

动态统计算法的系统开发策略研究

王志龙 韦秀长 (合肥工业大学计算机网络系统研究所 230009)

摘要:文章针对统计对象为动态统计算法的一类问题的系统特点,提出了面向数据对象的系统开发策略,并对其设计思想和主要技术问题作了讨论,给出应用实例。

关键词:统计对象 动态统计算法 面向数据对象

1. 引言

管理信息系统的基本功能之一就是依据统计对象中各项目的统计算法对系统中的数据进行加工处理,以满足特定的应用要求。就统计对象的统计算法而言,在开发 MIS 时,我们经常碰到的是统计对象为静态统计算法的一类设计问题,这类统计对象其统计算法一旦确定,则在多个统计期内保持不变,统计算法变更的频率低。因此,这样一类统计对象软件设计时只需要按照统计程序即可。但是,在 MIS 开发中我们也会经常碰到统计对象为动态统计算法的另一类设计问题,这类问题其统计算法在各个统计期变更频率高,需要随时确定,这样在开发这类统计对象为动态统计算法的系统时,就不能沿用前一种系统开发方法,而必须寻求一种新的系统开发策略来解决。下述便是本文采用的面向数据对象技术来解决这类问题的一些探讨。

2. 设计思想

根据系统的特点将系统中的数据分为若干个数据类,设计一套统计算法表示规则及相应的统计算法解释机制,建立记录统计对象项目的统计算法库;通过统计算法预生成功能,自动生成统计算法库中统计对象各项目当前统计期的初始统计算法;利用统计算法定义功能最终确定统计对象各项目在当前统计期的统计算法;使用统计顺序确定功能自动确定统计对象各项目统计计算的先后顺序;最后通过统计算法解释机依次解释各项目的统计算法,从而实现统计对象统计结果的自动生成。

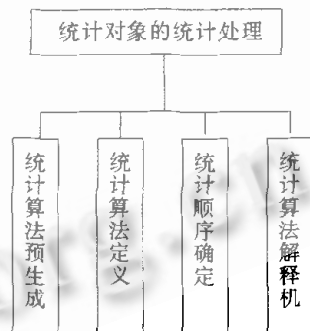
基于上述设计思想,实现统计对象为动态统计算法的统计功能的基本构件如下:

3. 若干技术问题

采用面向数据对象的设计方法时,主要应解决好以下技术问题:

(1)数据类设计。数据类设计的目的是为了使用户

在算法定义时快速地选取数据对象。合理地设计数据类是有效使用面向数据对象的动态算法定义方法的前提。它决定了用户快速进入到具体数据对象的路径。由于我们在系统设计时将系统中的数据分放在不同的数据库中,因此,数据类的划分应根据数据来源和性质并结合数据所存放的数据库等因素进行逐级分类,以使得每一小类保持适度量的数据对象。为了便于对数据类操作,我们通常设计一个数据类库,其一般性结构包括数据类别码、数据类名称等字段。



(2)算法表示规则设计。算法表示规则不仅关系到用户进行算法定义时的操作简便性和算法的可读性,而且决定了算法解释机设计的难易性。一般地,所设计的算法表示规则应能方便地显示操作运算符,参与统计计算的数据对象所属的数据类。其一般模式为:〈分隔符〉〈数据类别码〉〈数据对象〉〈分隔符〉[〈运算符〉]。值得注意的是,这里的数据类别码应同数据类库中的数据类别码相一致,数据对象表示应易于用户理解,能据此唯一地确定参与统计的数据值,还应能对其实现快速查找。

(3)统计算法库设计。统计算法库是记录各统计项目的统计算法的,它是进行统计计算的依据。由于各个

统计期的统计算法可能有所差异,为便于了解不同统计期的统计算法以及对已过去的某一时期的数据因某种原因而必需进行重新统计,因此,该库中除了能记录每一统计项目的统计算法、各项的计算顺序外,还应记录各项的统计算法所属统计期。一般地,统计算法库的结构设计时应包括统计期、统计项目、统计算法、计算顺序等字段,并按统计期和计算顺序字段建立索引文件。

(4)统计算法定义与检查。系统自动显示统计对象各项的初始算法,用户可根据约定的统计算法表示规则,直接写入或通过数据类的菜单向导选取某一数据类中的数据对象,以确定某个项目新的统计算法。此外,必须采取各种保证措施,自动对用户所定义的统计算法进行算法表示规则的合法性检查、数据类和数据对象的存在性检查、统计项目间是否存在算法的循环定义检查等,以保证用户所定义的统计算法正确、有效。

4. 应用实例

炼油企业的特点是工艺过程连续、生产装置投入的来源和产出的去向多变。每班生产统计的主要任是根据收集的油罐收付、装置仪表的动态参数,统计生成装置生产台帐和油品生产台帐。在企业生产过程中,由于受到计量仪表不一定齐备、计量仪表发生故障、计量数据不准确等大量不确定因素的影响,统计人员在做每班的油品生产情况统计时,都必须根据生产调度指令和经验,确定生产装置的投入和产出量的统计算法,也就是说其统计算法是动态的。

基于前述的面向数据对象的设计思想,我们进行了如下设计工作:划分油品收付(A)、装置生产记录(B)和装置投入产出(C)三个数据类;算法表示规则采用“|〈数据类别码〉〈数据对象〉|〔算术运算符〕”模式;装置投入产出项目统计算法库的结构如下所示,并按日期(RQ)+班次(BC)+计算顺序(JSSX)建立索引。

字段名	类型	长度	小数位	中文含义
RQ	D	8		日期
BC	C	1		班次
ZZBM	C	3		生产装置编码
ZZXMM	C	2		生产装置的投入产出项目代码
JSGS	C	24		统计算法表达式
JSSX	N	2		统计项目计算顺序

每班统计操作四步曲为:班产算法预生成、班产算法

调整、项目统计顺序确定和班产统计。

(1)班产算法预生成:用户输入统计日期和班次后,系统将自动进行该日期和班次的统计算法是否已经存在等方面的检查,以防用户误操作。若条件成立,系统则根据生产调度系统下达的本班调度指令和一般性统计规则,自动生成本班各生产装置的投入产出项目的初始算法。

(2)班产算法调整:屏幕显示出生产装置的投入产出项目,画面格式如下:

日期 97.07.01		班次: 1
装置名称	统计项目名称	统计算法
常减压(111)	处理量(10) 初顶(11) 常顶(12)	A0101 付常减压 + A0106 付常减压 C11110 * 0.4
F4—算法修改		F5—菜单向导

(注:A—油品收付记录类码 0101、0102—油罐号
C—装置投入产出类码 111—常减压装置代码 10—常减压处理量项目代码)

统计人员可按↑、↓键查看不同项目的统计算法,当按下F4键即可修改当前光标所在项目的算法。修改算法时,可按F5通过菜单向导选取生成或直接写入算法表达式,在退出该项目修改过程时,系统自动对所定义的算法进行是否符合算法表达规则、是否存在统计项目循环定义等一系列的合法性检查,以确保用户定义的算法合法、有效。

(3)项目的统计顺序确定:系统首先检查统计算法表达式中所有的数据类别码,若不含有标志装置投入产出项目类码(C)的项目自动赋为1,否则,按照一定的算法规则确定出该项目的计算顺序号。根据统计计算的先后次序依次赋为2,3,...

(4)班产统计:当用户输入待统计的日期和班次后,系统自动做本班次的统计算法是否已经定义、是否重新进行统计等合法性检查之后,调用统计算法解释机,按照已定的统计顺序依次对各项的统计算法进行解释求解,并显示最终的统计结果。

5. 结束语

如何解决好统计对象为动态统计算法的一类问题对于系统的开发具有十分重要的意义,本文提出的面向数据对象的开发策略已成功的应用到前述的某炼油厂生产统计等多个系统开发中,实践证明实施效果良好。

(来稿时间:1997年8月)