IDEF1x数据建模方法调研报告

IDEF的含义是集成计算机辅助制造（Integrated　Computer－AidedManufacturing，ICAM)DEFinition。最初的IDEF方法是在美国空军ICAM项目建立的。根据用途，可以把IDEF族方法分成两类：第一类IDEF方法的作用是沟通系统集成人员之间的信息交流。主要有：IDEF0、IDEF1、IDEF3、IDEF5。IDEF0通过对功能的分解、功能之间关系的分类（如按照输入、输出、控制和机制分类）来描述系统功能。IDEF1用来描述企业运作过程中的重要信息。IDEF3支持系统用户视图的结构化描述。IDEF5用来采集事实和获取知识。第二类IDEF方法的重点是系统开发过程中的设计部分。目前有两种IDEF设计方法：IDEF1X和IDEF4。IDEF1X可以辅助语义数据模型的设计。IDEF4可以产生面向对象实现方法所需的高质量的设计产品。

IDEF1X是IDEF系列方法中IDEF1的扩展版本，是在E-R(实体联系)方法的原则基础上,增加了一些规则，使语义更为丰富的一种方法。概念模型设计常用IDEF1X方法，它就是把实体-联系方法应用到语义数据模型中的一种语义模型化技术，用于建立系统信息模型。IDEF1的最初形式是在P.P.S (Peter) Chen的实体联系模型化概念与P.P.(Ted) Codd的关系理论的基础上发展起来的。IDEF1X是IDEF1的扩展版本，除在图形表达和模型化过程方面的改进外，还对语义进行了增强和丰富。例如：分类联系 (Categorization Relationships)的引入。

IDEF1X是语义数据模型化技术，它具有以下的特性：(1)   支持概念模式的开发。IDEF1X语法支持概念模式开发所必需的语义结构，完善的IDEF1X模型具有所期望的一致性、可扩展性和可变换性。(2)  IDEF1X是一种相关语言。IDEF1X对于不同的语义概念都具有简明的一致结构。IDEF1X语法和语义不但比较易于为用户掌握，而且还是强健而有效的。(3)  IDEF1X是便于讲授的。语义数据模型对许多IDEF1X用户都是一个新概念。因此，语言的易教性是一个重要的考虑因素，设计IDEF1X语言是为了教给事务专业人员和系统分析人员使用，同样也是教给数据管理员和数据库设计者使用的。因此，它能用作不同学科研究小组的有效交流。(4)   IDEF1X已在应用中得到很好地检验和证明。IDEF1X是基于前人多年的经验发展而来的，它在美国空军的一些工程和私营工业中充分地得到了检验和证明。(5) IDEF1X是可自动化的。IDEF1X图能由一组图形软件包来生成。商品化的软件还能支持IDEF1X模型的更改、分析和结构管理。

IDEF1X把实体－联系方法应用到语义数据模型化中，IDEF1的最初形式是在P.P.S (Peter) Chen的实体联系模型化概念与P.P.(Ted) Codd的关系理论的基础上发展起来的。IDEF1X是IDEF1的扩展版本，除在图形表达和模型化过程方面的改进外，还对语义进行了增强和丰富。例如：分类联系 (Categorization Relationships)的引入。

IDEF1X模型的基本结构和ER模型基本类似，主要有以下元素：  
（1）实体(如人、地点、概念、事件等)用矩形方框表示；（2）实体之间的关系(联系)，用方框之间的连线表示；（3）实体的属性，用方框内的属性名称来表示。

能建立完整的IDEF1x概念模型并支持直接将模型转换为物理数据库的结构。实体之间的关系可以分为确定关系和不确定关系。确定关系又分为连接关系和分类关系。连接关系也称“父子关系”，它是两个实体之间的联系或连接，一个实体(子实体)依赖于另一个实体(父实体)。分类关系表示实体间的一种分层结构，一个实体(类属实体)表示这些事物的全集，其它(分类实体)则为其子集。不确定关系又称“多对多关系”，两个实体间相互存在着一对多的联系。连接关系又分为标识关系和非标识关系。判别一个关系是标识关系还是非标识关系只要区分子实体的主键，看是否需要父实体的外键来共同作为主键，需要则为标识关系（Identifying）；如果子实体自己的主键就可唯一标识则它为非标识关系（Non-Identifying）。在标识关系中的子实体称为依赖实体，用圆角矩形表示；其它用方角矩形表示的就是独立实体。分类关系根据表示分类的实体集是不完全分类还是完全分类的又可以分为不完全分类关系和完全分类关系。

很多的数据库建模工具都支持IDEF1x方法，如CA公司的ERWin、Sybase公司的PowerDesigner以及微软公司的Visio等。

使用IDEF1X方法创建E-R模型的步骤如下：

（1）初始化工程

这个阶段的任务是从目的描述和范围描述开始，确定建模目标，开发建模计划，组织建模队伍，收集源材料，制定约束和规范。收集源材料是这阶段的重点。通过调查和观察结果，业务流程，原有系统的输入输出，各种报表，收集原始数据，形成了基本数据资料表。

（2）定义实体

实体集成员都有一个共同的特征和属性集，可以从收集的源材料—基本数据资料表中直接或间接标识出大部分实体。根据源材料名字表中表示物的术语以及具有“代码”结尾的术语，如客户代码、代理商代码、产品代码等将其名词部分代表的实体标识出来，从而初步找出潜在的实体，形成初步实体表。

（3）定义联系

IDEF1X模型中只允许二元联系，n元联系必须定义为n个二元联系。根据实际的业务需求和规则，使用实体联系矩阵来标识实体间的二元关系，然后根据实际情况确定出连接关系的势、关系名和说明，确定关系类型，是标识关系、非标识关系（强制的或可选的）还是非确定关系、分类关系。如果子实体的每个实例都需要通过和父实体的关系来标识，则为标识关系，否则为非标识关系。非标识关系中，如果每个子实体的实例都与而且只与一个父实体关联，则为强制的，否则为非强制的。如果父实体与子实体代表的是同一现实对象，那么它们为分类关系。

（4）定义码

通过引入交叉实体除去上一阶段产生的非确定关系，然后从非交叉实体和独立实体开始标识侯选码属性，以便唯一识别每个实体的实例，再从侯选码中确定主码。为了确定主码和关系的有效性，通过非空规则和非多值规则来保证，即一个实体实例的一个属性不能是空值，也不能在同一个时刻有一个以上的值。找出误认的确定关系，将实体进一步分解，最后构造出IDEF1X模型的键基视图（KB图）。

（5）定义属性  
    从源数据表中抽取说明性的名词开发出属性表，确定属性的所有者。定义非主码属性，检查属性的非空及非多值规则。此外，还要检查完全依赖函数规则和非传递依赖规则，保证一个非主码属性必须依赖于主码、整个主码、仅仅是主码。以此得到了至少符合关系理论第三范式的改进的IDEF1X模型的全属性视图。

（6）第五步——定义其他对象和规则

    定义属性的数据类型、长度、精度、非空、缺省值、约束规则等。定义触发器、存储过程、视图、角色、同义词、序列等对象信息。

建模工具可以根据这些规则自动生成物理数据库中更新、插入和删除的触发器。

IDEF1x和传统的E-R方法相比，具有很多的优点，主要表现在：（1）IDEF1X模型语义更为丰富和精细，可充分而清楚地表达企业的复杂数据信息及其业务规则；（2）IDEF1X模型具有更强的一致性和更高的规范化程度；（3）IDEF1X定义的逻辑模型更利于向物理模型转换。IDEF1X定义的符合第三范式的逻辑模型已表达出了企业的数据信息和业务规则，可直接向物理模型转换。

    IDEF1x方法中建模的概念和E-R方法的概念基本类似，但拥有更丰富的语义和规则、更加完善的语法、增强的图形表达能力、规范的开发过程、规范的文件格式以及大量软件建模工具的支持。而且IDEF1x模型能自动实现键的迁移，从父实体到子实体的关键字迁移是强制的。IDEF1x方法中的1:1和1:N联系有明确的HAS\_A语义，分类联系有IS\_A语义，和面向对象中聚合和继承的概念相对应，所以将来可以方便地将面向对象的概念引入IDEF1x建模方法。IDEF1x方法中的非确定联系必须引入交叉实体，转化为两个1:N联系，可以直接向物理模型转换。IDEF1x中的域的概念也在大多数物理数据库中得到实现。所有这些，在传统E-R建模方法中都是难以实现的。

使用IDEF1x语义建模方法对信息系统进行数据建模，并用建模工具对其进行需求、逻辑和物理设计，充分地保证了数据的一致性和完整性。并且能够实现将数据库的分析、概念模型设计和物理数据库结构设计有机结合起来，大大地提高了系统的开发效率。

目前，IDEF1x方法在国内外信息系统建模领域特别在CIMS大型、复杂的综合网络信息系统中得到越来越广泛的应用。IDEF1x方法还处于不断的发展过程中，未来的研究方向是将面向对象的建模方法引入IDEF1x的模型，使其更易于复杂对象的建模和类型的扩展。另外我们还需要为IDEF1x方法提供如同关系模型一样强大的理论依据。