

FoxBASE+索引功能的讨论

宣仲良 (苏州教育学院)

摘要:本文指出了“FoxBASE+索引功能的扩展”一文中若干欠妥之处,并提出了改进方案。

“FoxBASE+索引功能的扩展”一文[1](下面简称原文)主要介绍了在多重索引下,D型字段和带负值N型字段的索引方法,以及实现双向索引的途径和通用程序。原文在运用函数和某些算法以扩展FoxBASE+的索引功能方面,对数据库程序设计者确有启发。但原文忽略了FoxBASE+的一些基本假定,以至留下了多处“隐患”。本文指出这些欠妥之处并给出其改进方法。

一、D型字段的正确索引问题

原文认为:用SET DATE ANSI将D型数据格式转换为YY.MM.DD后,再用DTOD(<D型表达式>)就能在多重索引下正确索引D型字段。

这种解决方法虽能在一定条件下(习惯于使用ANSI日期格式,且无非二十世纪日期数据)有效,但明显缺乏“通用性”(在欧洲就不宜用),还可能无谓地增加索引关键字(即索引表达式的值)的长度(在SET DATE ON时)。其实,FoxBASE+的DTOD()函数本身就提供了在多重索引下,任何允许格式(五种)的D型表达式均可被正确索引的功能-DTOD(<D型表达式>,1)。此函数值是C型有世纪前缀的年月日数字串,长度固定为8个字符;在SET CENT OFF时系统也会自动加上二十世纪前缀“19”。因此含有非二十世纪的日期数据时,必须设定SET CENT ON且年份用四位数字表示,才能保证D型表达式被正确索引。

二、双向索引通用程序的几个问题

1.建立正确的索引表达式的判断问题

原程序仅以“索引表达式长度<236(字符)”作为判断依据,这有失偏颇。

FoxBASE+规定:索引关键字长度必须<100(字符);同时又规定任何命令行长度必须<254(字符)。因此,实质上FoxBASE+对索引表达式本身的长度和值(即索引关键字)的长度都有限制。同时满足这两方面限制的索引表达式,是正确建立索引文件的必要前提。原文中的通用程序缺乏对索引关键字长度的判断,因而是片面的。

在原文所考虑的范围,即按若干字段建多重索引表达式,

索引关键字长度可按下式计算:(设SET CENT ON)

索引关键字长度 = $10 + \sum (C \text{ 或 } D \text{ 型索引字段长度} + 1) + \sum (D \text{ 型索引字段长度} - 1)$

上式中常数10是FoxBASE+允许的最大记录指针值(即数字1000000000)的长度;+1为防止C、N型数据经运算并通过函数进行类型转换后,形成索引关键字(或其一部分)时数值溢出;-1是因为经DTOD(<D型表达式>,1)转换后形成的索引关键字9或其一部分)的长度比原D型字段长度短2个字符(-2+1=-1)。

2.索引文件名问题

原文的通用程序中,对索引文件的命名采用了固定名称:TEMP(原程序倒数第三行为:INDE ON &MEX TO TEMP),并且设置了SET SAFE OFF。这意味着用该程序不管建立了多少个索引文件,最终只保留最后一个。这显然不妥。

作为通用程序,有必要增加一条用一个内存变量(设为AA)接受用户输入的需建立的索引文件名,再将建索引文件命令改为:

INDE ON &MEXTO &AA

3.原程序的使用条件和期限问题

按照原作者的设计,原程序只能在SET CENT OFF和D型字段中无非二十世纪日期数据的前提下,才能有效使用至1999年12月31日。过了这个期限,D型字段的升、降序索引都会出错。升序索引出错的原因在本文的第一部分已论及,降序索引出错的原因是原程序使用了如下语句:

MEXP = '+STR(CTOD("99.1.2.31")-' + MF + ',5)'

显然将CTOD("99.1.2.31")作为被减数,减数大于它时,差为负值。(请注意:CTOD("99.1.2.31")的值,在SET CENT OFF或ON时均为1999.1.2.31)这样,“D型字段降序索引”就变为“带负值N型字段索引”,按原文第一部分“多重索引下负值字段索引偏差的纠正”所分析的理由,上述语句肯定出现索引偏差。

避免这种偏差的方法是:使用FoxBASE+允许的最大日期数:9999.1.2.31,并且将转换成C型的数字串长度改为7,即原语句改为:

MEXP = '+STR(CTOD("9999.1.2.31")-' + MF + ',7)'

顺便指出一下:目前市面上流行的FoxBASE+(2.00版或2.10版)中的CTOD()函数允许使用的日期范围是:公元1582年10月15日至9999年12月31日,而不是使用手册和有关书籍中介绍的公元100年1月1日至9999年12月31日。因此两个D型数据相减,被减数为最大允许值、减数为最小允许值时,其差<3074323(天),这也就是C型数字串长度定为7的理由。

最后需要说明的是:原文中提出的降序索引的算法是可行的,但不是最简便的。更好的算法,带筛选器索引和逻辑字段索引等问题的讨论,拟另文阐述。