式表示的数据用标准数据形式来存放

   （2） 数值型数据可以由程序员任意选定存放形式。

   （3） 数据在计算机内进行运算，都是化成二进制数以后再进行的。

一、 用法子句（USAGE 子句）(SAMPLE7-1)

使用用法 (USAGE) 子句可以使程序设计者自由选择数据在内存中的存放形式。

USAGE 子句的一般格式为

```
[ USAGE IS ] { DISPLAY
             { COMPUTATIONAL
             { COMP
```

其中“USAGE IS DISPLAY”的意思是“显示型的用法”，即此数据适合打印、显示，它采用标准数据类型。COMPUTATIONAL 和 COMP 是同一意思，表示“计算型的用法”表示此数据类型适合计算，它采用适用于计算用的定点二进制形式或内部浮点形式。

标准 COBOL 只列出 DISPLAY 和 COMPUTATIONAL 两种用法的格式。各种计算机还规定了它可采用的有哪几种数据存放形式以及用什么来代表他们，在使用前，请查阅所用计算机的 COBOL 说明书。

说明：（1）USAGE 子句是用来指定数据项在内存中的存储形式的。

（2）如果省略 USAGE 子句，则隐含表示用 DISPLAY 形式。

（3）如果对组合项描述为一种存储形式，则表示这个组合项的下属各初等项都是这种形式

（4）USAGE 子句指定的数据存储形式不应与 PIC 子句指定的数据类型矛盾。

（5）长、短浮点形式已确定了内存的长度，不应再用 PIC 子句。

（6）在传送或运算时不同存储形式的数据型数据间可互相转换。

（7）在用 DISPLAY 语句显示数据项的内容时，如果数据项的 USAGE“用法中”不指定 DISPLAY，则在显示前，计算机自动将内存中的数据形式转换（EBCDIC 码或 ASCII 码，试计算机系统而定），然后以字符形式显示。

（8）如果用 WRITE 语句，则直接输出，不进行转换。

二、符号子句(SIGN 子句)

SIGN 子句用来指定数值型数据描述体中运算符号的状态和位置。

说明：（1）在没有 SIGN 子句时，数值的符号是存放在数据项最后一个字节中的。

（2）用 SIGN 子句可以指定符号在数值的前部还是后部。

（3）指定符号单独占一个字节，用“SEPARATE”可选项。

SIGN 子句的一般形式

[SING IS] { LEADING / TRAILING } [SEPARATE CHARACTER]

例：**77 A PIC S9(3) VALUE －125 SIGN LEADING(TRAILING) SEPARATE.**

注意：（1）它只能用于 PIC 字符串中含有“S”的数值型数据描述体中。

（2）使用 SIGN 子句的数据项的用法(USAGE)应当是 USAGE DISPLAY。不能用于计算型用法的数据项

（3）用 SEPARATE 可选项时，内存中增加了一个字节，用来放符号标志。

（4)如果一个数据项的描述体中包含 SIGN 子句,则数据项的值应包括正或负的符号，否则会出错。

三、重定义子句(REDEFINES 子句) (SAMPLE7-2:冒泡排序)

不同的数据项可以共用内存中的同一段空间。

重定义子句的一般形式：

层号　数据名 1　REDEFINES 数据名 2

例: 01 A.

02 A1 PIC X(6).

02 B1 REDEFINES A1.

03 B11 PIC X(4).

03 B12 PIC 99.

02 C1 REDEFINES A1 PIC 9(6).

说明：（1）数据名 2 与数据名 1 的层号必须相同。即它们应是同一层次的。REDEFINES 子句不能用于 88 层和 66 层。

（2）用 REDEFINES 子句的描述体应紧跟在被重新定义的数据项的描述之后，中间不能插入其他项的描述说明。

（3）可以多次重定义，但必须紧跟出现，而且要求使用最初定义的数据名。

（4）REDEFINES 子句不能用于文件节的 01 层中，因为文件节中 01 层描述的是记录，但工作单元节中的 01 层是可以用 REDEFINES 子句重新定义的，因为这里的 01 层不是指输入输出文件的记录，而是指组合项。

（5）用 REDEFINES 子句可以改变数据的结构，但两个数据名的长度应相同。

（6）REDEFINES 子句应在其它子句之前。

（7）内存中的值为数据名 1 和数据名 2 共享。也就是说，重定义后两个数据名的名称和两种数据结构同时存在，都有效。程序中可使用其中任何一个。他们在内存中为同一段存储单元。如果改变了内存内容，则二者的值都因而改变。

（8）重定义子句所在的数据描述体中不能使用初值子句赋初值。

五、重命名子句(RENAMES 子句)

用 REDEFINES 子句可以在不改变数据项的长度的前提下,重新定义数据区的名称和数据结构的形式（包括重新定义初等项的类型和长度)。用重命名（RENAMES）子句可以把原来已定义

的某些数据项重新组合成一个新项，并以一个新名字来代表它。但重命名子句不能改变原来初等项的类型、长度和属性。

重命名子句的一般形式

66 数据名 1 RENAMES 数据名 2 [THRU 数据名 3]

例：01 A.

```
   02 B.
     03 G PIC X(2).
     03 H PIC 99.
   02 C
     03 I PIC 99V99.
     03 J PIC S999.
   02 D PIC 9(5).
   02 E PIC XX.
  66 K RENAMES G THRU I.
  66 M RENAMES B THRU C.
  66 N RENAMES E.
```

说明：（1）层号只能用 66，它必须紧跟在 01 层记录中最后一个数据描述体之后，因为它是对记录中有关部分重新组合和命名的。

（2）如无 THRU 部分，则数据名 1 和数据名 2 代表的是同一内容。

（3）用 THRU 时，数据名 2 在记录中的位置应在数据名 3 之前，而且数据名 3 不应包括在数据名 2 之中。

（4）RENAMES 子句只能用在工作单元节中，不能用于文件节中。

六、遇零置空子句

BLANK 子句的作用是：当数据项的值为零时，使它的内容改变为空白。这个子句只能用于数值型或编辑数值型的初等项。

例： 03 A PIC $(5).99 BLANK WHEN ZERO.

七、对齐子句(JUSTIFIED 子句)

JUSTIFIED 子句的一般形式

{JUSTIFIED / JUST} RIGHT

77 B PIC X(5) JUST RIGHT.

说明：JUSTIFIED 子句只能用于字母型和字符型数据，而不能用于数值型数据项和编辑型数值项；因为后者是按小数点位置对齐的方式定位。

八、同步安置子句(SYNCHRONIZED 子句)

在许多定字长的计算机中，一个机器字往往指定为四个字节。两个机器字之间为“自然边界”，从内存中取数据是，以机器字为单位。

用同步安置子句(SYNCHRONIZED 子句)可以指定数据项在内存中如何按自然边界来安置。

同步安置子句的一般形式

{SYNCHRONIZED / SYNC} [LEFT / RIGHT]

如果用 SYNCHRONIZED LEFT，表示数据项从一个机器字的左边自然边界开始存放，如果数

据项的长度不是一个机器字，则右边补零（数值型数据项）或空白（非数值型数据项）。后继数据项不能使用这些空余的字节，而必须从下一个机器字的左边自然边界开始存放。这些空余的字节称作“填补字节”。如果用 SYNCHRONIZED RIGHT 则正好相反。如果程序运行时间很长，使用 SYCN 子句来减少执行时间是值得的。

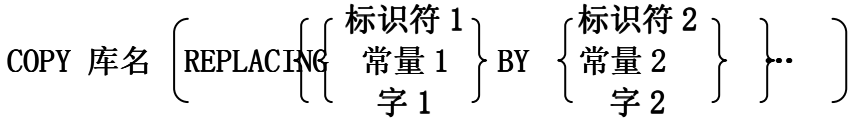
九、多格式数据记录—记录取得重叠(SAMPLE7-3)

每一个文件在内存区都相应的分配一个记录区，该记录区的数据结构应在数据部的文件节中描述。但有时从文件中读入的数据不是属于同一种格式的，它们代表不同的内容。如果两个不同格式的记录长度不同，则内存记录区的长度按长者确定。

十、复写语句(SAMPLE7-4)

COBOL 程序中的 DATA DIVISION 部分往往是很长的，因为其中包含许多的数据项的描述。可以利用 COPY 语句使某些记录描述和数据描述为不同的程序共用。为此要建立一个“源程序库”，将上述这些共同使用的程序中的某一部分事先存入库中。

复制语句的一般形式



“字”的含义是不超过 30 个字符组成的字符序列。它包括数据名、条件名、过程名、表意常量、助忆名或保留字等。复写语句可以用在除标识部以外的其他三个部中，但在环境部中使用复写语句的意义不大。使用复写语句的目的是有利于在不同文件中用标准化的记录书写程序，减少写程序的工作量，便与程序的保存和修改。