

# 第五章 数据流模型分析与设计

饶元

2019.10

# 内容

- 与数据建模的关系
- 数据流模型的设计理念
- 数据流模型的具体方法

# 课堂设计实例

个人基本情况									
姓 名	林碧兰	性 别	女						
民 族	汉族	出生年月	1991-11-2						
政治面貌	中共党员	身体状况	良好						
参加工作 时 间	2014 年 7 月	教 龄	0 年						
学 历	本科	学 位	学士学位						
毕业时间	2014 年 7 月	毕业学校 及专业	广州师范大学						
现任岗位	专业教师		任职时间	2014 年					
职 称	讲师		职称取得时间	2008 年					
技能水平	工程师		技能证取得时间	2010 年					
学习工作经历									
起止时间	所在学校/单位		职务						
2001-07~2004-07	国际旅游商务学校		教师						
2004-03~2004-07	黑龙江农业工程职业学院		兼职任课教师						
2004-07~2011-07	黑龙江旅游商务学校		园林系系主任						
2012-07~2014-07	哈尔滨现代艺术设计职业技术学院		兼职任课教师						
<div>+</div> <th colspan="5">任教情况</th>					任教情况				
课程名称		授课起止日期		授课对象		学期课时			
《园林工程施工图绘制》		2014 年 5 月—6 月		15 秋园林		64			

基本信息表

用户ID
用户姓名
性别
民族
出生年月
政治面貌
身体状况
参加工作时间
教龄
最高学历
最高学位
毕业时间
毕业学校
技能水平
技能证取得时间
现任岗位
任值时间
职称
职称取得时间

学习工作经历

学习工作ID
开始时间
结束时间
所在学校/单位
职务
用户ID (FK)

任教经历表

任教经历ID
课程名称
授课开始时间
授课结束时间
授课对象
学期课时
用户ID (FK)

学历

学历ID
学历名称
学位名称
毕业时间
毕业学校
用户ID (FK)

能力评价表

能力评价ID
技能水平
技能证取得时间
现任岗位
任值时间
职称
职称取得时间
用户ID (FK)

基本信息表

用户ID
用户姓名
性别
民族
出生年月
政治面貌
身体状况
参加工作时间
教龄

学习工作经历

学习工作ID
开始时间
结束时间
所在学校/单位
职务
用户ID (FK)

任教经历表

任教经历ID
课程名称
授课开始时间
授课结束时间
授课对象
学期课时
用户ID (FK)

# 提出的问题

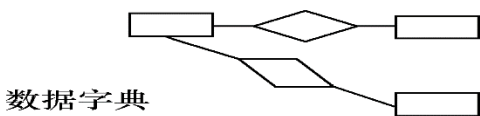
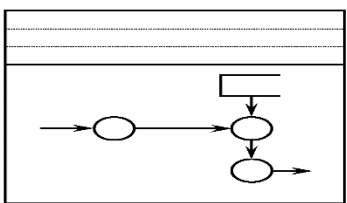
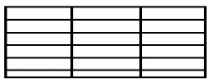
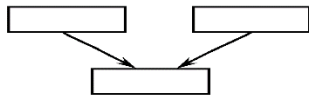
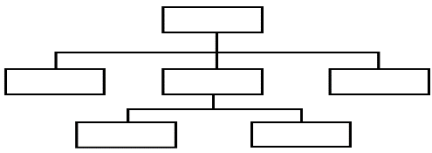
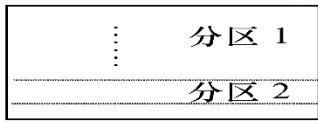
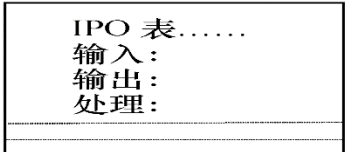
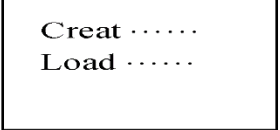
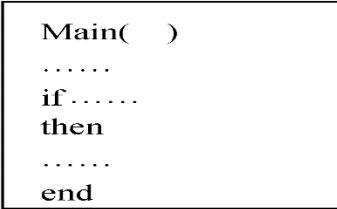
- 1、如何来分析和确定系统中存在的功能呢？
- 2、如何建立数据与功能之间的联系呢？
- 3、如何通过对数据的建模以及对数据的操作来实现相应的功能呢？

# 构建数据-地点（角色）的CRUD矩阵

- 在每一个地点需要实体和属性的哪些子集来完成工作；
- 需求什么级别的访问？
- 可以创建实体实例吗？
- 可以读取实体实例吗？
- 可以删除实体实例吗？
- 可以修改吗？

[illegible]

# 数据设计和处理设计同时进行

设计阶段	设计描述	
	数 据	处 理
需求分析	数据字典、全系统中数据项、数据流、数据存储的描述	数据流图和判定表（判定树）、数据字典中处理过程的描述
概念结构设计	概念模型（E-R图） 	系统说明书包括： ① 新系统要求、方案和概图 ② 反映新系统信息流的数据流图 
逻辑结构设计	某种数据模型 关系  非关系 	系统结构图 （模块结构） 
物理设计	存储安排 方法选择 存取路径建立 	模块设计 IPO 表 
数据库实施阶段	编写模式 装入数据 数据库试运行 	程序编码、编译联结、测试 
数据库运行和维护	性能监测、转储 / 恢复 数据库重组和重构	新旧系统转换、运行、维护（修正性、适应性、改善性维护）

# DFD建模

- 数据流分析（Data Flow Analysis，简称DFA）方法是一种以数据流技术为基础的、自顶向下、逐步求精的系统分析方法。
- 从用户视角下来审视数据被操作与流转的过程。

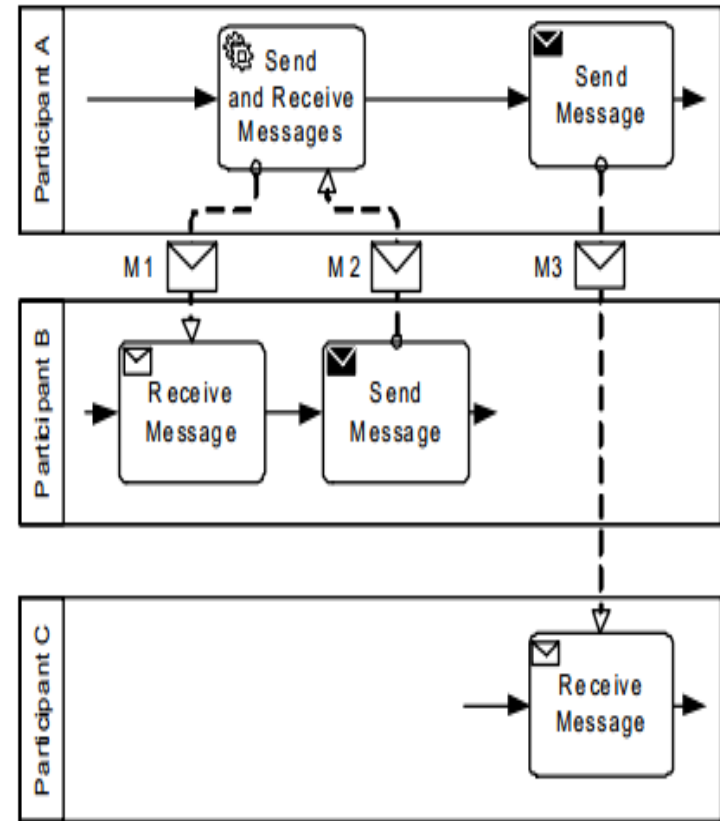


# 内容

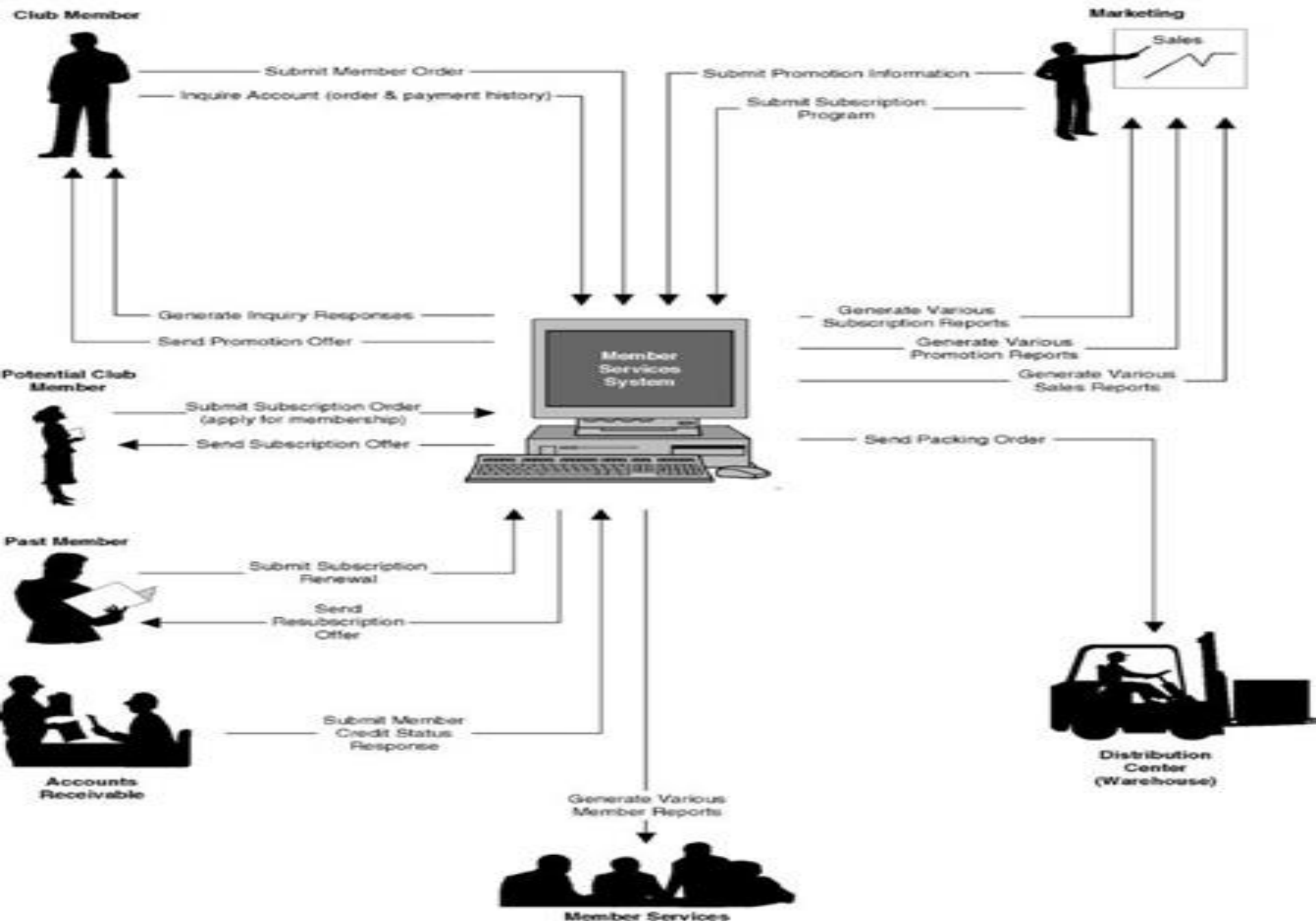
- 与数据建模的关系
- 数据流模型的设计理念
- 数据流模型的具体方法

# 1、数据流图 (DFD: Data Flow Diagram)

- 数据流图是一种描述数据以及对数据进行加工转换的图形工具
- 表达数据在系统内部的**逻辑流向**以及系统的**逻辑功能**和数据的**逻辑变换**
- 数据流分析: 就是用户视图的流动。

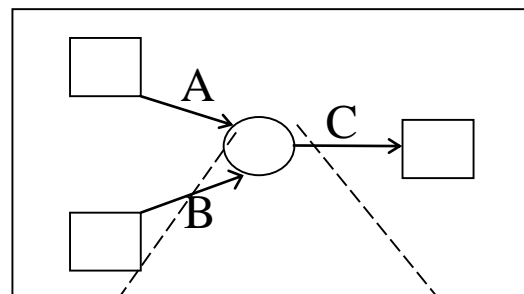


# Member Services Context Diagram

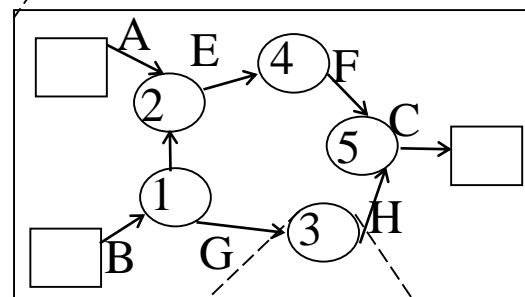


# 数据流图绘制过程示意图

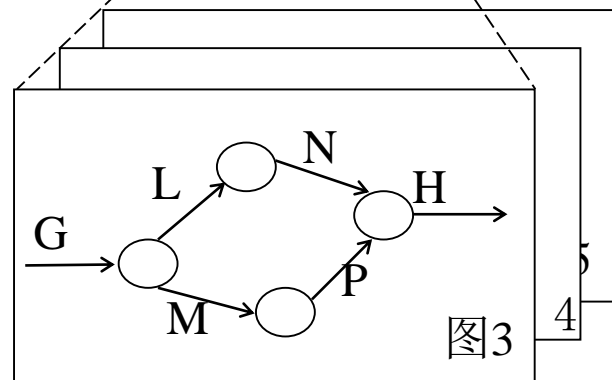
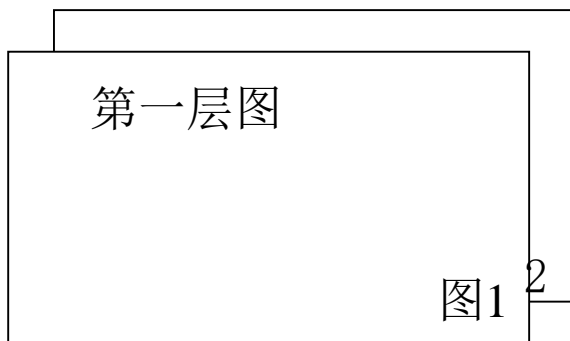
[a] 关联图



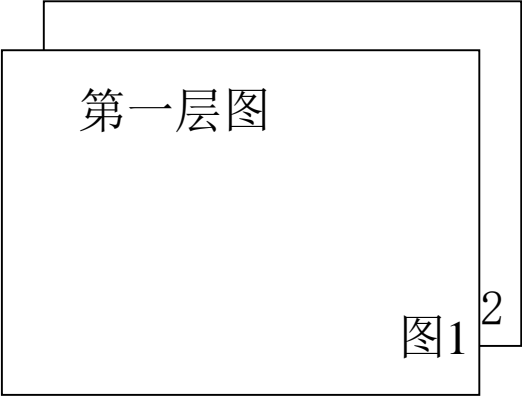
[b] 顶层图



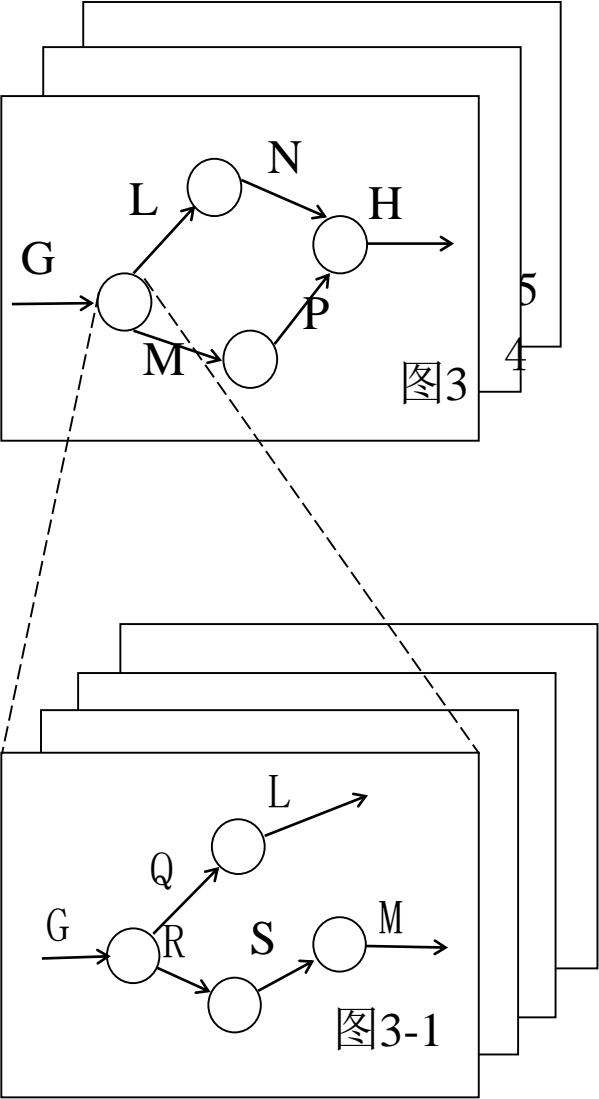
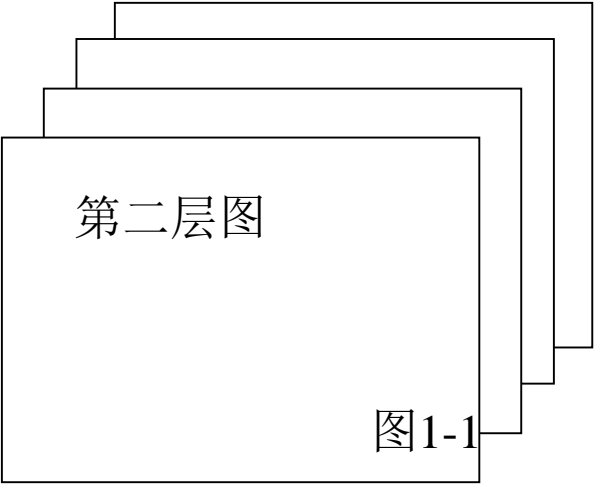
[c]



[c]

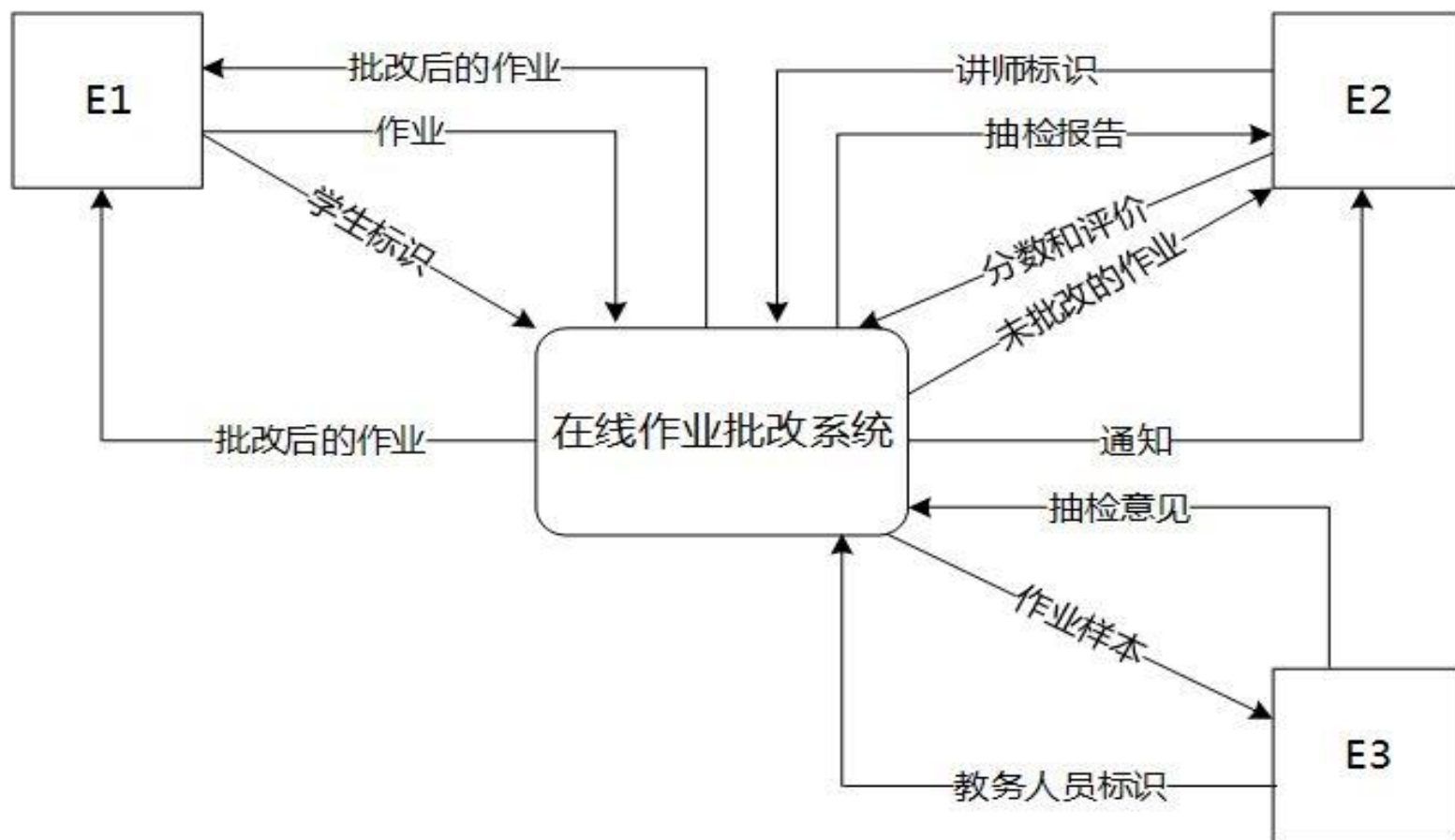


[d]

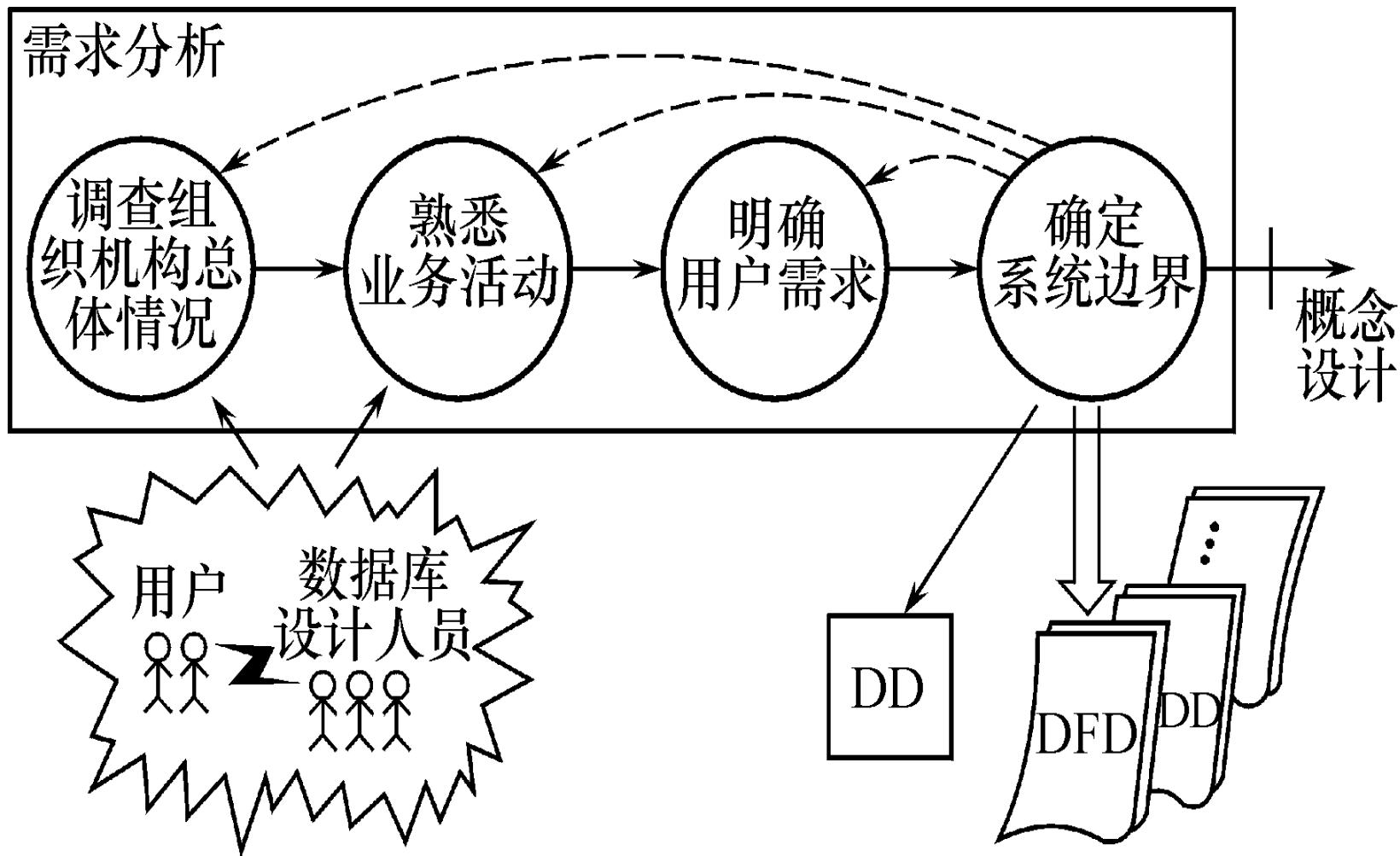


# DFD设计思想 (1)

- 由外而内： 即外部的需求是所有工作的起点
- 系统边界的确定： 来自于外部用户的需求确定



# 注意：需求分析过程



# 示例：比较系统的复杂度



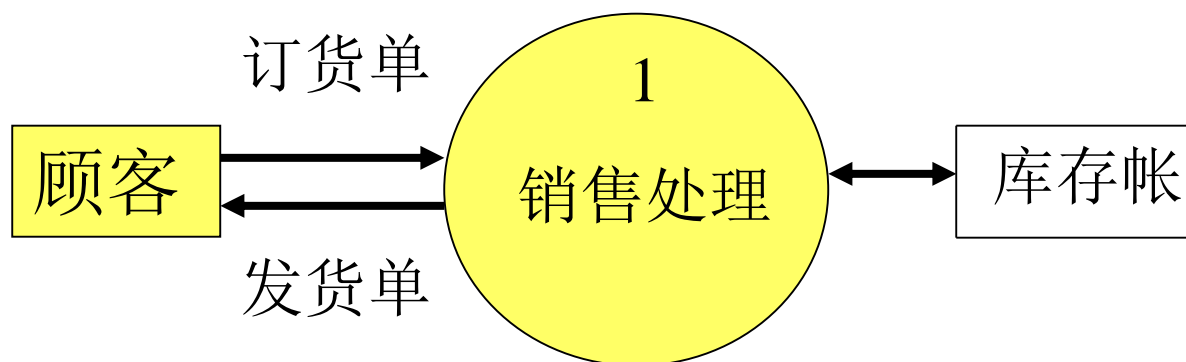


## 示例1:

销售部门接到顾客送来的订货单后， 根据库存情况向用户发货

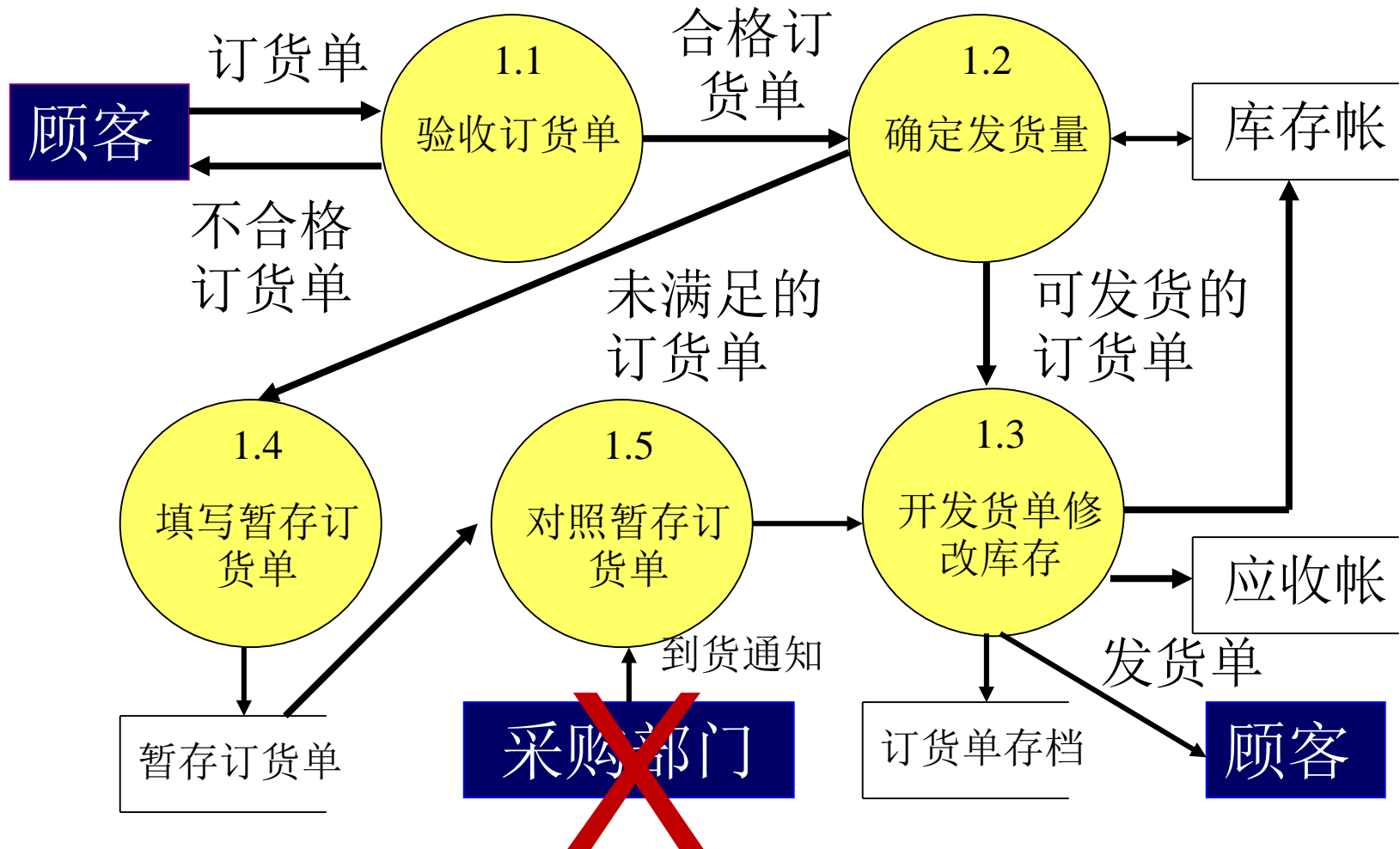
- 订货单处理
- 确定发货量
- 开发货单及其相关处理
- 填写暂存订货单
- 对照暂存订货单

### 订货处理—第1层DFD



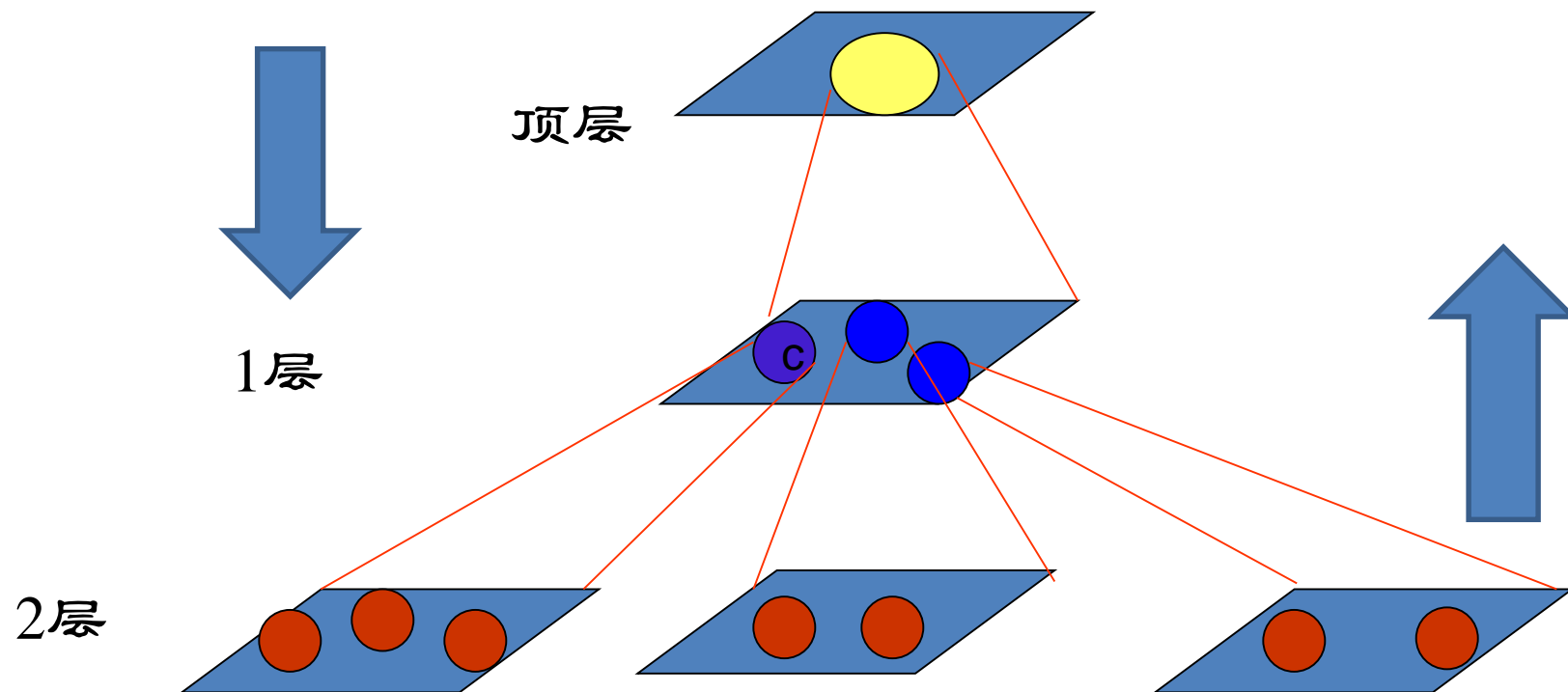
## 订货处理—第2层DFD

请问目前的设计存在什么问题？

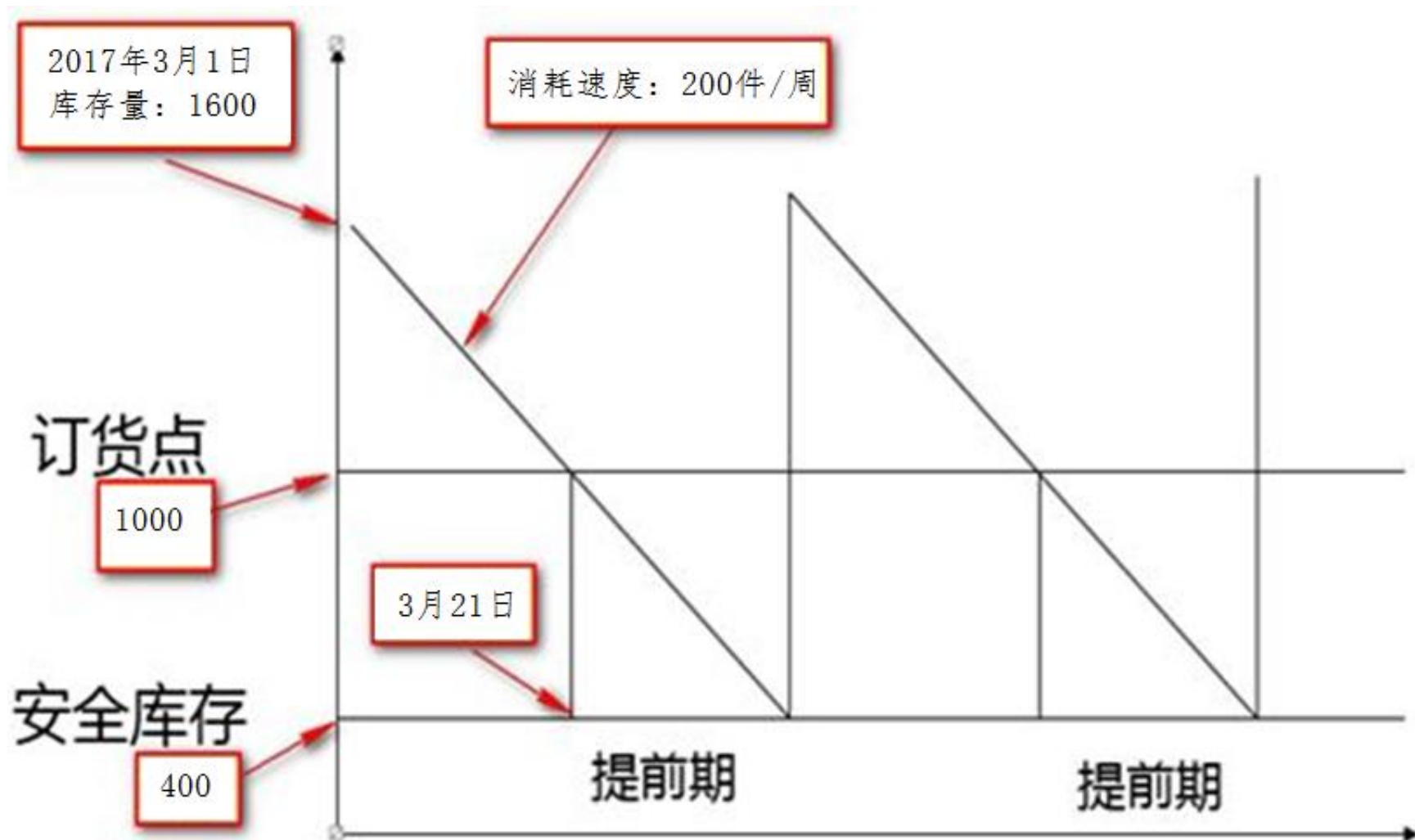


## DFD设计思想 (2)

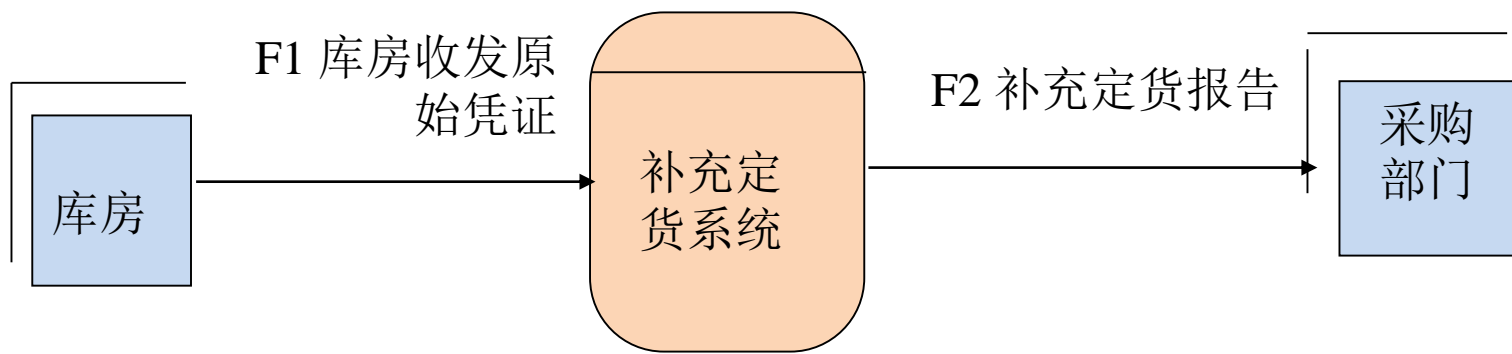
- DFD的核心特征是：自顶向下，逐步求精
- 核心：“分解”和“抽象”



## ➤ 例2： 补充定货系统

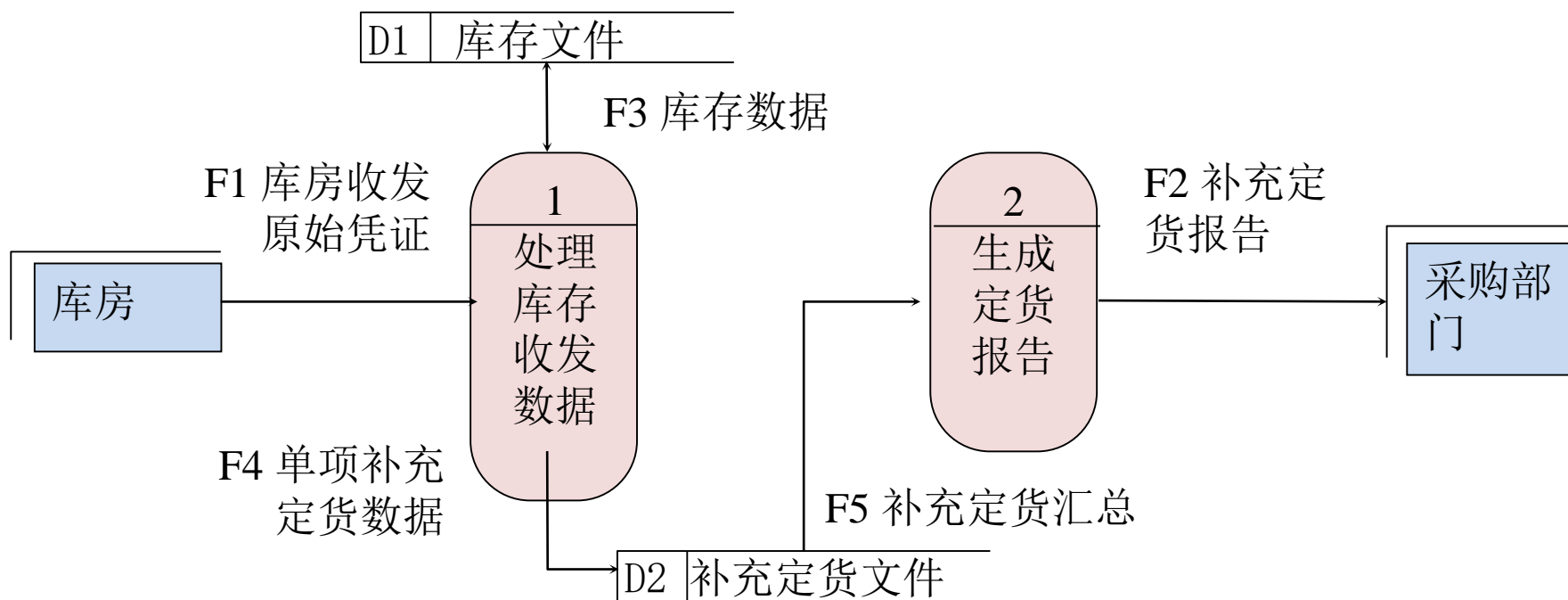


## ➤ 例2： 补充定货系统

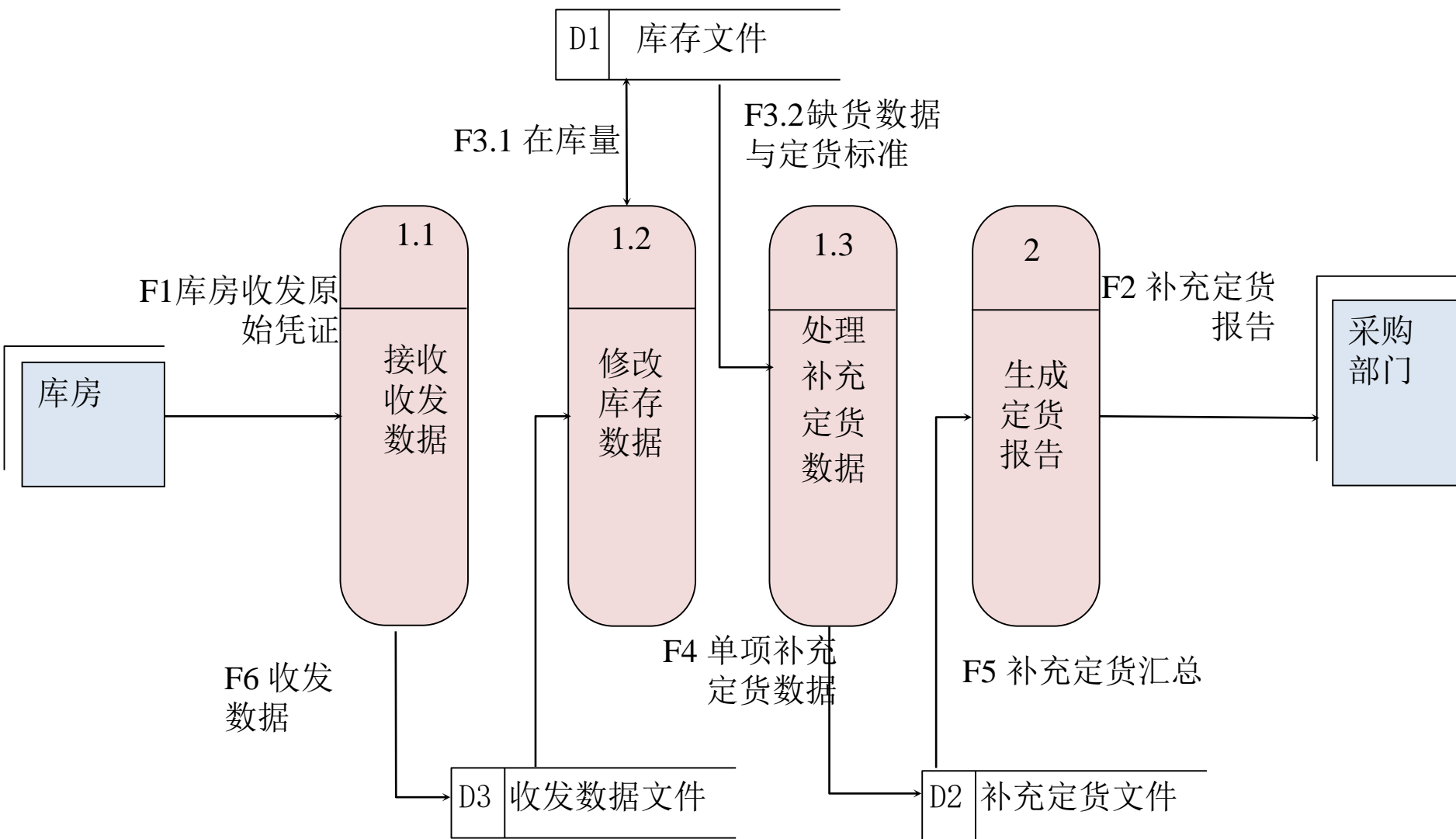


补充定货系统关联图

# 补充定货系统的顶层图

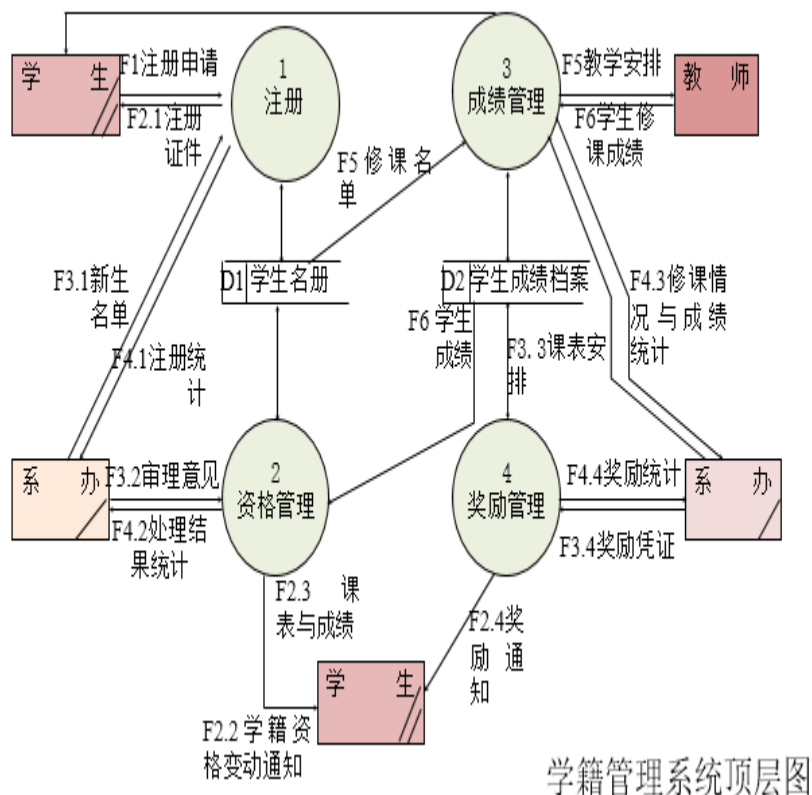


# 补充定货系统第一层数据流图



# DFD设计思想 (3)

- DFD数据流图与数据字典相结合
- 核心：相互补充和具体实现



学籍管理系统顶层图

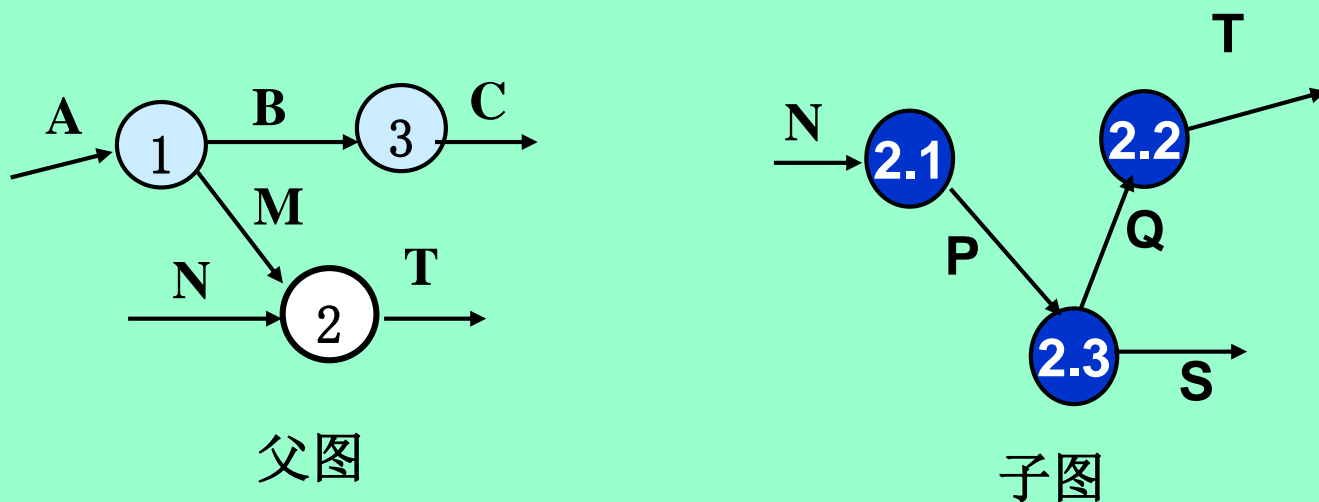
数 据 流					
系统名：学籍管理		编号：			
条目名： 学生成绩通知		别名： 成绩通知单			
来源：成绩管理		去处：学生			
数据流结构： 学生成绩通知：{学号+学生姓名+{课程名称+成绩}该生本期所修课程 +（补考课程名称+补考时间+补考地点）}所有在册学生					
简要说明： 学生成绩通知在每学期期末考试结束后一周至下学期开学前一 周期间内发给所有本期在校学生。					
修改纪录：		编写	张XX	日期	2017年4月20日
		审核	李XX	日期	2017年4月20日

数据词典《数据流》条目举例



## DFD设计思想 (4)

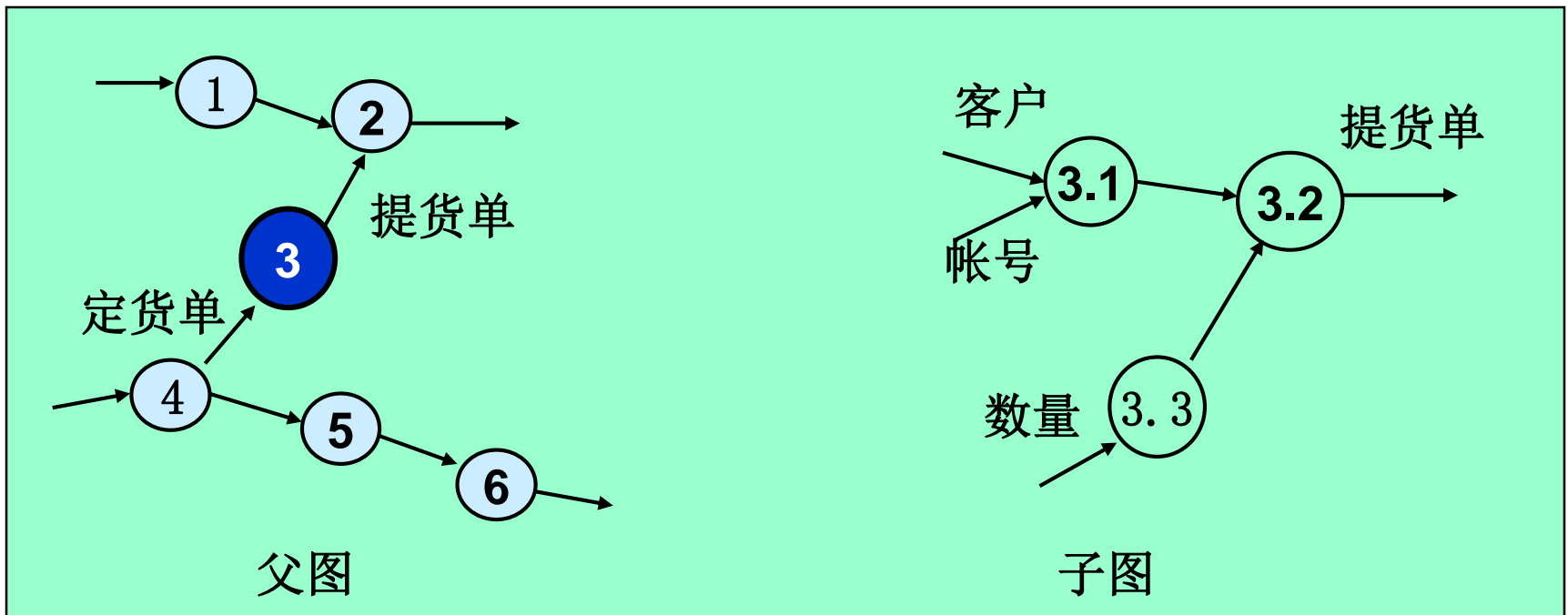
- DFD数据流上下层之间的流入与流出数据一致
- 核心：平衡与一致



信息的封装与隐藏

## DFD设计思想 (4)

- 考察下图中父图与子图是否平衡？



如果有如下数据条目： 订货单=客户+帐号+数量，  
则子、父图平衡

# 数据流图的建立—建立原则

- 数据流图的建立过程必须遵循自顶向下、逐层分解的原则
- 数据流图只反映系统逻辑功能
- 分层的数据流图总是由顶层、中间层和底层组成的（或：上下文图+0级图+n级图）：
  - ◆ 顶层数据流图确定了系统的**边界**
  - ◆ 中层图描述了某个处理过程的**分解**，而组成部分要进一步被分解
  - ◆ 底层图描述的是无须分解的基本处理过程
  - ◆ 每个加工至少应有一个**输入数据流**（反映被处理数据的来源）和一个**输出流**（反映加工的结果）

# 数据流图的建立—建立原则

建立原则：

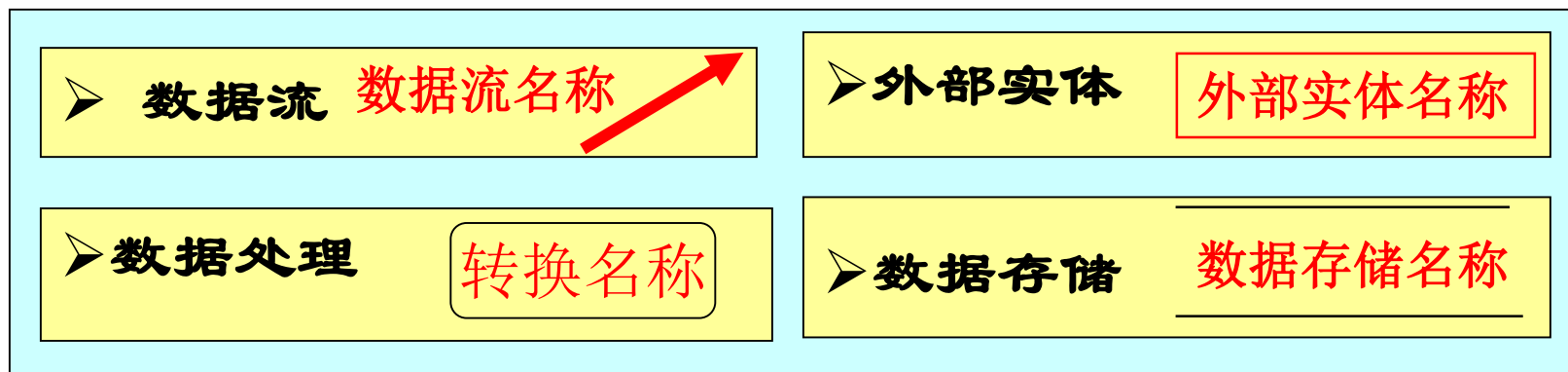
- 数据流图的建立过程必须遵循自顶向下、逐层分解的原则
- 数据流图只反映系统逻辑功能
- 分层的数据流图总是由顶层、中间层和底层组成的（或：上下文图+0级图+n级图）：
  - ◆ 顶层数据流图确定了系统的边界
  - ◆ 中层图描述了某个处理过程的分解，而组成部分要进一步被分解
  - ◆ 底层图描述的是无须分解的基本处理过程
  - ◆ 每个加工至少应有一个输入数据流（反映被处理数据的来源）和一个输出流（反映加工的结果）

# 内容

- 与数据建模的关系
- 数据流模型的设计理念
- 数据流模型的具体方法

# 数据流图的基本组成

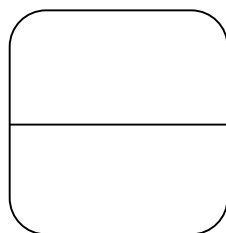
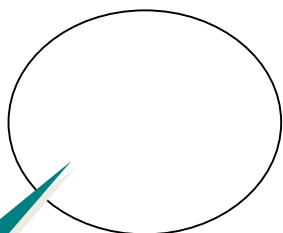
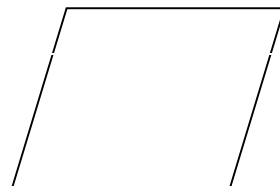
- **外部项(外部实体)**: 外部项在数据流图中表示所描述系统的数据来源和去处的各种实体或工作环节。
- **加工(数据处理)**: 又称数据处理逻辑, 描述系统对信息进行处理逻辑功能。
- **数据存储**: 逻辑意义上的数据存储环节, 即系统信息处理功能需要的、不考虑存储物理介质和技术手段的数据存储环节。
- **数据流**: 与所描述系统信息处理功能有关的各类信息的载体。



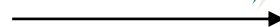
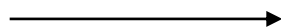
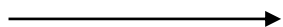
# 数据流图的基本组成

## 数据流图

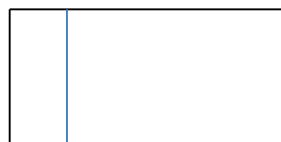
外部项



数据流



处理过程



数据存储

# 数据流图

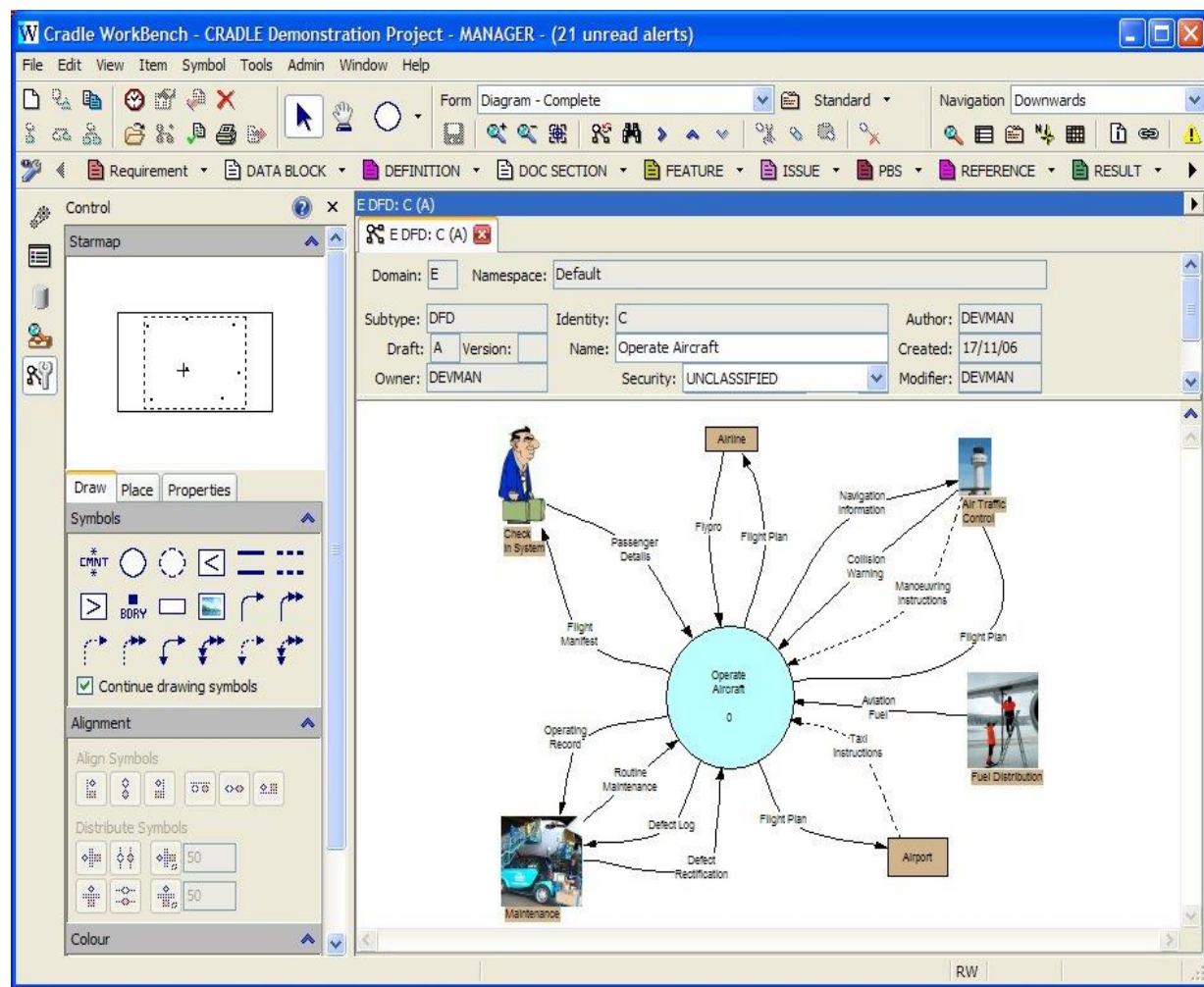


确定系统的外部项，实际上就是确定系统与外界的分界线

## ■ 外部项

(external

entity) 是指不受系统控制的，在系统**以外的人、程序、机构或其他实体**，外部项与系统通过数据交互，表达了该数据的外部来源或去处



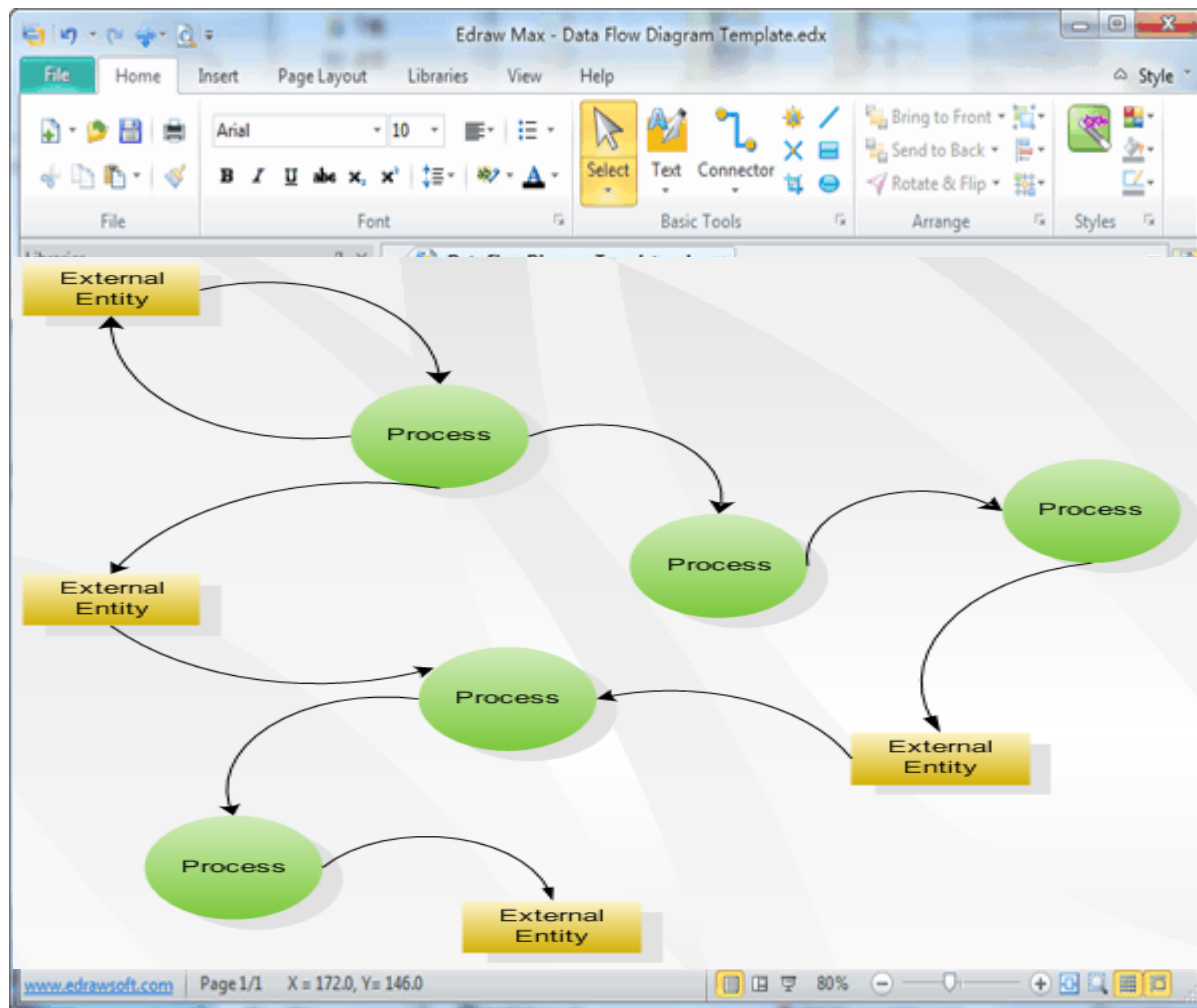


# 数据流图

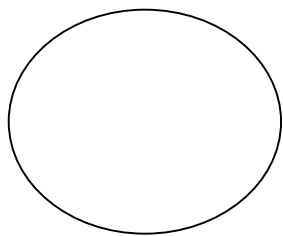
对每一条数据流都要给予简单的描述

## ■ 数据流

(data flow)  
就是一束按特定的方向从源点流到终点的数据，它指出了数据及其流动方向

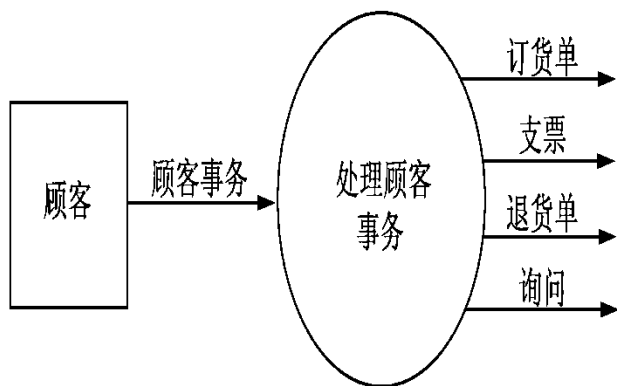


# 数据流图

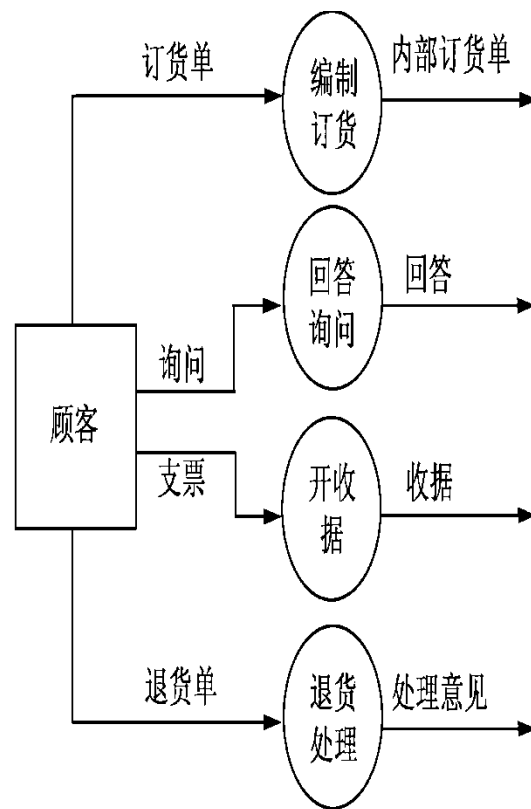


变换数据的结构，如将数据的格式重新排列；  
在原有数据内容基础上产生新的数据内容，如对数据进行  
累计或计算平均值

■ 处理过程  
(process)  
是对数据进行  
变换操作，即  
把流向它的数  
据进行一定的  
变换处理，产  
生出新的数据



(a)

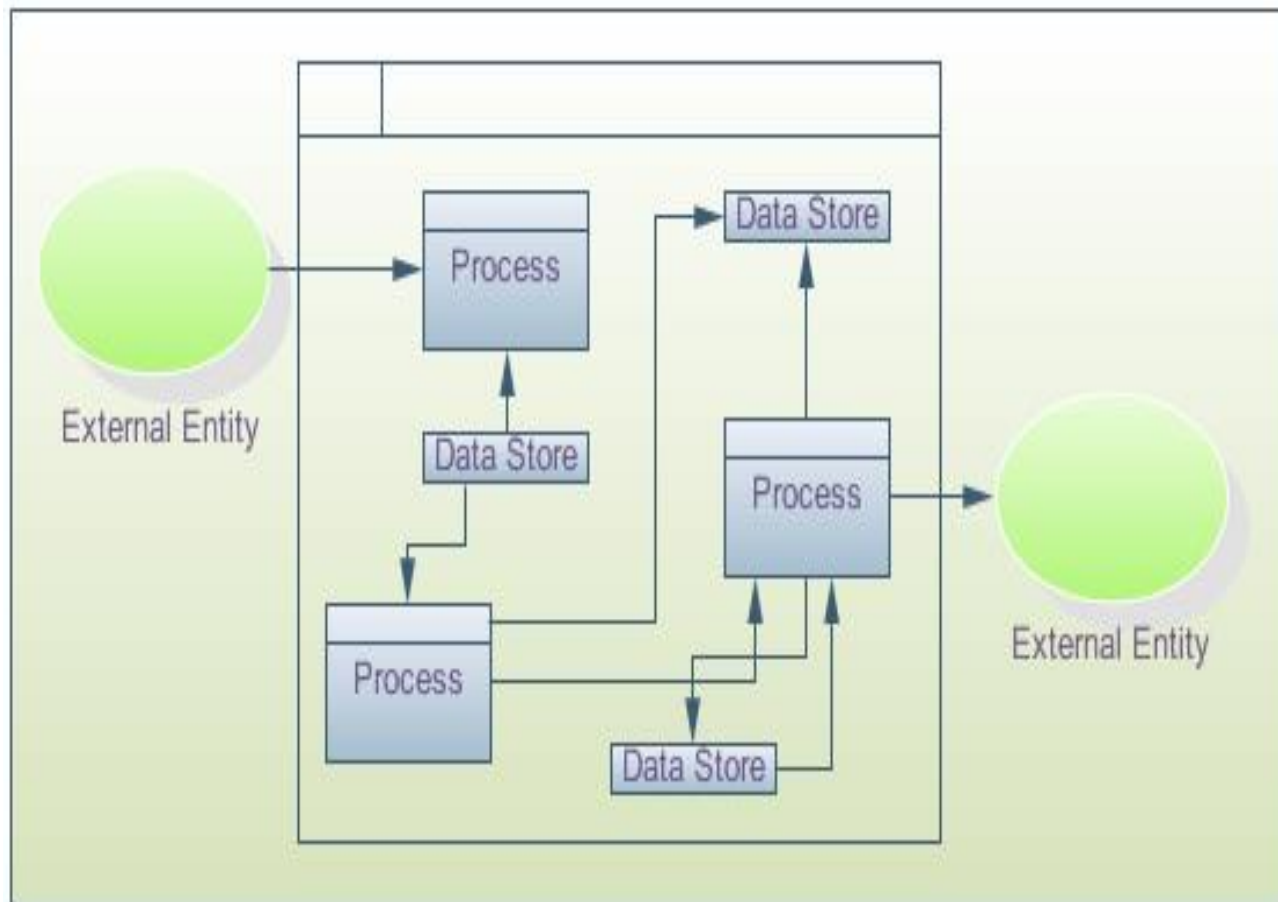


(b)

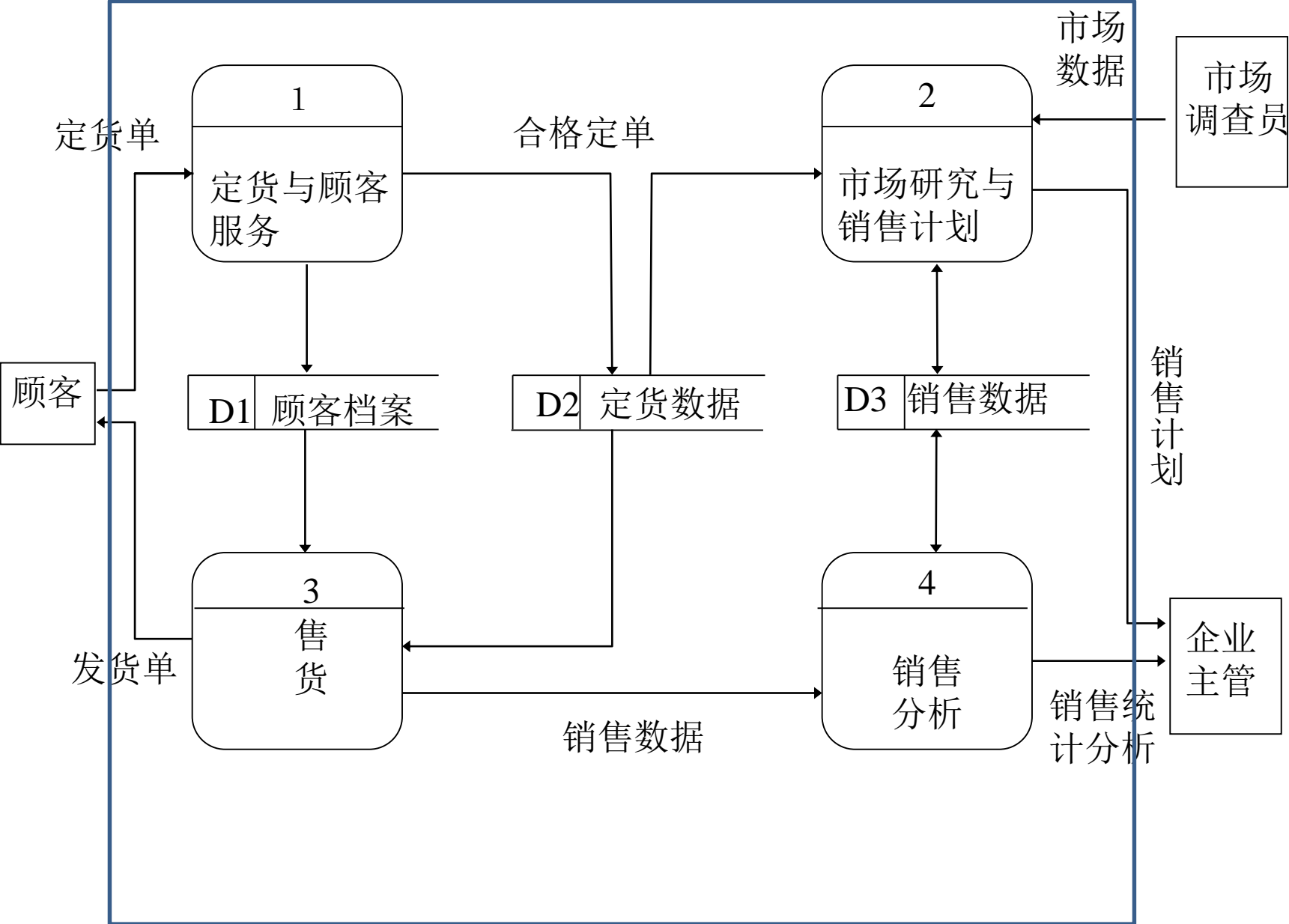
# 数据流图

并不指保存数据的物理地点或物理存储介质，而是数据存储的逻辑描述

■ 数据存储  
(data store)  
指出了数据保存的地方

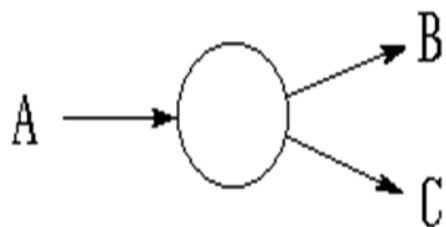


# 示例

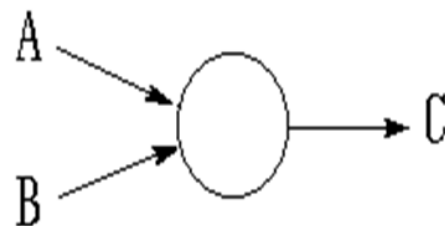


# 数据流与数据处理间存在的基本操作

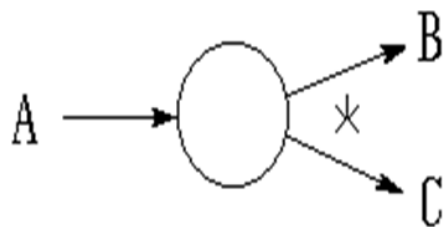
+ “与”，即同时输入/出； \* “或”，即至少一项输入/出  
 $\oplus$  “异或”，即非同时输入/出



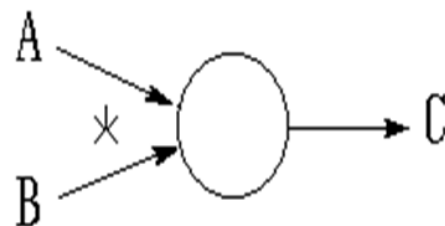
有A则有B或C,  
或两者都有



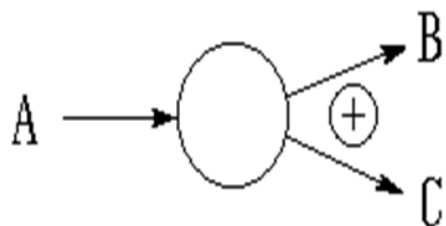
当A或B有一个  
存在, 就有C



有A则有B与C,  
两者同时有



当A与B都存在,  
就有C



有A则有B或C, 但  
不会同时有B与C

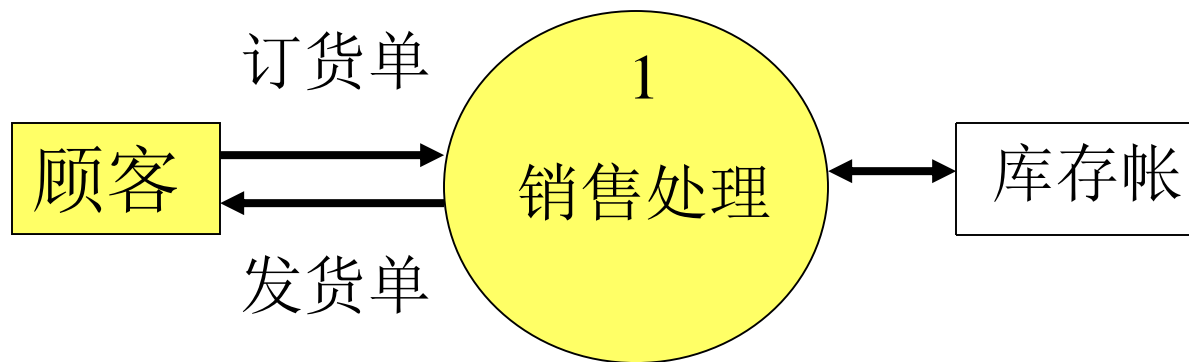
# 数据流图的建立—自顶向下扩展

示例：

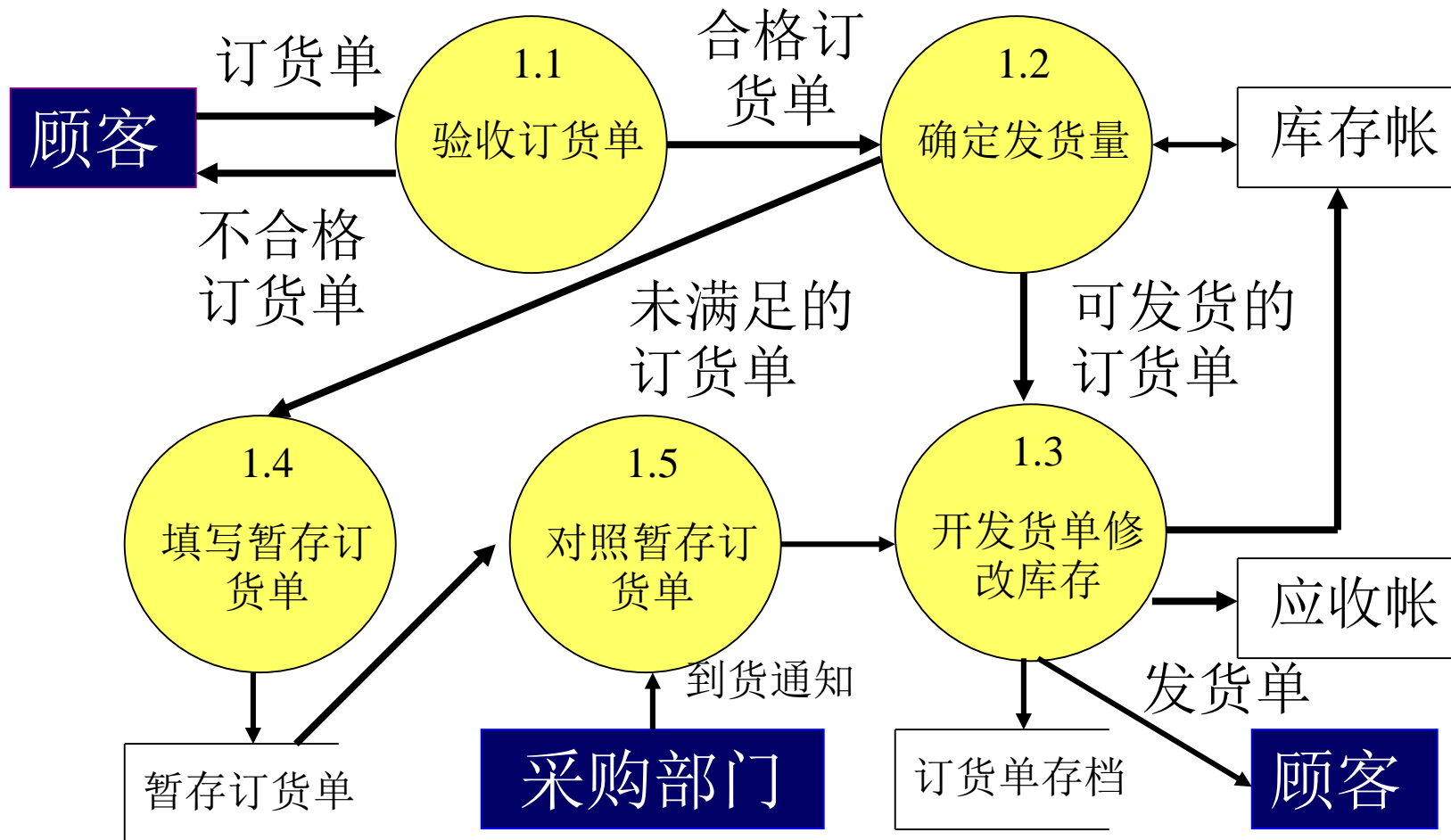
销售部门接到顾客送来的订货单后，根据库存情况  
向用户发货

- 订货单处理
- 确定发货量
- 开发货单及其相关处理
- 填写暂存订货单
- 对照暂存订货单

## 订货处理—第1层DFD



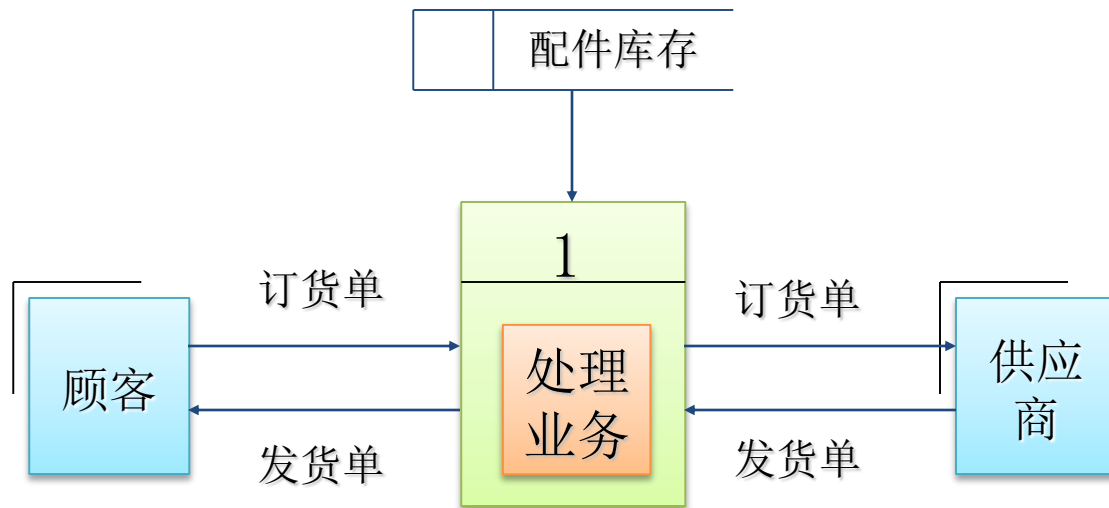
## 订货处理—第2层DFD



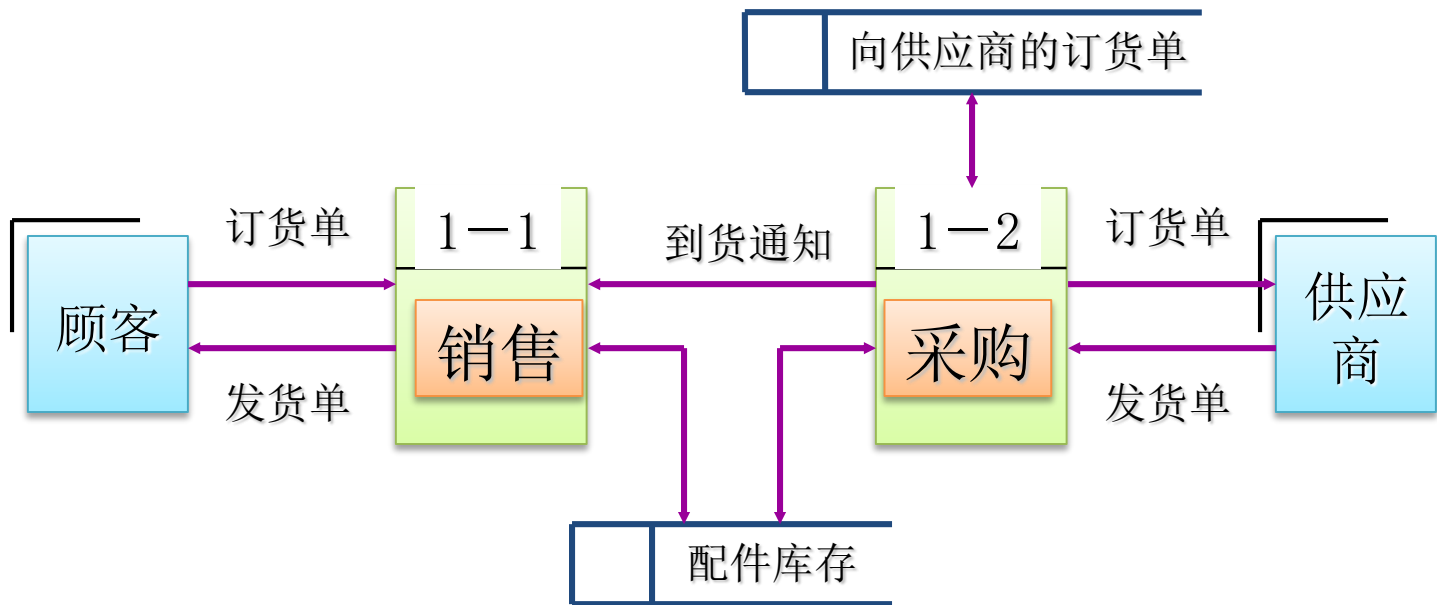


# 例子3:汽车配件公司数据流图建模

- 关联层数据流程图

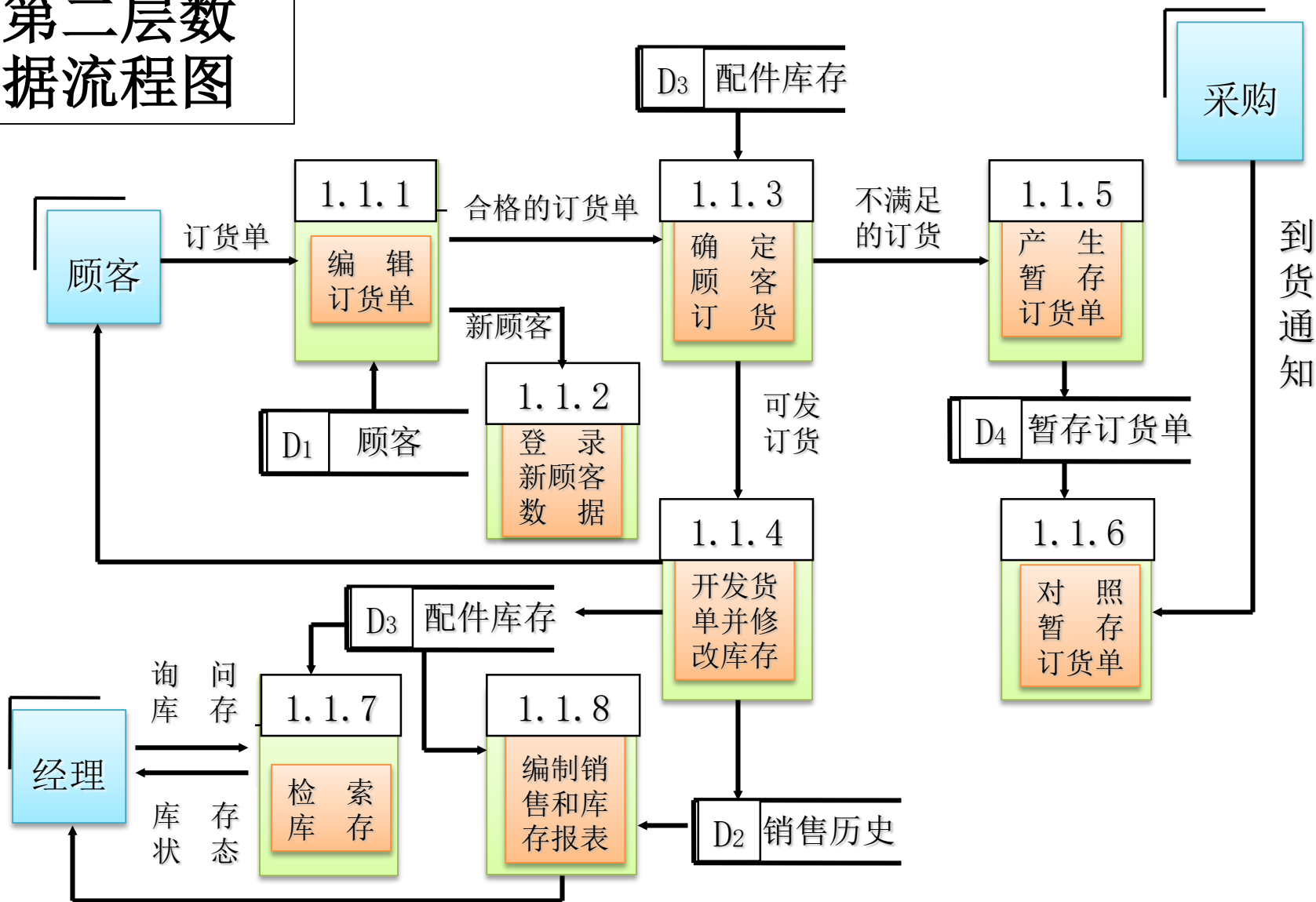


# 例子3: 汽车配件公司数据流图建模

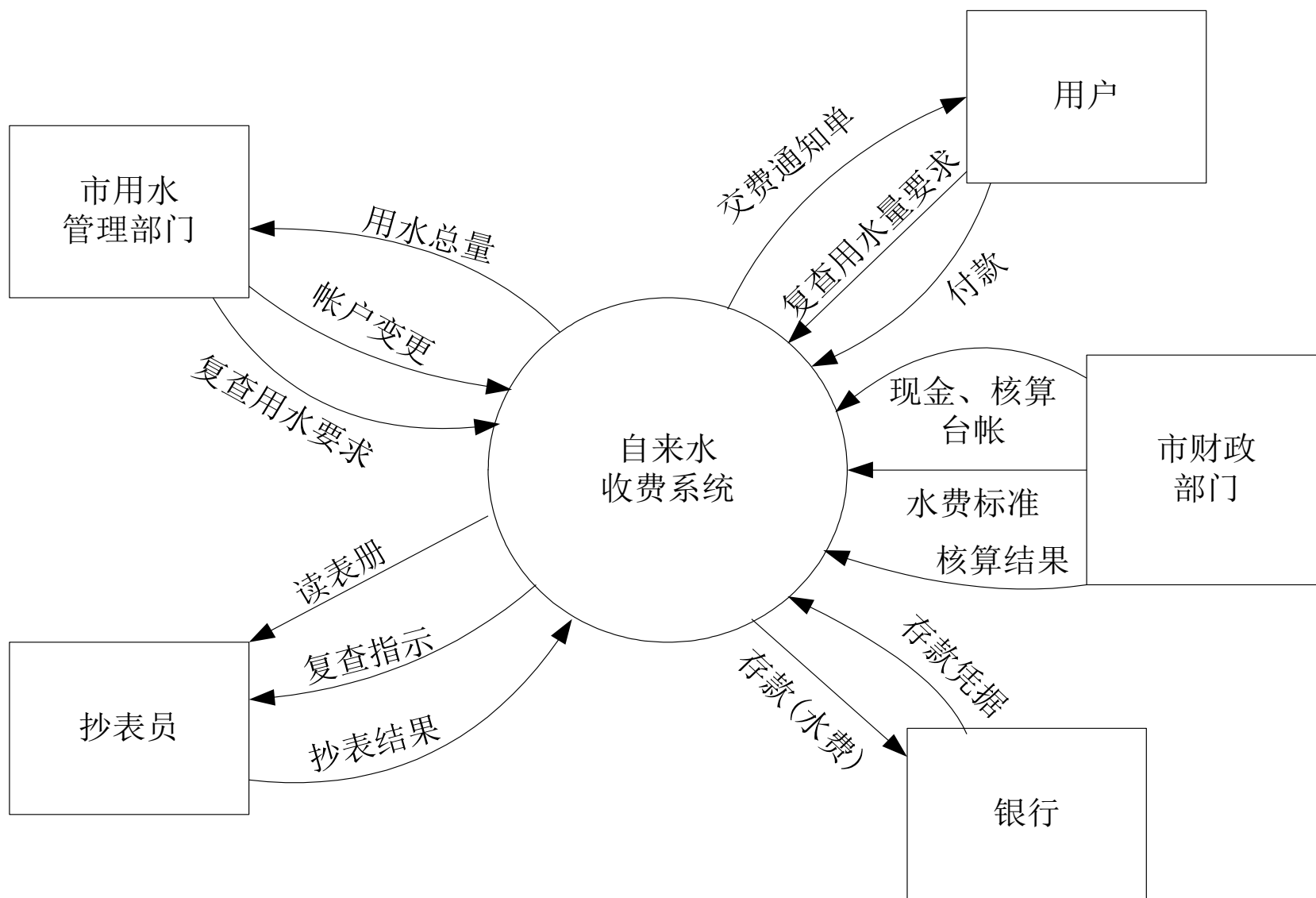


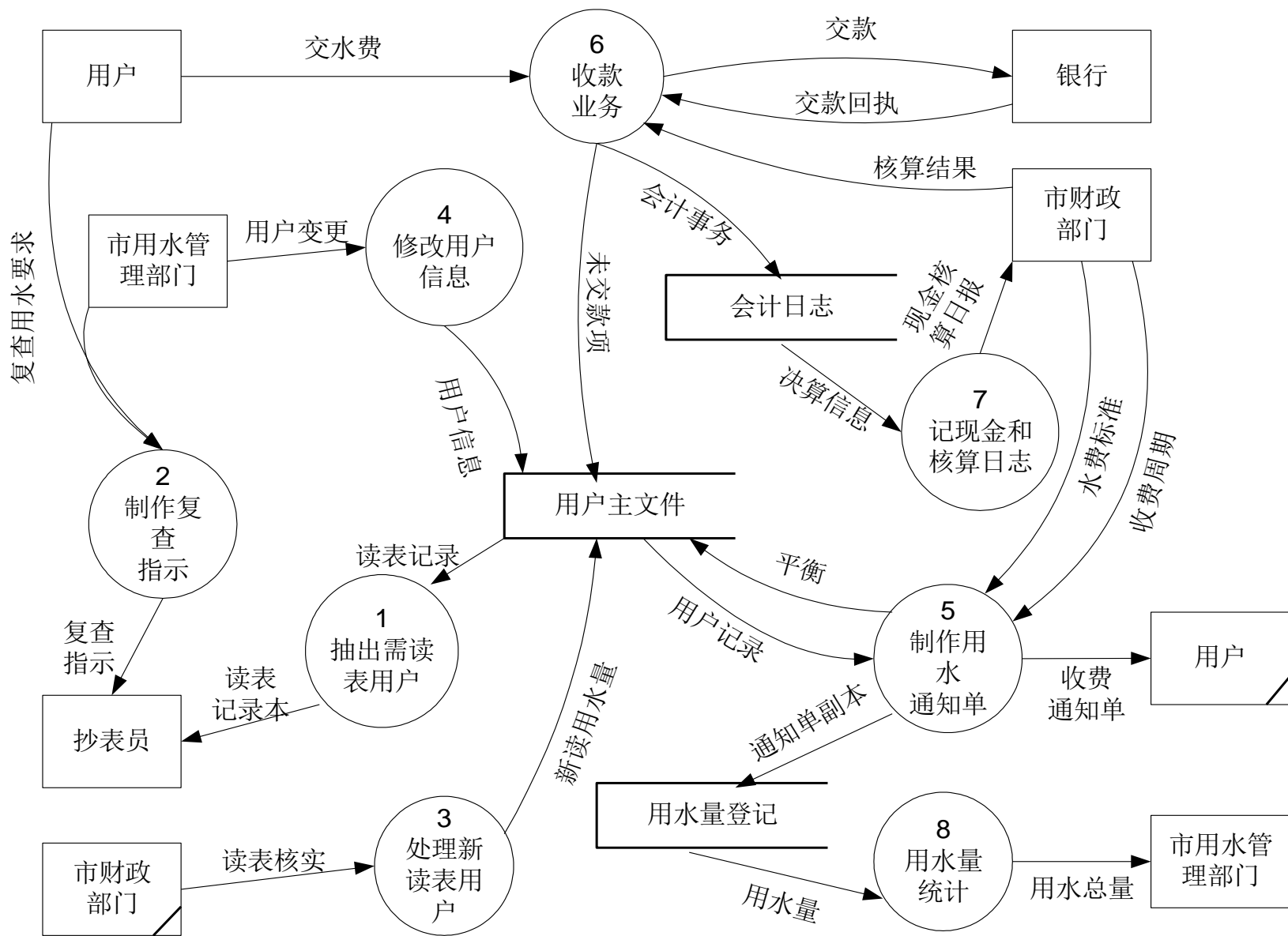
# 例子3: 汽车配件公司数据流图建模

## • 第二层数据流程图



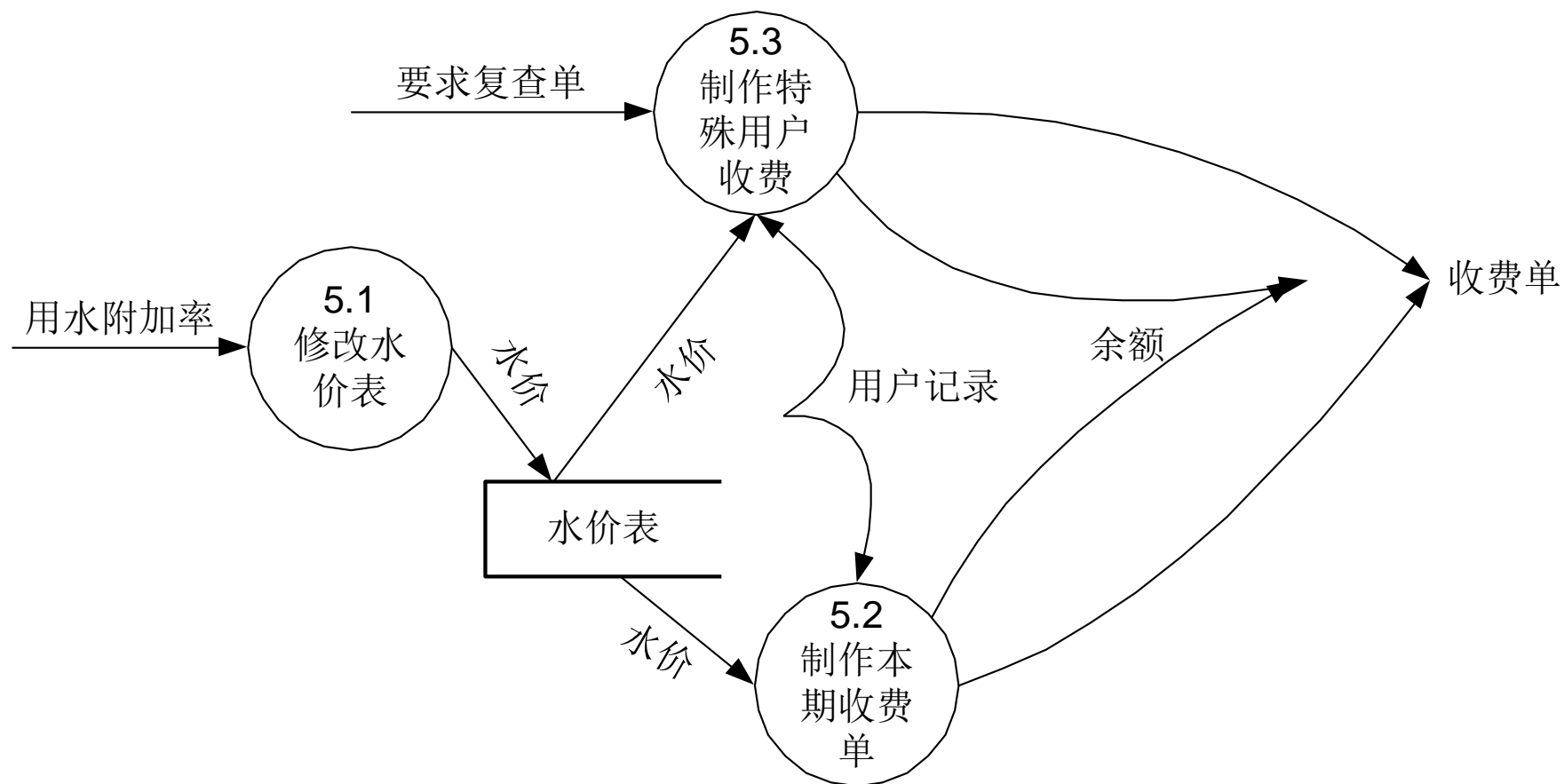
# 数据流图的建立—例4



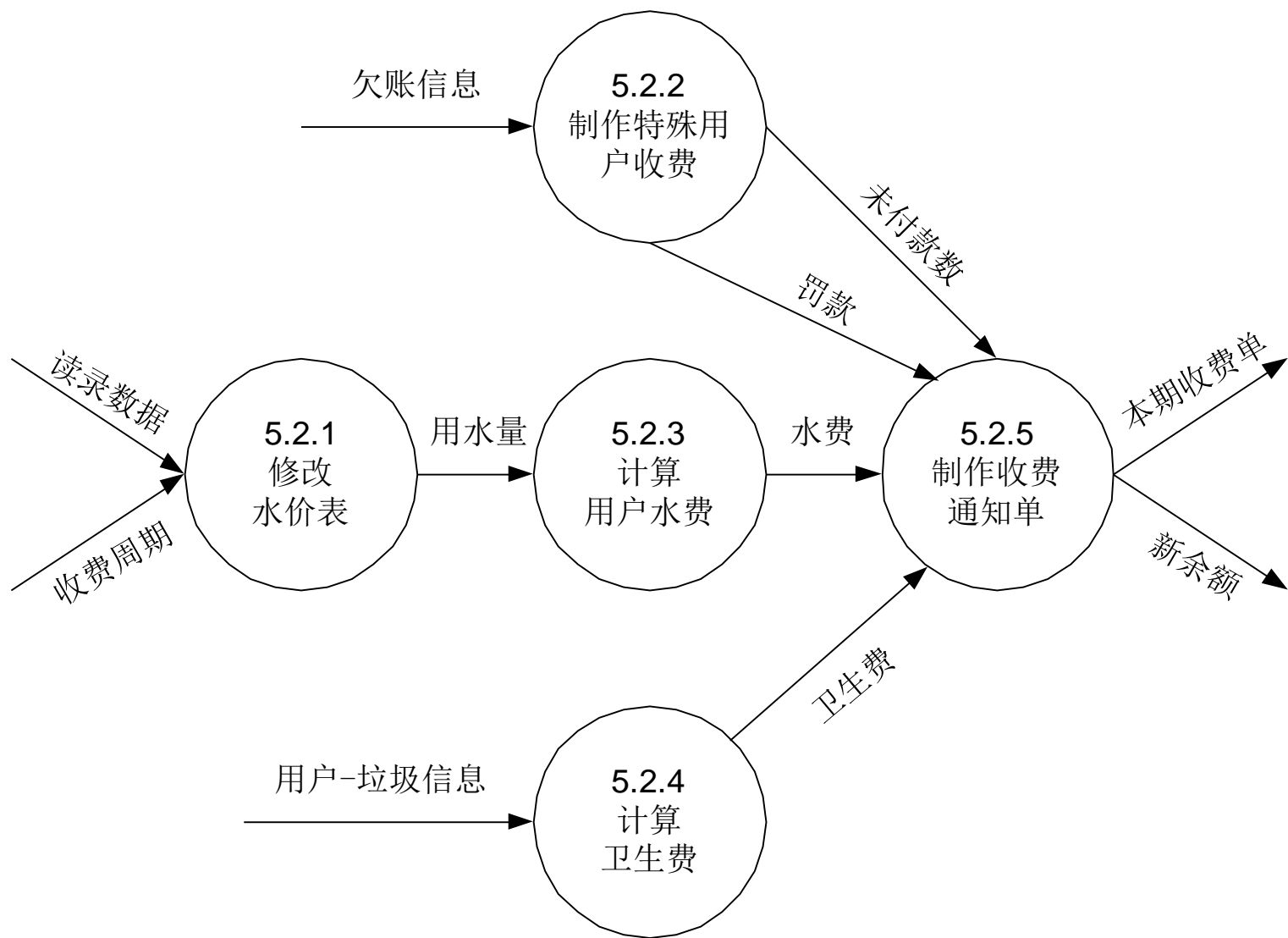


# 数据流图的建立—例子

## 数据流分析技术



# 数据流图的建立—例子



# 数据词典(DD: Data Dictionary)

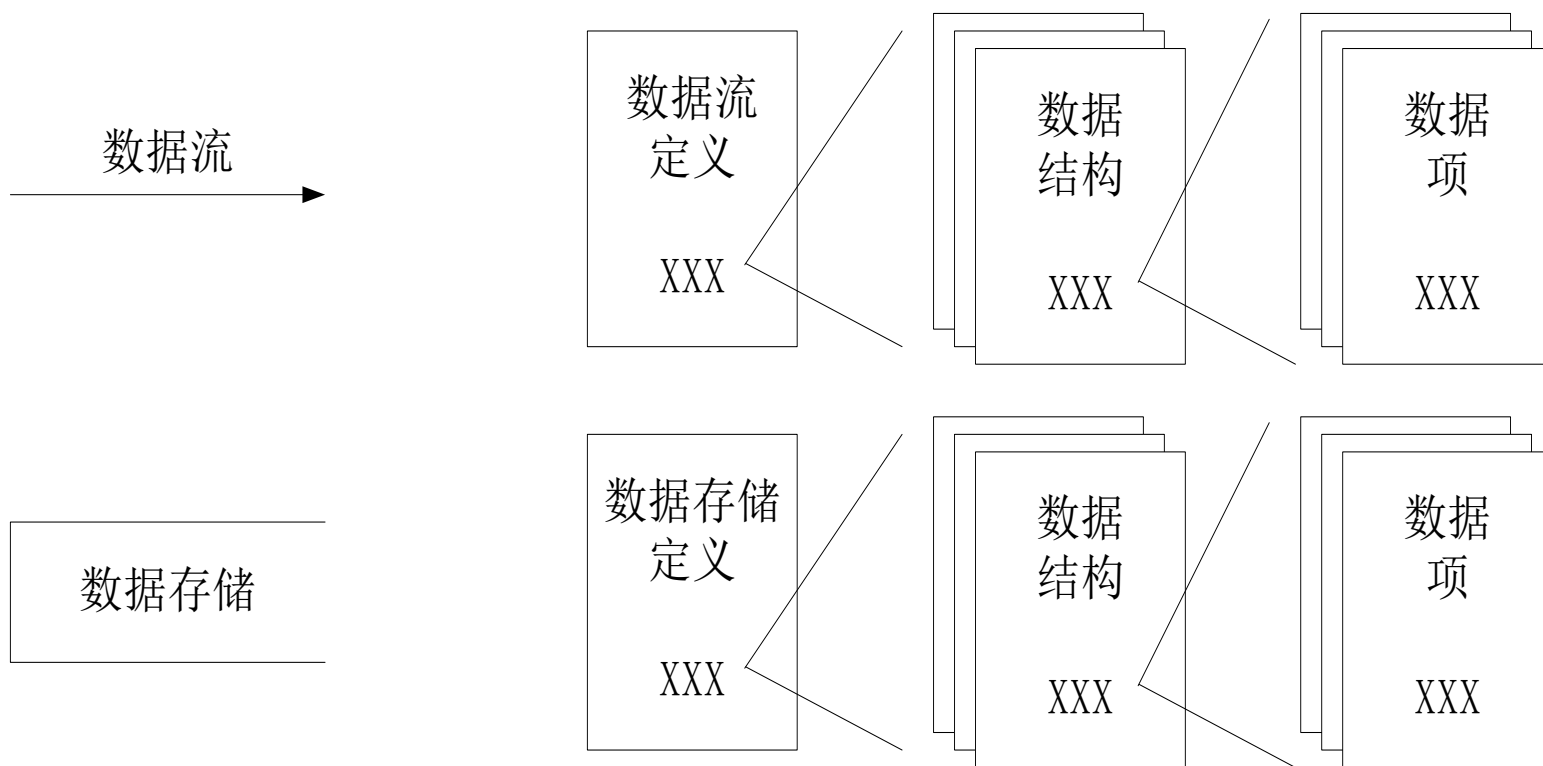
- 仅仅一套数据流图并不能构成系统说明书，只有当图中出现的每一个成分都**给出详细定义之后**，才能较全面地描述一个系统
- 数据流图中所有名字的定义及描述就构成了一本字典，它包括数据流、数据存储、外部项和处理过程的详细条目
- 数据流、数据存储等数据型条目构成数据字典（data dictionary）
- 数据字典是数据流图的辅助资料，对数据流图起注解作用。数据字典主要用于描述数据流和数据存储的逻辑内容，以及外部项和处理过程的某些数据特性



## ■ 数据字典通过以数据项和数据结构的定义来描述数据流、数据存储的逻辑内容

数据流图

数据字典



# 数据字典—数据流

数据流的定义有以下内容：

- 数据流的名称
- 数据流的简述
- 数据流的来源
- 数据流的去向
- 数据流的组成
- 数据流的流通量
- 高峰时的流通量

数据流名称：选课单

简 述：学生填写并交给系里的所选课程列表

数据流来源：学生

数据流去向：“选课统计”过程

数据流组成：学号，课程表  
课程，教师

流 通 量：20份/天

高峰流通量：学期开始时，100份/天

■ 数据结构用来定义数据项之间的组合关系，是对数据的一种逻辑描述

■ 数据结构的定义有以下内容：

- 数据结构的名称
- 数据结构的组成

数据结构名称：课程

简述：用于记录有关课程的基本信息

组成：课程名，学时  
学分， 教材

组织：按课程名顺序存放

有关的数据流/数据结构：教师档案、教材表

有关的处理过程：排定课程表、教学查询

## ■ 数据存储的定义：

- 数据存储的名称以及必要时所给的编号
- 流入/流出的数据流
- 数据存储的组成，即它所包含的数据结构

数据存储名称：学生成绩

编号：D20

简述：记录学生所考各门课程的考试成绩

流入的数据流：“考试成绩单”，来源是“登记成绩单”

处理过程

流出的数据流：“成绩”，去向是“成绩统计”处理过程

数据存储的组成：学号

课程号

成绩

# 数据存储结构规范化

- 信息系统中，仅描述数据项是不够的，更重要的是以最优的方式组织
- 用规范化方法来设计数据存储的结构
  - 力求简化数据存储的数据结构
  - 提高数据的可修改性，完整性和一致性
- 规范化的形式：3NF
- 规范化的作用：减少冗余，避免操作异常
- 数据存取要求特性
  - 可预测性
  - 数据的更新程度
  - 数据立即存取要求

# 数据存取要求分析

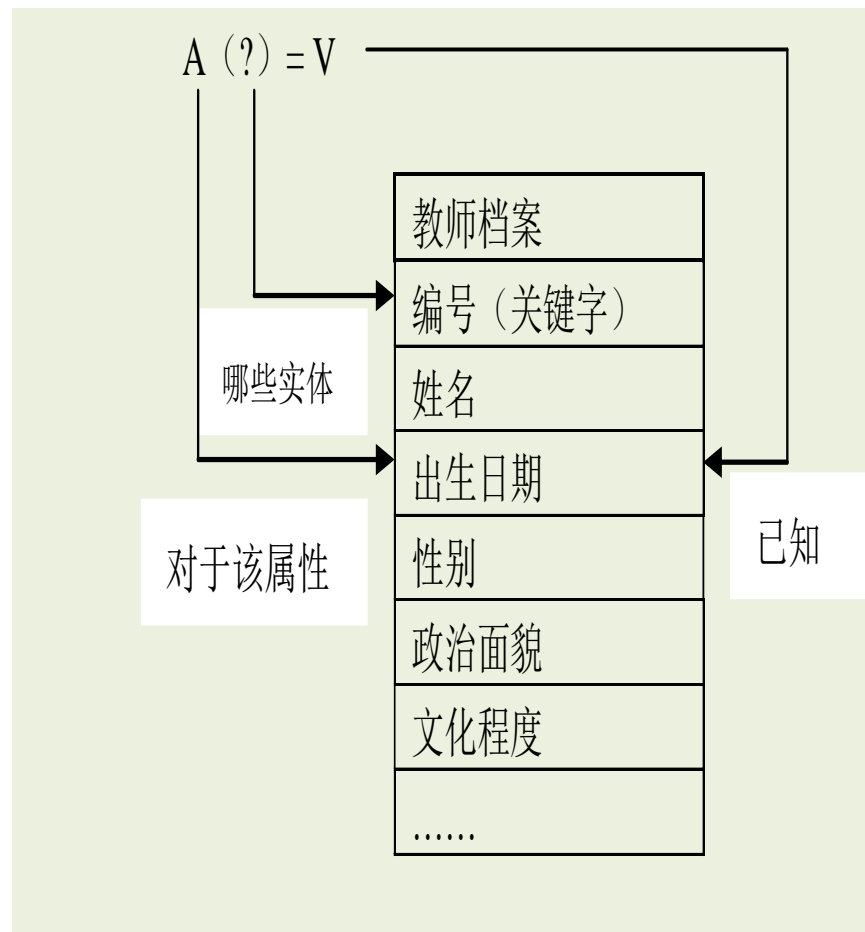
- 对数据存取要求的描述，可以用实体—属性—属性值（E—A—V）这三种符号，以某种表达式或数据存取分析图来表示

- 三种表达式示例：

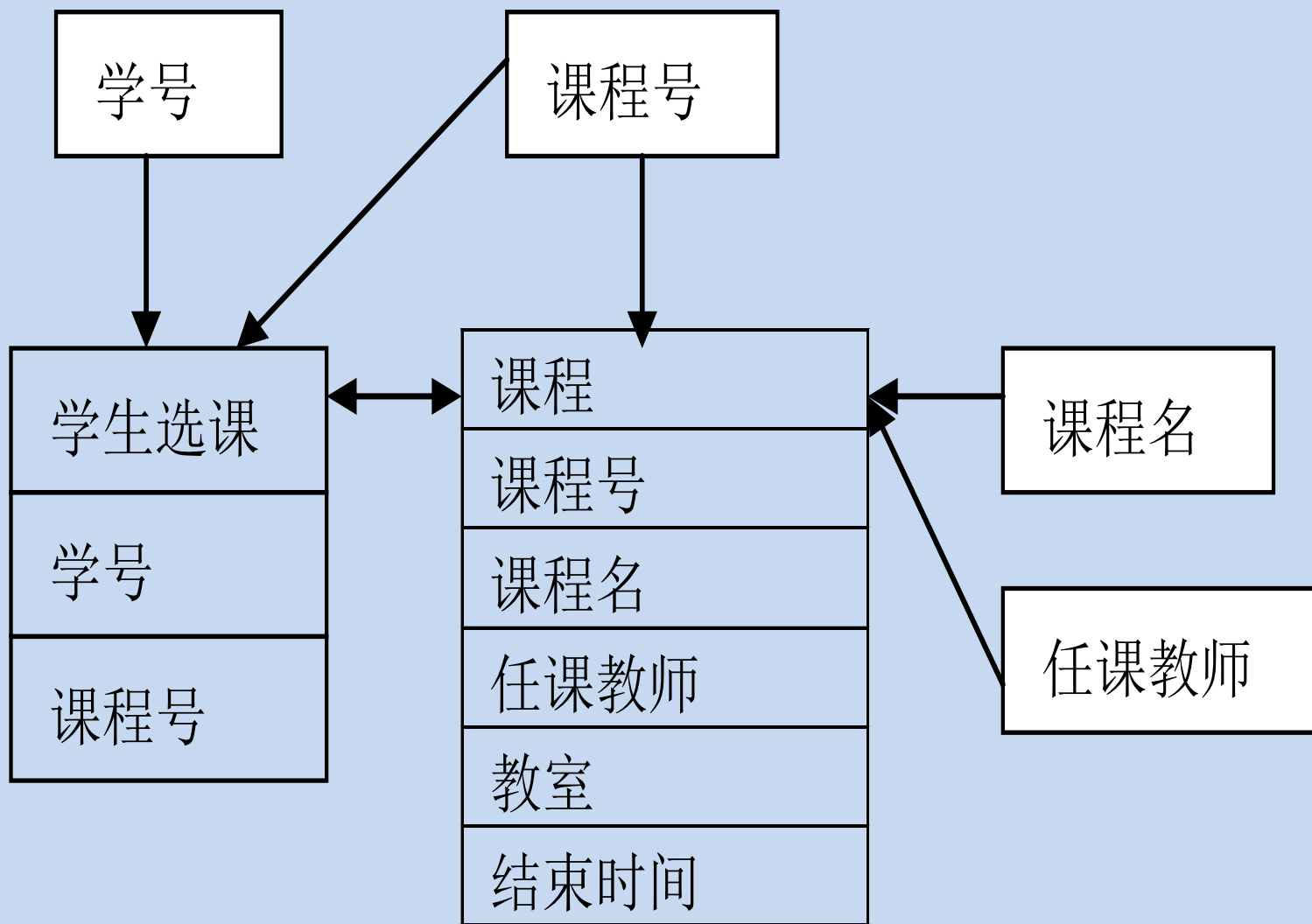
——A (E) = ?

——A (?) { = ≠ < > } V

——? (E) = ?



# 立即存取分析



■ 对处理过程中具体操作的描述，不属于数据字典的范围，这里仅对处理过程的部分数据特性作简单的描述

## ■ 处理过程的定义：

- 处理过程在数据流图中的名称、编号
- 处理过程的简单描述
- 处理过程的输入数据流、输出数据流及其来源与去向
- 其主要功能的简单描述

处理过程名称：编辑学生成绩单

编号：3.1.1

简述：将学生某门课程的考试成绩录入系统，并产生某学生已考课程的成绩单

输入：课程成绩单，来源为外部项“教师”

处理：按一定的格式将所有学生该门课程的考试成绩一次录入“考试成绩”数据存储，根据学号将该课程的考试成绩分别转入该学号的“成绩单”数据存储

输出：考试成绩，去向为“成绩单”数据存储、“确定

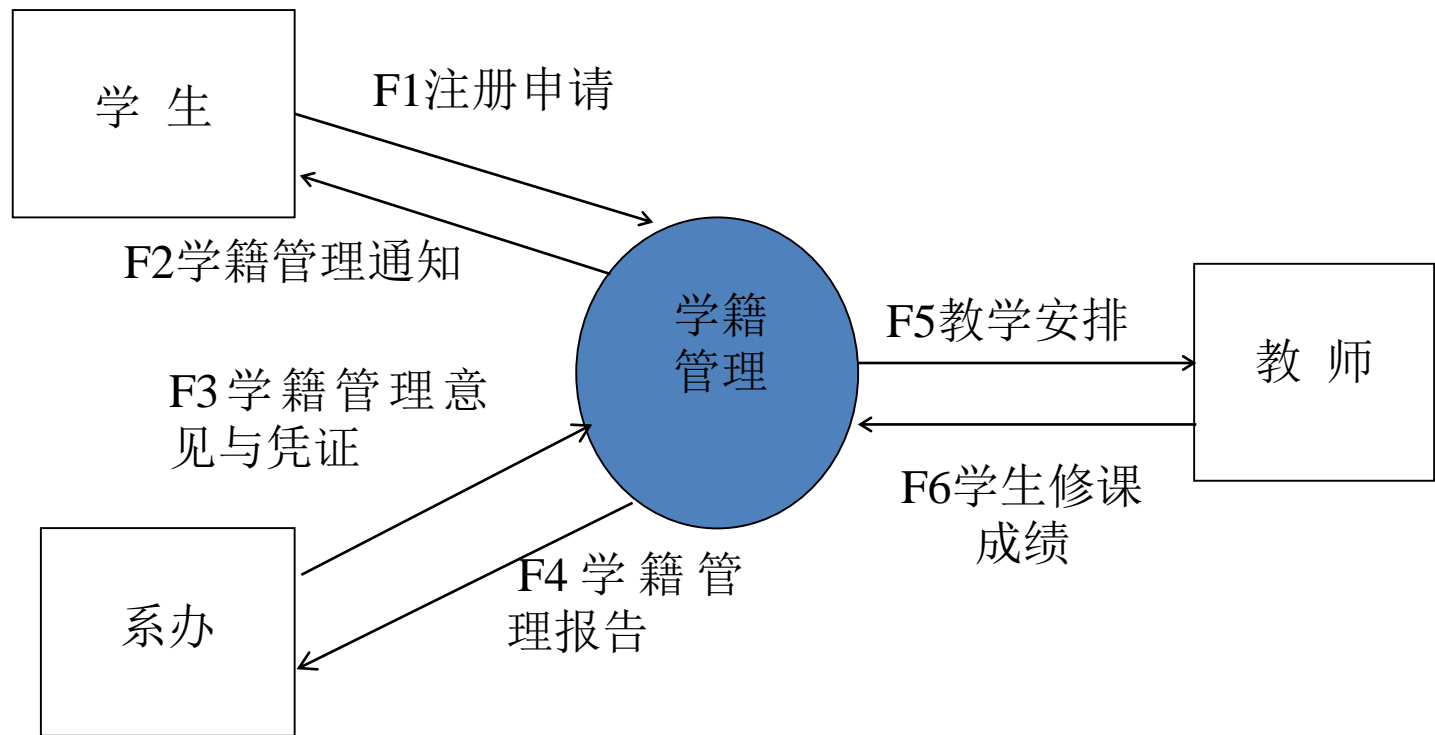
补考”处理过程



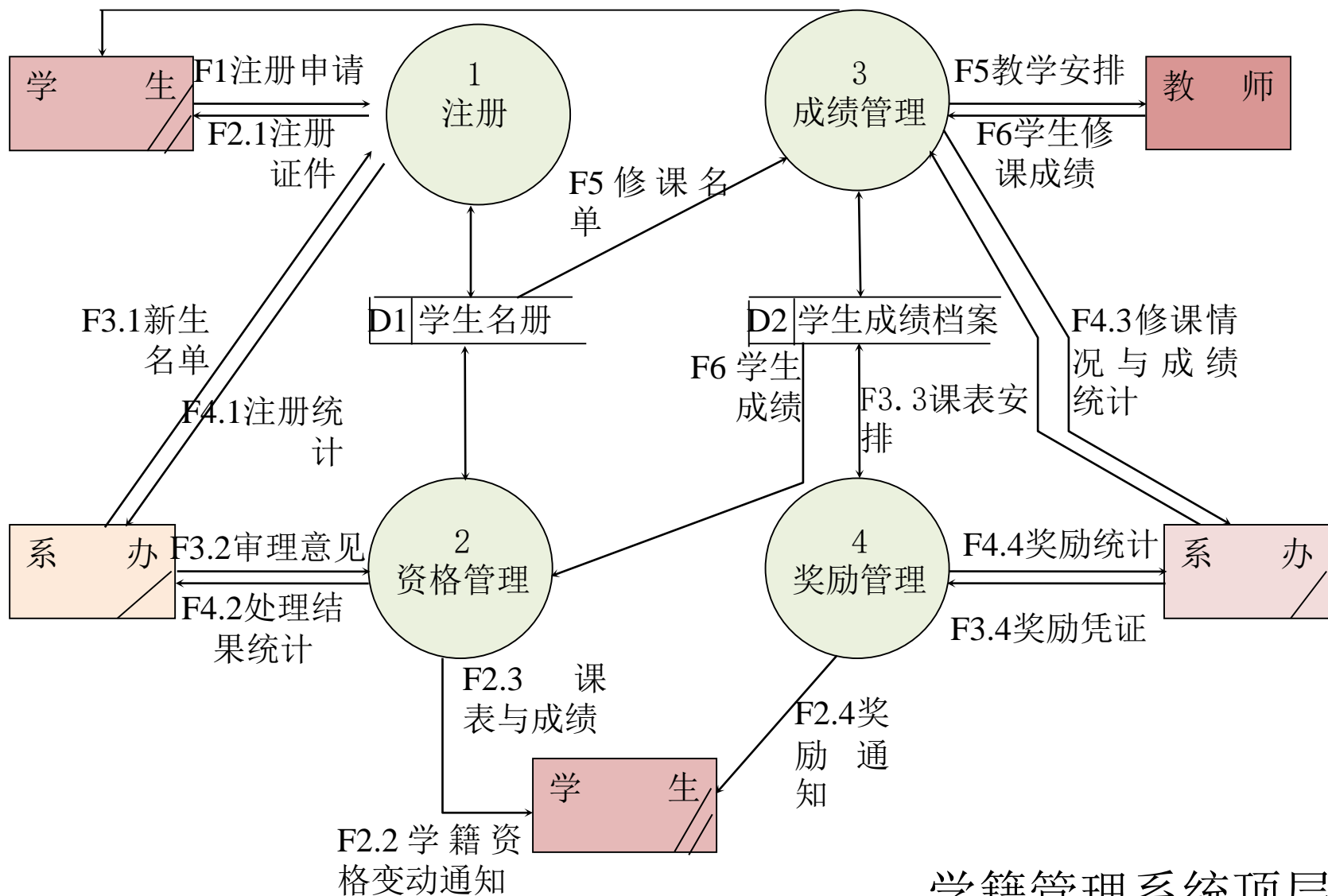
# 课堂练习： 学籍管理系统需求：

- **学生** 可以进行注册申请,获得注册证件, 并获得相关的各类通知;
- **教师** 可以进行教学活动安排,同时完成课程后进行考试,并将成绩单进行统一管理;
- **系办人员** 可以对学生的注册情况（注册名单、注册统计）、学籍情况和资格管理（学籍资格变动通知，选课表安排）、成绩管理（修课名册， 教学安排， 学生成绩， 修课情况与成绩统计）以及学习的成绩进行奖励和认证管理;

# 例4:学籍管理系统



学籍管理系统关联图



学籍管理系统顶层图

数 据 流				
系统名： 学籍管理		编号：		
条目名： 学生成绩通知		别名： 成绩通知单		
来源： 成绩管理		去处： 学生		
数据流结构： 学生成绩通知： {学号+学生姓名+{课程名称+成绩}该生本期所修课程 +（补考课程名称+补考时间+补考地点） }所有在册学生				
简要说明： 学生成绩通知在每学期期末考试结束后一周至下学期开学前一 周期间内发给所有本期在校学生。				
修改纪录：	编写	张XX	日期	2013年5月10日
	审核	李XX	日期	2013年5月10日

数 据 元 素

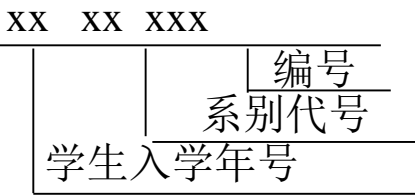
系统名： 学籍管理  
条目名： 学号

编号：  
别名：

属于数据流：  
F1-F7

存储处： D1 学生名册  
D2 学生成绩

数据元素结构：  
代码类型          取值范围          意义  
字符          00010001—992999  
(由数字组成的字符串)



简要说明：  
学号是学生的识别符，每个学生都有唯一的学号。

修改纪录：

编写	张XX	日期	2017. 4. 10
审核	李XX	日期	2017.4.10

数 据 存 储				
系统名：学籍管理 条目名：学生名册		编号： 别名：		
存储组织： 每个学生一条纪录 按学号顺序排列		纪录数：约800 主关键字：学号 数据量：约72KB 辅关键字：学生排名		
纪录组成： 项名：学 姓 性 出生 注册 修课.....修课 备 号 名 别 年月 学期 代号1 代号7 注 近似 长度 7 10 2 4 4 6 ..... 6 20 (字节)				
简要说明： (1)学籍变动（留级，转专业）在备注中说明。 (2)重修课程在备注中说明。				
修改纪录：		编写	张XX	日期 2017. 4. 10
		审核	李XX	日期 2017.4.10

加 工

系统名：学籍管理  
条目名：成绩管理

编号： 3  
别名：

输入：学生修课名单  
课程名称 学生成绩

输出：教学安排 学生成绩通知单  
学生修课情况与成绩统计

加工逻辑：

1. 从学生名册中获取修同一课程的学生名单；
2. 计每门课程的修课人数并报系机关；
3. 从系机关获取课程安排数据，包括各门课程的上课时间。地点；
4. 形成教学安排数据，其中包括各门课程的修课学生名单。上课地点，通知有关任课教师。
5. 接收任课教师的学生成绩数据，并登录在学生成绩档案中；
6. 进行成绩统计，计算每门课程成绩优良。及格。不及格。缺考各项人数及比率，计算各科平均成绩并向系机关报告；
7. 向学生发出学生成绩通知，并附补考安排。

简要说明：

课程安排由系机关中教学管理人员直接向学生公布。

修改纪录：

编写 张XX

日期

2017年4月10日

审核 李XX

日期

2017.4.10

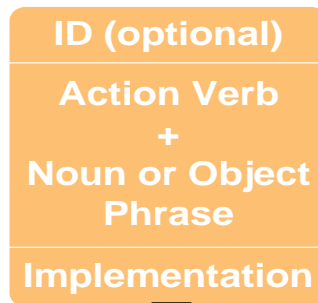
外 部 项				
系统名： 学籍管理		编号： W3		
条目名： 教师		别名： 任课教师		
输入数据流： 教学安排		输出数据流： 学生成绩		
主要特征： 教师：即本系统中为修课学生授课的任课教师，其主要特征是：教师姓名。讲授课程名称。联系地址。				
简要说明： 本系统负责下达教师的教学任务，只是根据系机关课程安排通知教师有关教学安排。				
修改纪录：	编写	张XX	日期	2017. 4. 10
	审核	李XX	日期	2017.4.10



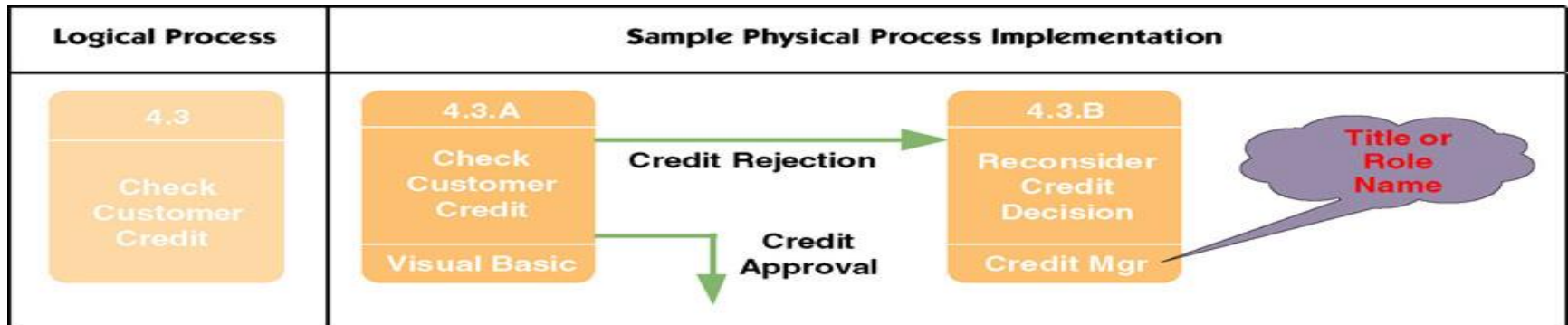
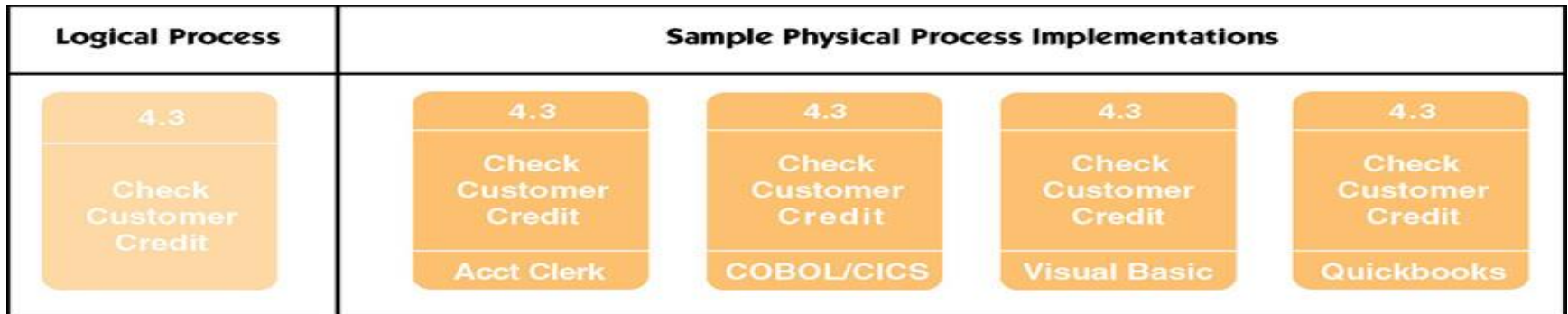
# 总结:利用DFD建模步骤

- 步骤1: 建立顶级DFD
- 步骤2: 自顶向下，功能分解
  - 自顶向下、逐步求精对顶级数据流图进行精化，生成下一级数据流图
  - 对用户的需求描述进行语法分析
    - 名词或名词短语构成潜在数据流、数据存储和外部实体
    - 动词构成潜在的处理功能
- 步骤3: 利用数据字典进行数据补充描述






















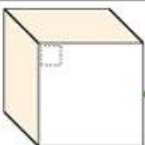

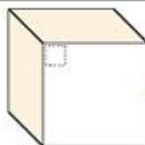

# 物理DFD过程



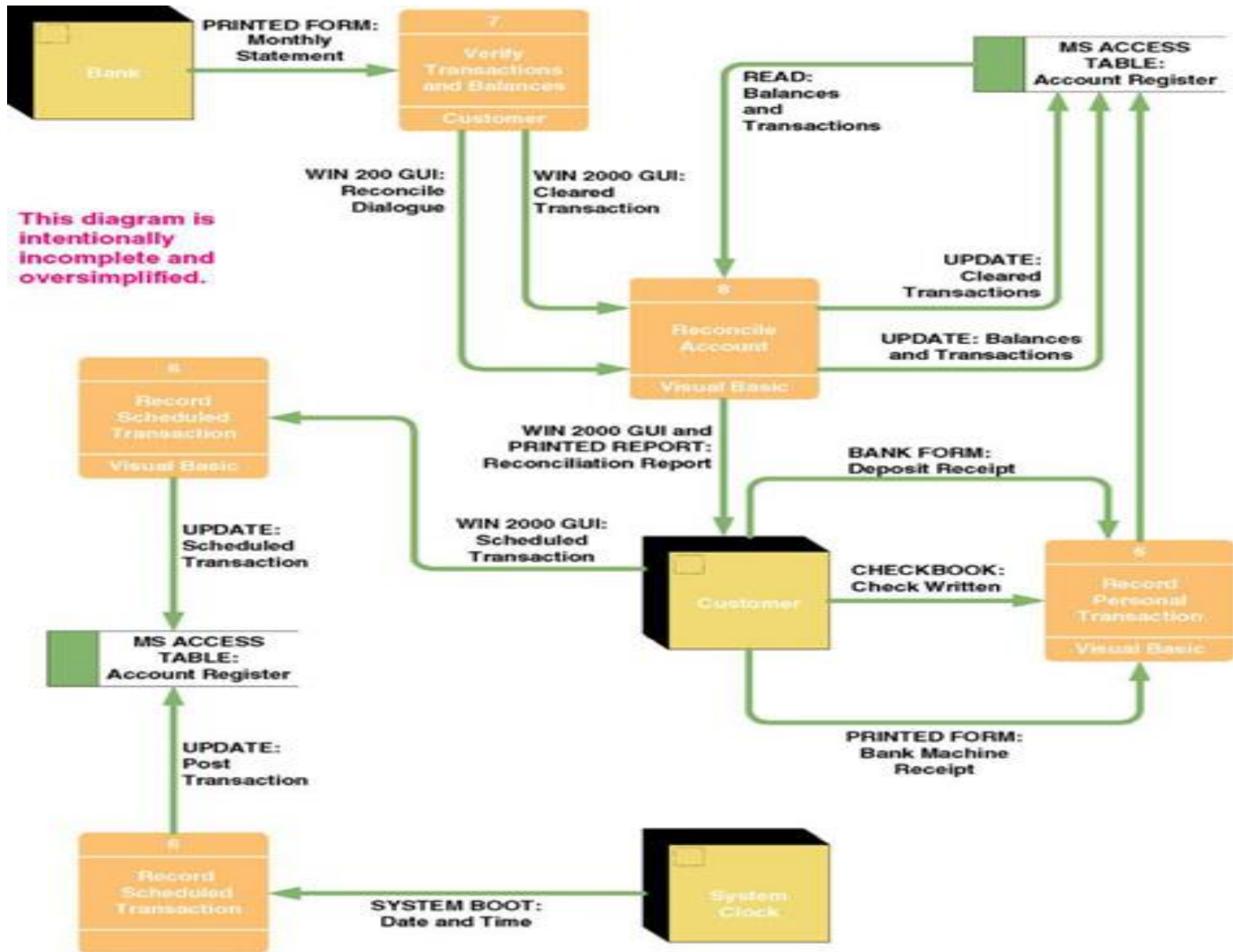
特定的技术处理；  
特定的参与者的处理名称；  
特定的技术或参与者处理顺序；  
冗余的处理、数据流、文件。



















# 1、物理数据流示例：

	Logical Data Flow	Implementation	Sample Physical Data Flow
	 Unfilled Orders →	Read records in a database	 SQL Select: Unfilled Orders →
	Update Credit rating → 	Update a record in a database	SQL Update: Credit Rating → 
	Delete Employee → 	Delete a record in a database	SQL Delete: Employee → 
	Insurance Accident Claim → 	Import a data file	IMAGE FILE: Insurance Accident Claim → 
	 Schedule of Classes →	Export a data file	 Comma Delimited File: Schedule of Classes →
	 Extended Cost → 	Pass data between modules of a program	 Extended Cost → 
	 Course Request → 	Pass a manual form	 Form 23: Course Request → 

## 2、物理DFD

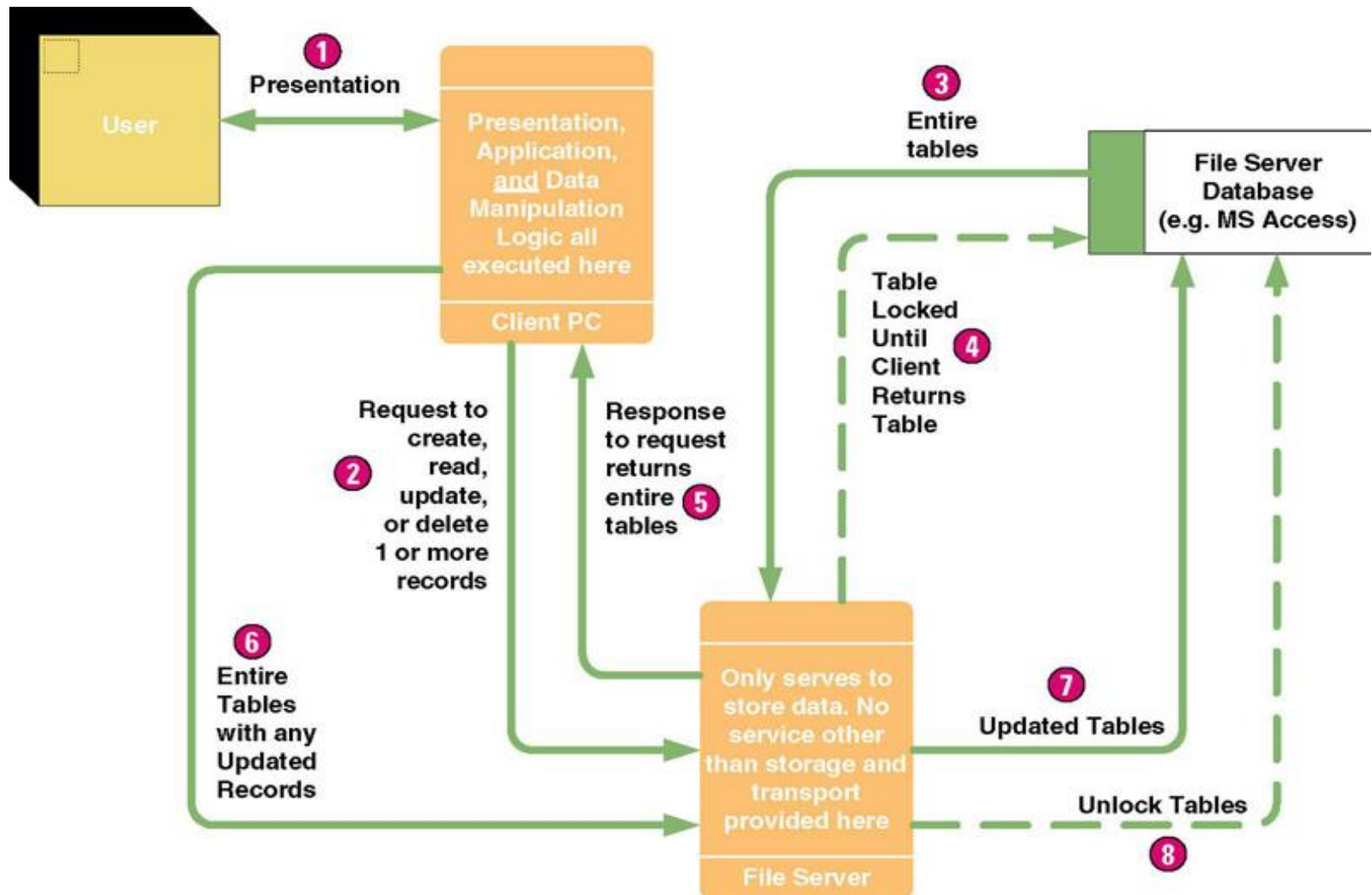


# Physical Data Store Notation

Logical Data Store	Implementation	Physical Data Store
 Human Resources	A database (multiple tables)	 Oracle : Human Resources DB
 Marketing	A database view (subset of a database)	 SQL Server: Northeast Marketing DB
 Purchase Orders	A table in a database	 MS Access: Purchase Orders
 Accounts Receivable	A legacy file	 VSAM File: Accounts Receivable
 Tax Rates	Static data	 ARRAY: Tax Table
 Orders	An off-line archive	 TAPE Backup: Closed Orders
 Employees	A file of paper records	 File Cabinet: Personnel Records
 Faculty/Staff Contact Data	A directory	 Handbook: Faculty/Staff Directory
 Course Enrollments By Date	Archived reports (for reuse and recall)	 REPORT MGR: Course Enrollment Reports

# 3、DFD应用： 信息技术架构

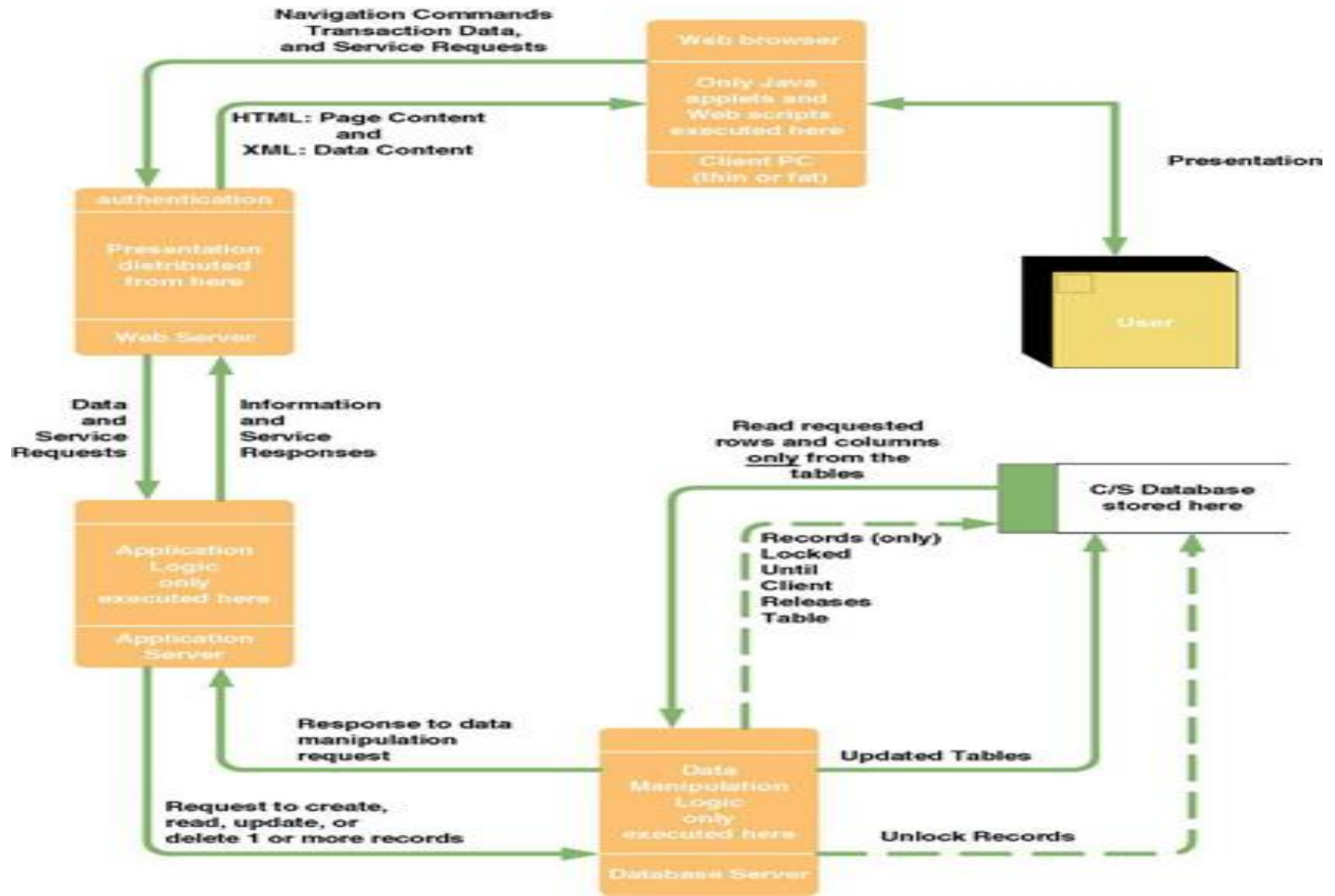
## File Server Architecture





# DFD应用：信息技术架构

## C/S Architecture



## 4、数据流图的作用与局限性

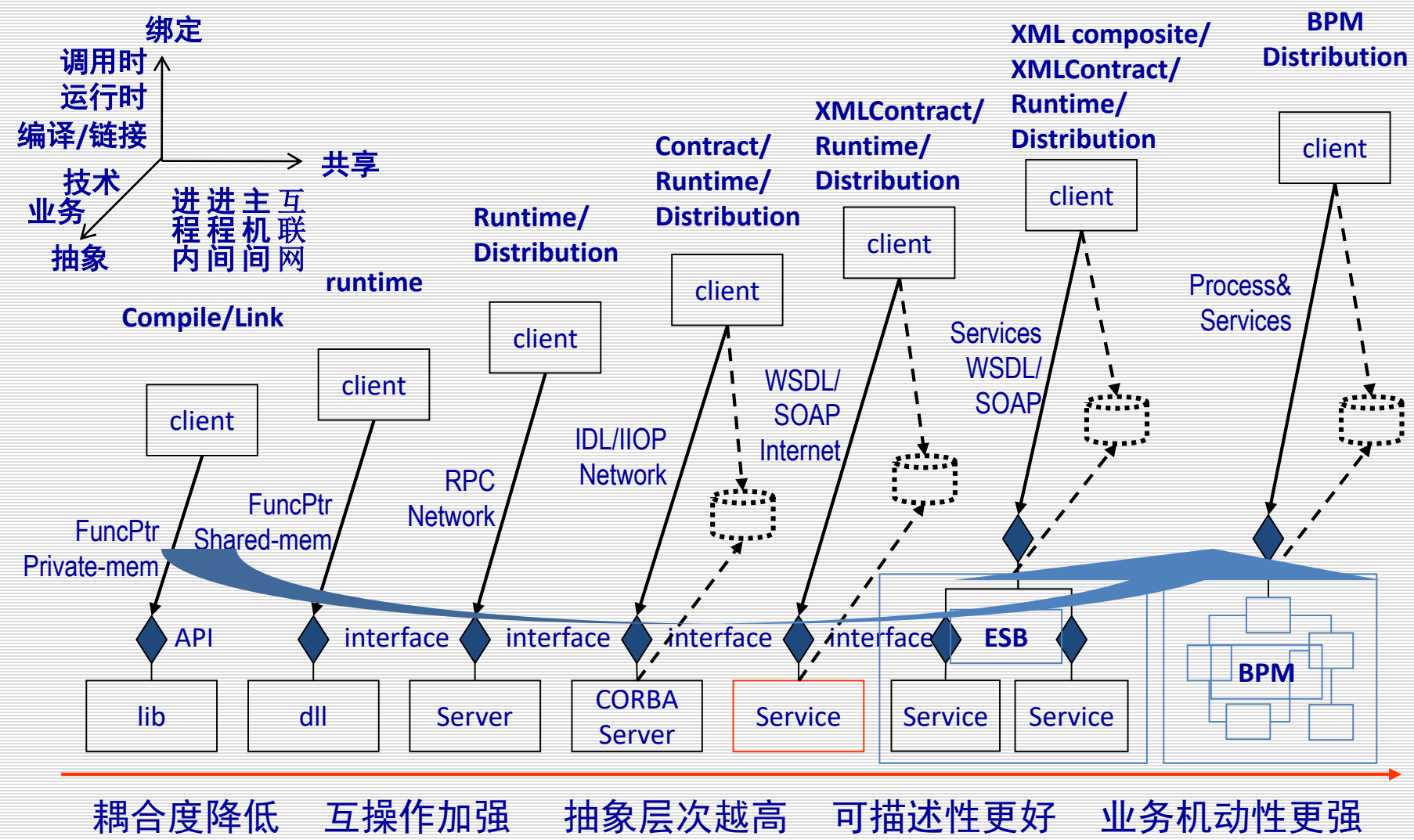
- 数据流图就是组织中信息运动的抽象，是管理信息系统逻辑模型的主要形式。
- 数据流图提供了**层次结构**，让分析人员能够方便地表示任意抽象级别上的信息系统或其子系统，并**支持问题分解、逐步求精的分析方法**。
- 这个模型不涉及硬件、软件、数据结构与文件组织，它与对**系统的物理描述无关**，用一种图形及与此相关的注释来表示系统的逻辑功能，即所开发的系统在管理信息处理方面要做什么。
- 数据流图是系统分析人员与用户进行交流的有效手段，也是系统设计(即建立所开发的系统的物理模型)的主要依据之一。



## 4、数据流图的作用与局限性

- 难以在数据流图上标识出数据流、数据存储，加工和外部项的具体内容，如数据流的组成元素，数据存储的数据结构，存取要求，数据量，加工的处理过程与算法等。
- 不能反映系统中的决策与控制过程。
- 难以对系统中人机交互过程以及信息的反馈与循环处理进行描述。

# 软件开发技术的演变:



# 结束语:

- 建模是系统分析的核心, 模型的优劣与否将决定着系统的各项性能.
- 不同的模型是从不同的角度来对系统进行分析 and 描述,
- 模型与建模过程是灵活的, 希望能够利用现有技术来描述更多的内容.