

软件可靠性测试

北京航空航天大学可靠性与系统工程学院 2010-5-13

回顾 (1/4) : 测试公理

- 测试只能证明错误的存在，而不能表明程序中没有错误。
- 缺陷分布存在“二八原则”，即20%的模块占有80%的缺陷。
- 软件测试最困难的问题之一是知道何时停止测试(When to stop testing?)。
- 自己测试自己的程序是不可能的。
- 每个测试用例必备的部分是描述预期的输出。
- 一个好的测试用例是一个能发现新缺陷的用例，而不是一个表明程序工作正确的用例。

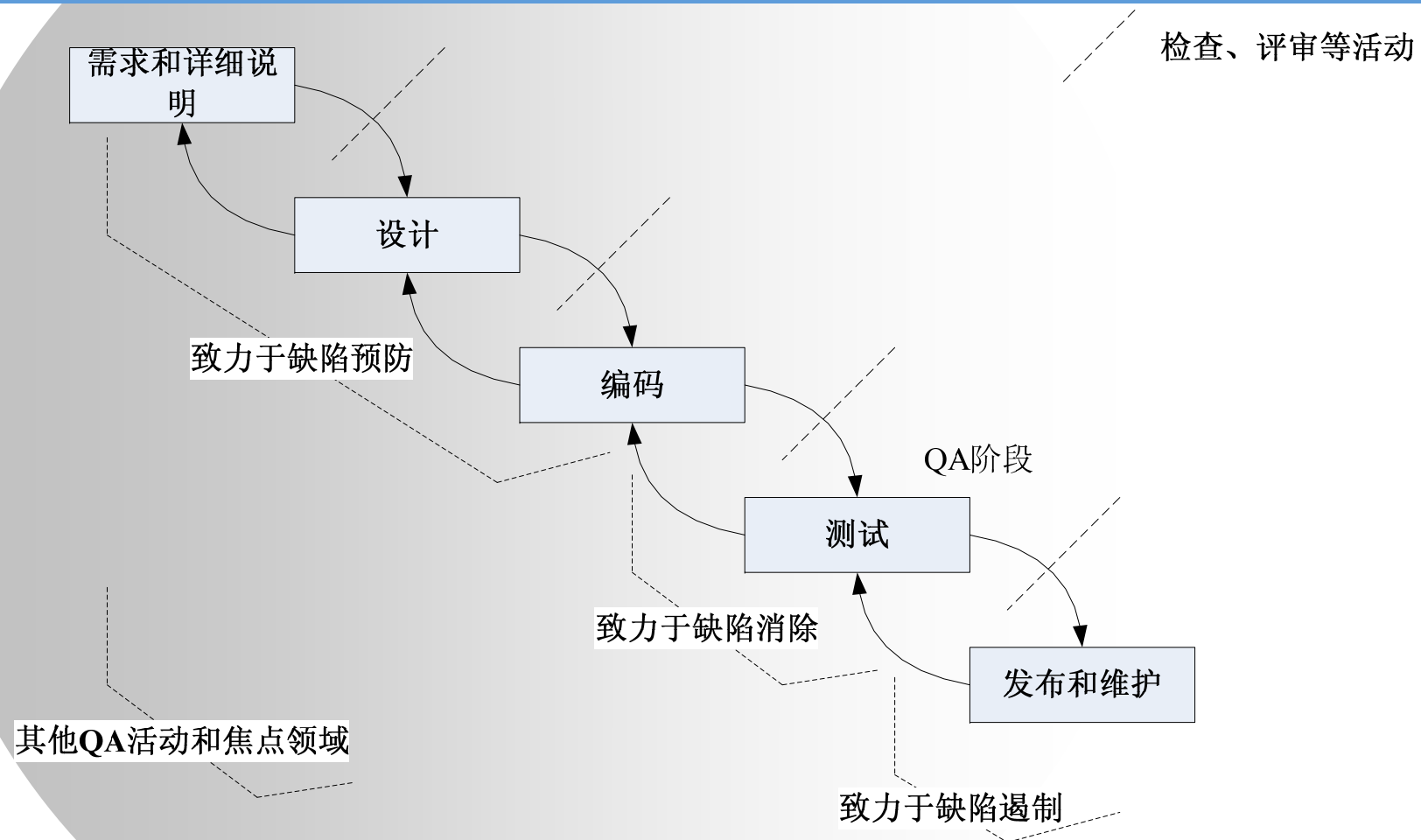


回顾(2/4): 已讲技术

- 配置管理技术
- 软件需求、设计、编码避错设计准则
- 软件静态测试、白盒测试、黑盒测试、软件可靠性测试



回顾(3/4): 与缺陷做斗争



回顾（4/4）：什么是软件可靠性

- 在规定的条件下，在规定的时间内软件不引起失效的概率。该概率是系统输入和系统使用的函数，也是软件中存在的缺陷的函数。系统输入将确定是否遇到已存在的缺陷（如果有缺陷存在的话）。
- 规定的条件：
 - 软件运行的软、硬件环境
 - 软件操作剖面：软件运行的输入空间及其概率分布
- 规定的时间：
 - 执行时间
 - 日历时间
 - 时钟时间



目标

- 软件可靠性测试基本概念
- 如何构造软件操作剖面
- 如何生成软件可靠性测试数据
- 软件失效数据种类及其收集方法



内容

- 软件操作剖面构造
- 软件可靠性测试数据生成
- 软件失效数据收集
- 软件可靠性测试实例



软件可靠性测试基本概念



软件可靠性测试的基本概念

1. 为什么要进行软件可靠性测试?
2. 什么是软件可靠性测试? 和一般测试有何异同?
3. 软件可靠性测试分类
4. 软件可靠性测试流程



什么情况下需要进行软件可靠性测试？

- 软件具有明确的可靠性定量要求，并需要进行验证
- 需要评估一个软件的可靠性定量水平
- 希望高效地达到可靠性目标要求

对软件可靠性进行

定量的评估或验证！



软件可靠性测试

为了达到和/或验证软件的可靠性定量要求而对软件进行的测试

主要特征：按照用户实际使用软件的方式测试软件



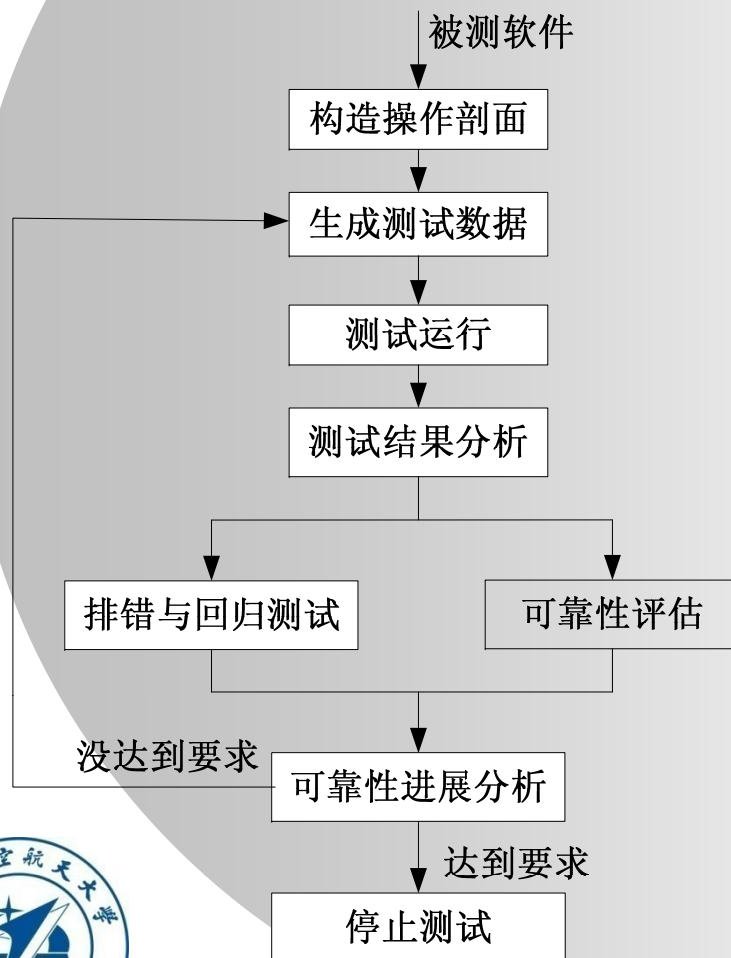
软件可靠性测试分类

- 软件可靠性增长测试
 - 测试目的：发现程序中影响软件可靠性的故障，并排除故障实现软件可靠性增长
 - 实施阶段：软件系统测试阶段的末期
- 软件可靠性验证测试
 - 测试目的：验证在给定的统计置信度下，软件当前的可靠性水平是否满足用户的要求
 - 实施阶段：软件验收阶段

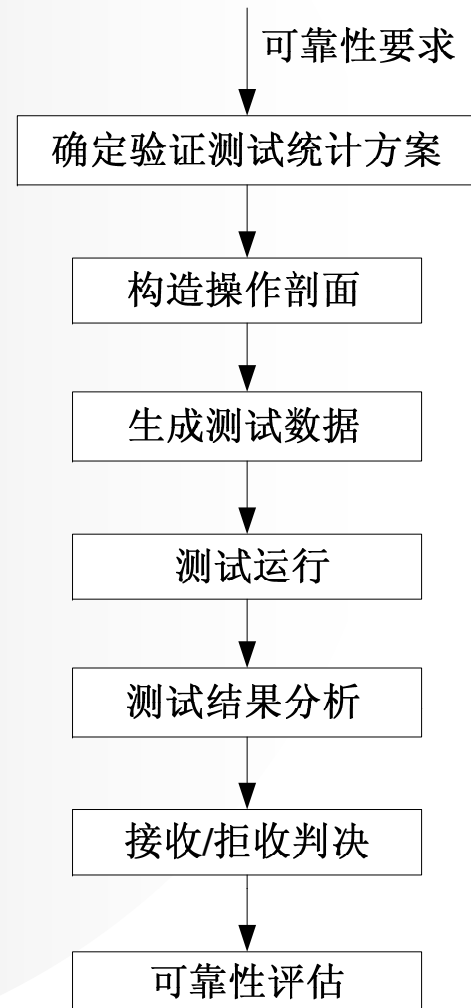


软件可靠性测试流程

软件可靠性增长测试



软件可靠性验证测试



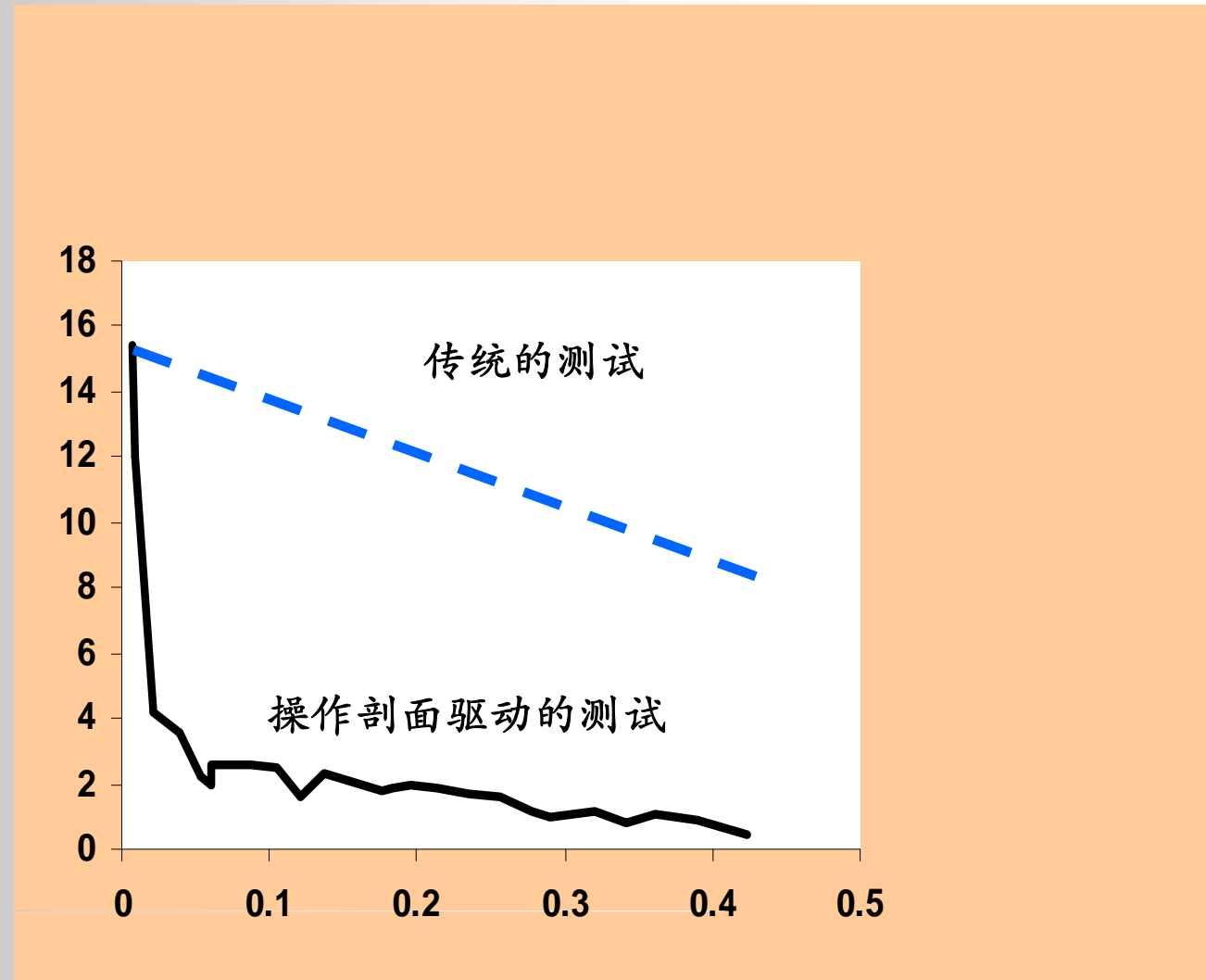
软件可靠性测试与一般测试的比较

比较项目	软件可靠性增长测试	一般软件测试
测试目的	评估软件可靠性水平 有效实现软件可靠性增长	发现软件的故障
测试效率	达到可靠性要求 较快	达到可靠性要求 较慢
测试数据生成方法	基于使用的测试， 根据软件的 使用状况 构造操作剖面然后生成测试用例	基于需求/结构的测试， 根据软件的 需求或结构 生成测试用例
数据收集	需要收集测试输出结果和失效时间等数据	只需收集测试输出结果
数据分析	通过 失效数据 进行可靠性分析	根据 用例执行情况 进行需求/结构覆盖分析
测试停止准则	满足可靠性要求	功能/性能测试：需求覆盖100% 结构测试：语句覆盖100%、 分支覆盖100%或满足其它结构覆盖要求



软件可靠性测试比传统测试更快接近目标值

失效率
/失效率目标值



百万次呼叫 Mcalls



构造软件操作剖面

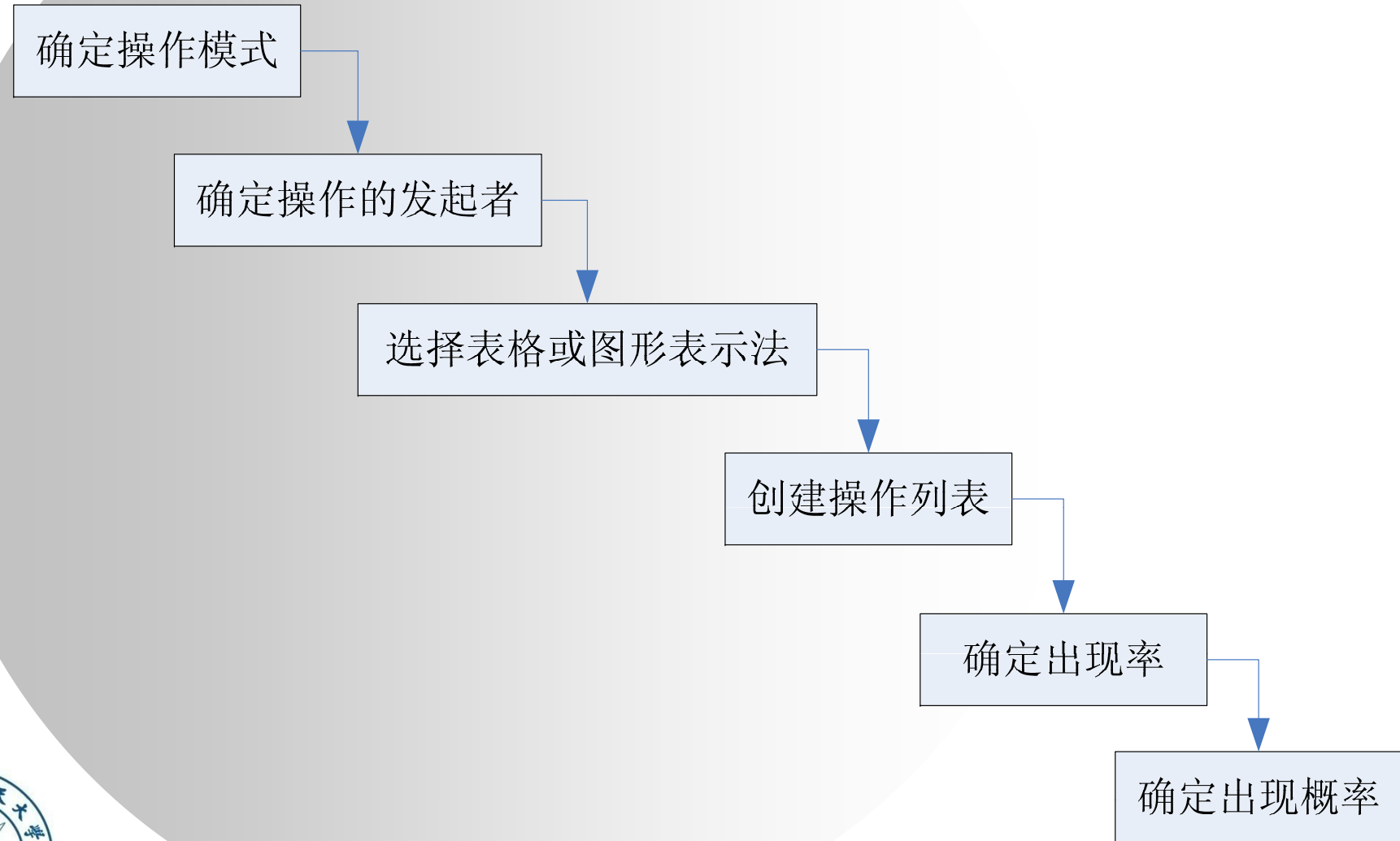


操作剖面的概念

- **剖面**：一组独立的称之为元素的可能情况和与之相关的发生概率，所有元素的概率和为1
 - 设A、B、C是相互独立的元素，某一时刻其发生的概率分别为：60%、30%、10%，则此时的剖面为：
[A,0.6...B,0.3...C,0.1]
- **操作剖面**：
 - Musa给出的定义：指一组操作及其发生的概率
 - **操作**：是一个主要的系统逻辑任务，持续时间短，结束时将控制权交还给系统，并且它的处理与其它操作有显著不同



操作剖面的构造过程



步骤1——确定操作模式

- 操作模式
 - 系统使用的不同模式
 - 需要单独测试的环境条件的集合
- 可能产生不同操作模式的因素
 - 一周的某天或一天的某段时间（主要时间与次要时间）
 - 一年的某段时间（金融系统的年末财政结算）
 - 业务量水平
 - 不同的用户类型（可能会以相同的方式使用系统的用户集合）
 - 用户的经验（专家和新手对系统的使用是不同的）
 - 系统的成熟度（数据库总数据量）
 - 精简的系统能力（对所有操作，或只对特定操作）



操作模式示例——BBS系统

根据不同的用户类型可以确定四个操作模式：

- 站务——负责整个BBS站点的管理工作，如开设版面、审核用户等。
- 版务——负责一个或多个版面的管理工作，如文章加精、置底、制作合集等。
- 用户——BBS注册用户，可执行发贴、投票、竞选版务/站务等操作。
- 游客——未注册人员，仅能浏览BBS的公开版面，不能执行发帖等操作。



操作模式示例——电话系统

- 高峰时段
 - 很大的呼叫和拨号通信量，不执行管理和检查操作
- 主要时段
 - 一般的呼叫和拨号通信量，执行管理操作，但不执行检查操作
- 低谷时段
 - 较低的呼叫和拨号通信量，较少的管理操作，大量执行检查操作



步骤2——确定操作的发起者

- 通过操作的发起者能够**系统地**得到软件的操作
- 操作的发起者
 - 系统用户
 - 外部系统
 - 系统自身的控制程序



发起者示例——BBS系统

- 根据用户类型可以确定BBS系统的操作发起者包括站务、版务、用户和游客。除此之外，BBS系统定期（每月或每季度一次）会进行数据库的备份操作，因此系统控制器也是一个操作发起者。



步骤3——选择表格或图形表示法

● 操作剖面表格表示法

- 由操作及其相关概率构成表格，如下表所示

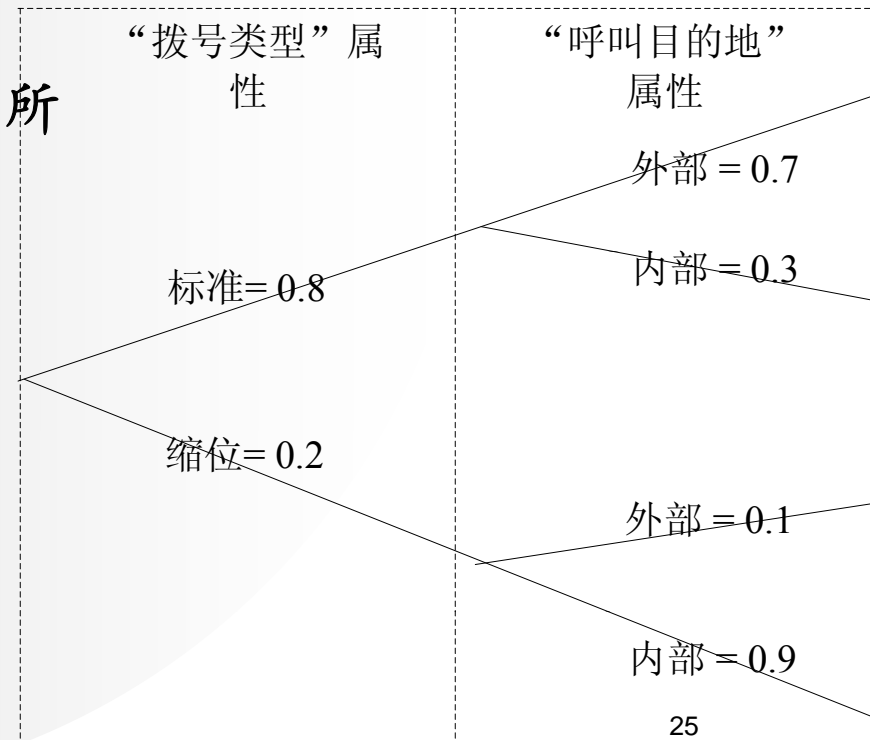
操作	发生概率
标准外部拨号	0.56
标准内部拨号	0.24
缩位外部拨号	0.02
缩位内部拨号	0.18

- 操作可以用**很少的**属性表示时选用
- 当前应用较广泛



● 操作剖面图形表示法

- 操作需要由多个属性来表示时选用
- 每个操作为：图中的一条路径
- 每个属性值都有一个关联的发生概率
- **操作**的发生概率由表示该操作的所有分支的发生**概率相乘**



步骤4——创建操作列表

- 操作列表创建方法
 - 为每个操作发起者创建一张操作列表，然后再合并
- 参考依据
 - 主要参考系统需求
 - 也可参考用户手册、系统原型以及以前的系统版本等
- 不同表示法的操作列表创建
 - 表格表示法：直接列出操作
 - 图形表示法：标出属性和属性值的方式间接列出操作



创建操作列表示例——BBS系统

操作发起者	操作
站务	审核用户信息
	新建版面
	仲裁
版务	删帖
	加精
	置底（或置顶）
	制作合集
用户	发帖
	发在线消息
游客	浏览文章
	注册
系统控制器	备份数据库

表中只列出了各操作发起者的一些主要操作，某些较少执行的操作如删除版面、封禁用户等未在表中列出。



步骤5——确定出现率

- 出现率
 - 操作在一段时间内的发生次数
- 数据来源
 - 最好使用真实的现场数据，例如通过系统日志得到
 - 相同或相似系统的数据
 - 没有可直接使用数据时，可依据相关信息合理估计
 - 无法估计操作的相对概率时，可假设所有操作的出现率相等
- 不同表示法出现率的确定
 - 表格表示法：确定操作的出现率
 - 图形表示法：确定属性值的出现率



确定出现率示例——BBS系统

操作	出现率（每天发生次数）
审核用户信息	50
新建版面	0.02
仲裁	0.04
删帖	15
加精	10
置底（或置顶）	2
制作合集	0.1
发帖	5
发在线消息	20
浏览文章	25
注册	60
备份数据库	0.05
合计	127.21



步骤6——确定发生概率

- 发生概率：利用出现率计算
- 不同表示法发生概率的确定
 - 表格表示法
 - 操作的发生概率=操作的出现率/所有操作的总出现率
 - 图形表示法
 - 属性值的发生概率=每个属性值得出现率/该属性的总出现率



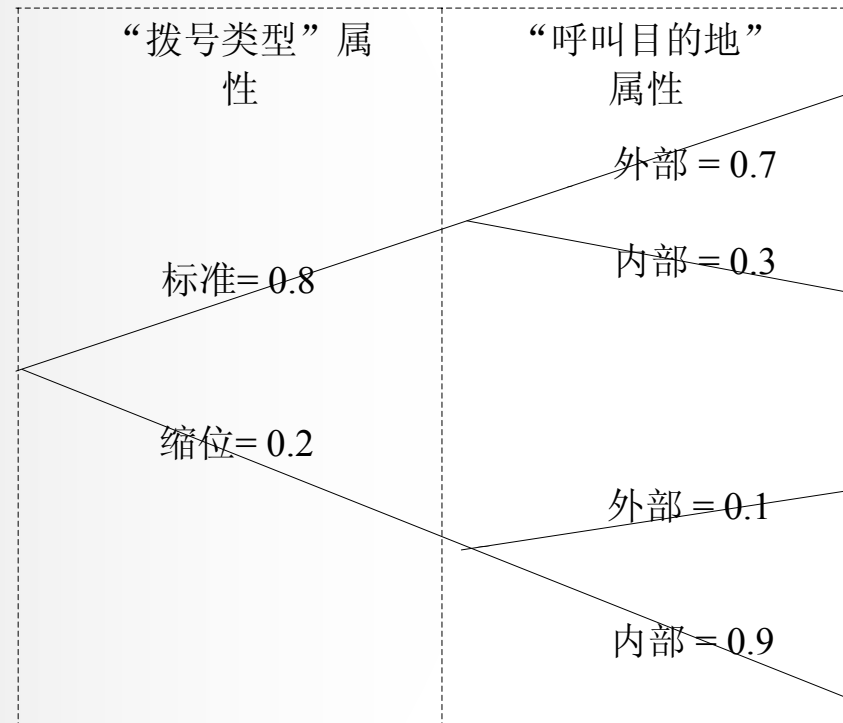
确定发生概率——BBS系统

操作	发生概率
审核用户信息	0.2822
新建版面	0.0001
仲裁	0.0002
删帖	0.0846
加精	0.0564
置底（或置顶）	0.0113
制作合集	0.0006
发帖	0.0282
发在线消息	0.1129
浏览文章	0.1411
注册	0.2822
备份数据库	0.0003
合计	1.0



操作剖面示例

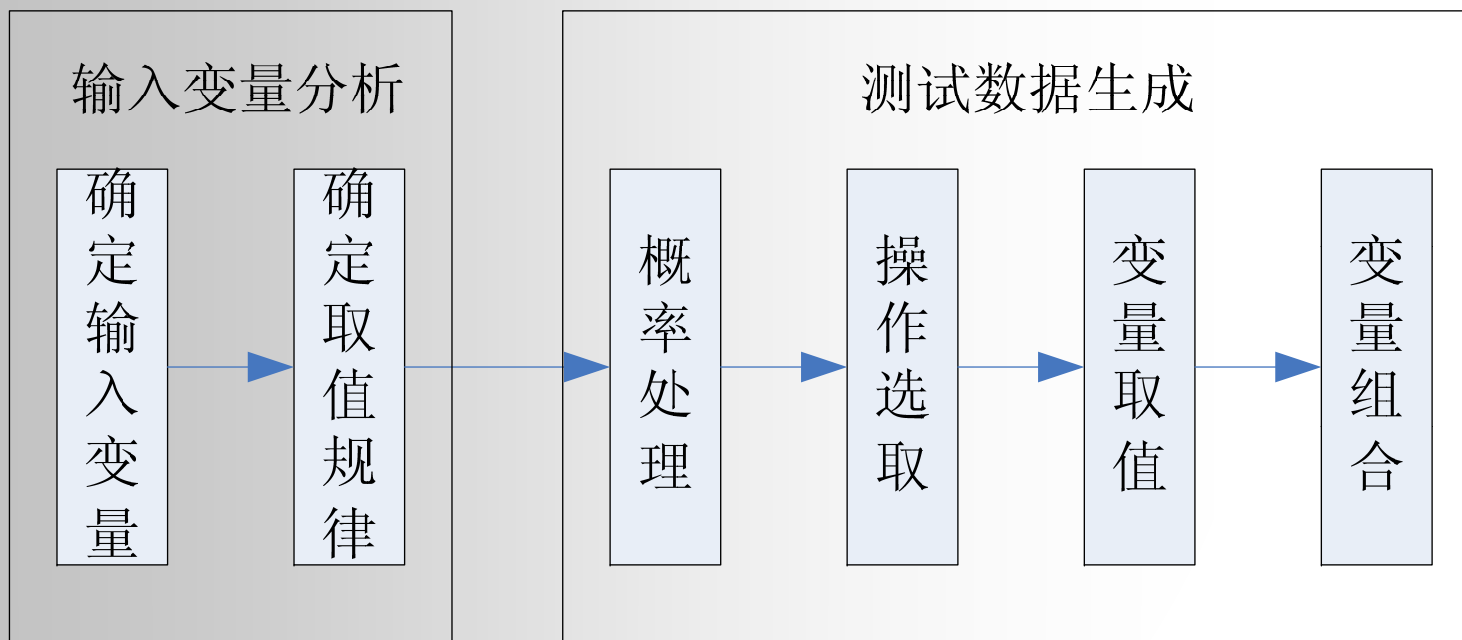
操作	发生概率
标准外部拨号	0.56
标准内部拨号	0.24
缩位外部拨号	0.02
缩位内部拨号	0.18



软件可靠性测试数据生成



软件可靠性测试数据生成流程



输入变量分析

- 目的
 - 确定操作的输入空间
- 输入变量
 - 任何存在于系统外部、同时影响此系统的数据元素
- 分析步骤
 - 确定输入变量
 - 确定取值规律



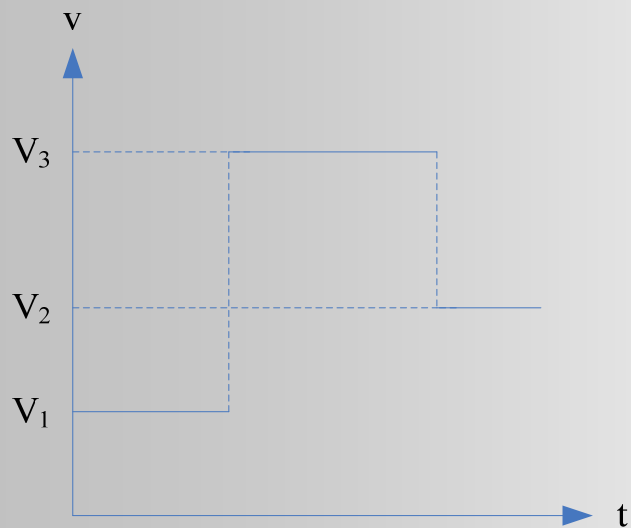
确定输入变量

- 分类方式：
 - 时间的相关性（相关、无关）
 - 取值的确定性（确定、不确定）
 - 取值范围数据特征（离散、连续）
 - 数据类型（整数型、实数型、字符型、逻辑型）

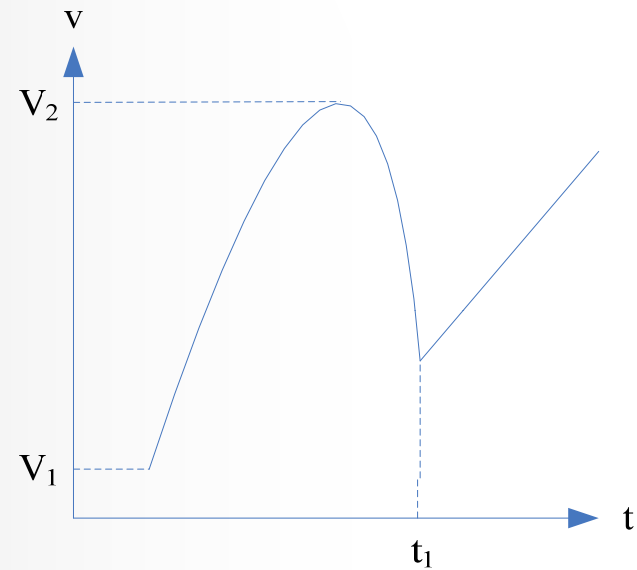
分类方式						输入变量类型
时间相关性		取值确定性		取值范围数据特征		
相关	无关	确定	不确定	离散	连续	
√		√		√		<u>时间相关确定离散型</u>
√		√			√	<u>时间相关确定连续型</u>
√			√	√		时间相关不确定离散型
√			√		√	时间相关不确定连续型
	√	√		√		时间无关确定离散型
	√	√			√	时间无关确定连续型
	√		√	√		<u>时间无关不确定离散型</u>
	√		√		√	<u>时间无关不确定连续型</u>



确定变量取值规律



时间相关确定离散型



时间相关确定连续型



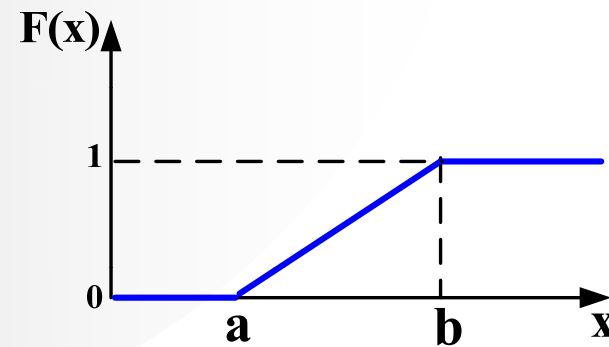
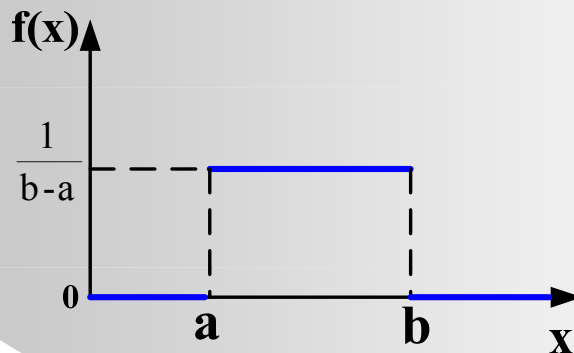
确定变量取值规律

- 时间无关不确定离散型

\mathcal{E}	X_1	X_2	...	X_n	...
$P(\mathcal{E} = X_i)$	P_1	P_2	...	P_n	...

- 时间无关不确定连续型

- 变量X值域: $X=[a, b]$
- 均匀分布的密度函数和分布函数



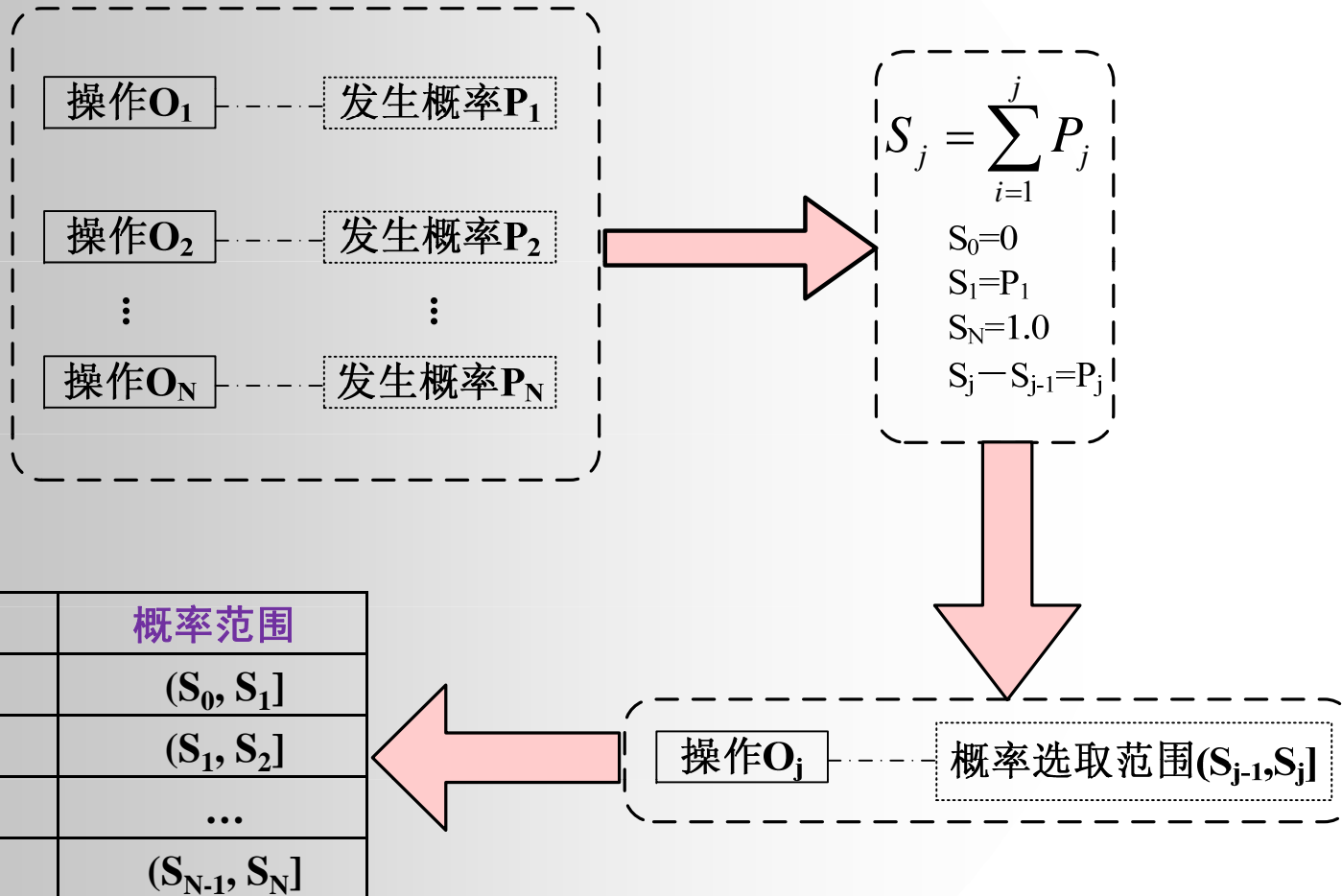
数据生成过程



如果根据操作剖面随机选取操作？

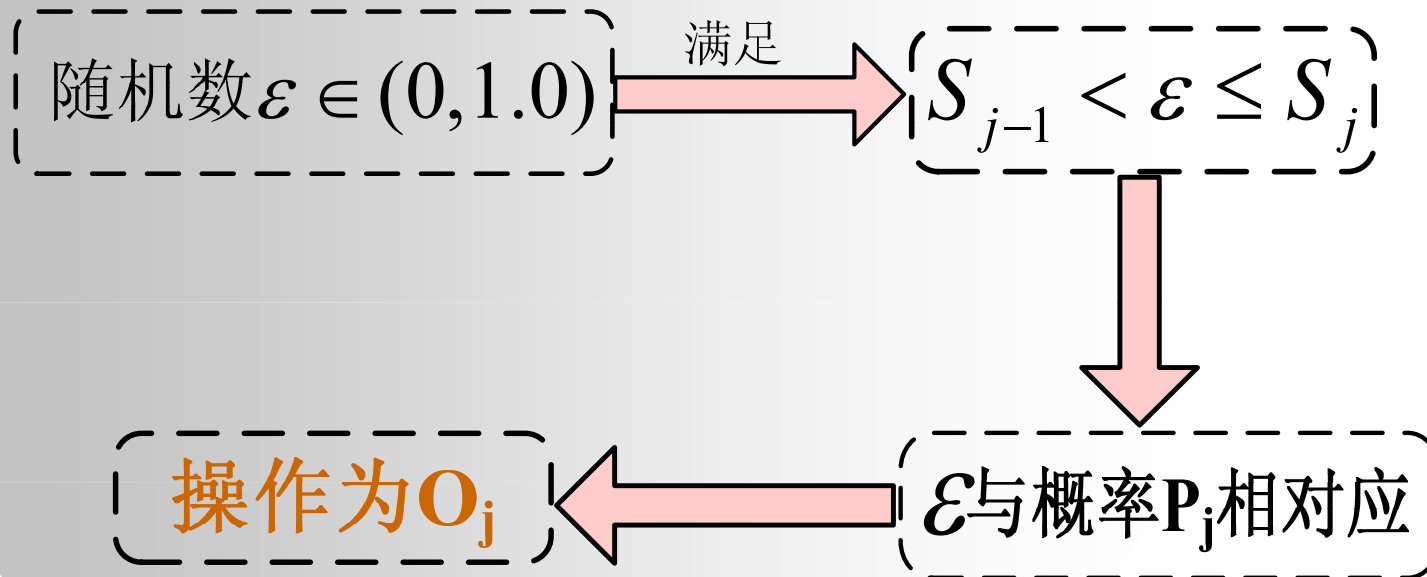


步骤1——概率处理



步骤2——操作选取

操作	概率范围
O_1	$(S_0, S_1]$
O_2	$(S_1, S_2]$
...	...
O_N	$(S_{N-1}, S_N]$



步骤3——变量取值

确定操作 O_j 中每个输入变量的具体取值

输入变量类型	取值方法
时间相关确定离散型	由 $V(t)$ 确定
时间相关确定连续型	由 $V(t)$ 确定
时间相关不确定离散型	由离散型随机过程确定
时间相关不确定连续型	由连续型随机过程确定
时间无关确定离散型	从可取值集合中指定
时间无关确定连续型	从取值实数域中指定
时间无关不确定离散型	根据可取值的概率分布随机确定
时间无关不确定连续型	根据取值实数域的概率分布密度随机确定



步骤4——变量组合

输入变量具体取值按顺序或时序组合得到软件可靠性测试数据



软件失效数据收集



软件失效数据收集

- 目的
 - 进行软件可靠性评估
- 收集的数据包括
 - 软件失效时间
 - 软件失效后果
 - 软件失效计数



软件失效时间：日历/时钟/执行时间

- 日历时间
 - 编年时间，包括计算机可能没有运行的时间
- 时钟时间
 - 从程序执行开始到程序执行结束完毕所经过的时钟时间
- 执行时间
 - 执行一个程序所用的实际时间
 - 执行程序所用的中央处理器时间
 - 程序处于执行过程中的一段时间



失效严重程度

根据系统能力影响划分的失效严重程度

失效严重程度	定义
1	用户不能进行一项或多项关键操作
2	用户不能进行一项或多项重要操作
3	用户不能进行一项或多项操作，但是有补救办法
4	一项或多项操作中的小缺陷



失效数据类型

● 完全失效数据

- 每一次失效发生的时间构成的失效数据，也称失效时间数据
- 表现形式
 - 失效间隔时间：相邻两次失效出现的时间间隔
 - 累计失效时间：将每次失效发生的时间相加得到的结果

● 不完全失效数据（失效计数数据）

- 由各时间段内发生的失效次数构成的失效数据，也称为失效计数数据
- 表现形式
 - 失效强度：单位时间内出现的失效数
 - 累计失效数：某一时刻为止总共出现的失效数



完全失效数据

失效序号	失效间隔时间 (小时)	累计失效时间 (小时)
1	0.5	0.5
2	1.2	1.7
3	2.8	4.5
4	2.7	7.2
5	2.8	10.1
6	3.1	13.1
7	1.8	14.8
8	0.9	15.7
9	1.4	17.1
10	3.5	20.6
11	3.4	24.1
12	1.2	25.2
13	0.9	26.1
14	1.7	27.8
15	1.4	29.2
16	2.7	31.9
17	3.2	35.1
18	2.5	37.6
19	2.1	39.6
20	4.5	44.1
21	3.5	47.6
22	5.2	52.8
23	7.2	60.1
24	10.7	70.1



不完全失效数据（失效计数数据）

时间（小时）	失效数	累计失效数
8	4	4
16	4	8
24	3	11
32	5	16
40	3	19
48	2	21
56	1	22
64	1	23
72	1	24



软件可靠性测试实例



超市信息管理系统简介

- 超市的信息管理系统可以对商品的销售、供应、库存等情况进行统一管理
- 数据库由商品、连锁店、仓库、销售、库存、供应六个相互联系的数据表组成
- 主要操作

- 基本信息
- 高级信息
- 高级查询
- 退出

基本信息 高级信息 高级查询 退出

商店
仓库
连锁店

基本信息 高级信息 高级查询 退出

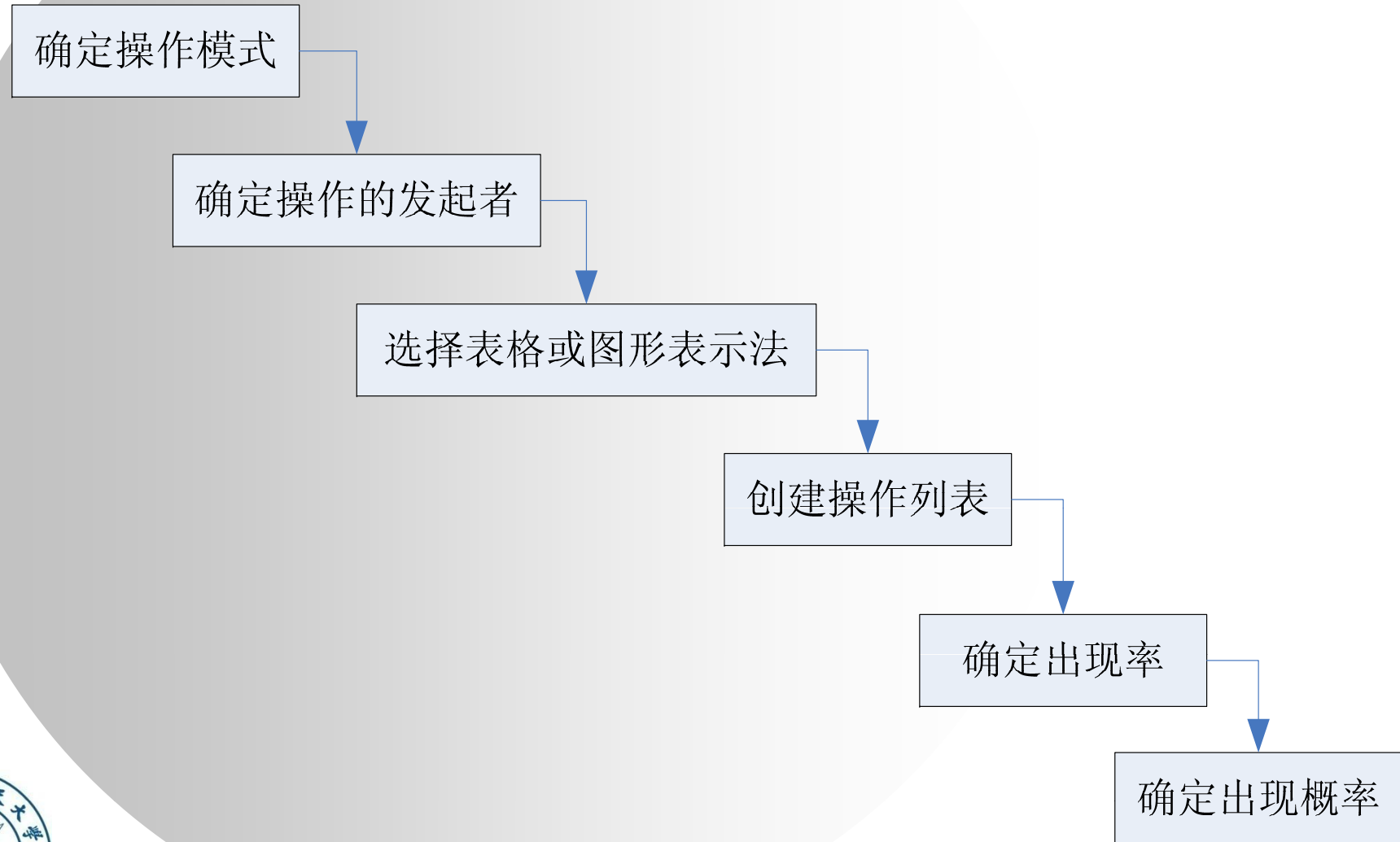
销售
供应
库存

基本信息 高级信息 高级查询 退出

统计信息



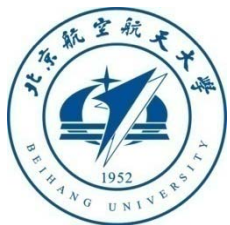
操作剖面的构造过程



构造操作剖面——确定操作模式

根据用户类型的不同确定操作模式

- 经理模式
 - 主要执行连锁店的管理、供应管理和各种查询操作
- 会计模式
 - 主要执行一些审计操作，如统计商品的销售、供应、库存等
- 售货员模式
 - 主要执行与销售相关的操作，如添加、更改、删除销售信息，统计供应情况等
- 仓库管理员模式
 - 主要维护库存信息，为此还需要执行一些基本信息的操作如商品、仓库的管理等



构造操作剖面——确定操作发起者

- 用户类
 - 经理、会计、售货员和仓库管理员
- 外部系统
 - 无，因为没有其它外部系统能够启动超市信息管理系统的操作
- 系统自身控制程序
 - 无
- 对于仓库管理员模式下的超市信息管理系统的操作发起者只有仓库管理员



构造操作剖面——选择表示法及创建操作列表

- 选择表示法
 - 由于操作比较简单，故选择表格表示法

- 创建操作列表

仓库管理员模式操作列表

操作发起者	操作
仓库管理员	添加库存信息
	删除库存信息
	查询库存信息
	添加仓库
	删除仓库
	查询仓库
	添加商品
	删除商品
	查询商品



构造操作剖面——确定出现率及发生概率

仓库管理员模式的操作出现率及发生概率

操作	出现率 (每天8小时的操作个数)	发生概率
添加库存信息	180	0.4147
删除库存信息	50	0.1152
查询库存信息	60	0.1382
添加仓库	5	0.0115
删除仓库	3	0.0069
查询仓库	6	0.0138
添加商品	30	0.0691
删除商品	20	0.0461
查询商品	80	0.1843
合计	434	1.0



生成测试数据——分析操作输入变量

操作“添加库存信息”输入变量分析

变量	类型	取值函数	
		取值范围	概率
仓库号	时间无关不确定离散型	有效仓库号	0.9
		无效仓库号	0.1
商品号	时间无关不确定离散型	有效商品号	0.9
		无效商品号	0.1
日期	时间无关不确定离散型	有效日期	0.85
		无效日期	0.15
库存量	时间无关不确定连续型	[0, 50000]	0.95
		其它	0.05



生成测试数据——数据生成

操作概率选取范围（仓库管理员模式）

操作	概率范围
添加库存信息	(0, 0.4147]
删除库存信息	(0.4147, 0.5299]
查询库存信息	(0.5299, 0.6682]
添加仓库	(0.6682, 0.6797]
删除仓库	(0.6797, 0.6866]
查询仓库	(0.6866, 0.7004]
添加商品	(0.7004, 0.7695]
删除商品	(0.7695, 0.8156]
查询商品	(0.8156, 1]



失效严重程度定义

失效严重程度	定义
1	用户不能进行一项或多项操作
2	用户不能进行一项或多项操作， 但是有补救办法
3	一项或多项操作中的小缺陷



测试执行并收集失效数据

本次测试共生成软件可靠性测试数据170个，发现失效14次。

系统可靠性测试失效数据（仓库管理员模式）

失效	累计失效时间（分钟）	失效严重程度
1	43	3
2	102	3
3	139	2
4	183	3
5	289	3
6	334	3
7	390	3
8	441	1
9	507	2
10	576	3
11	655	3
12	716	3
13	782	3
14	867	2



可靠性评估

- 由于失效时间基本符合指数分布，因此，根据指数分布可以计算不同失效严重程度的平均失效间隔时间
 - 严重程度为1级的MTBF₁:

$$MTBF_1 = T/n_1 = 867 \text{ 分钟}$$
 - 严重程度为1级和2级的MTBF₂₊:

$$MTBF_{2+} = T/n_{2+} = 216.75 \text{ 分钟}$$
 - 软件的MTBF:

$$MTBF = T/n = 61.9 \text{ 分钟}$$
 - 其中:
 - T=867分钟，为测试总时间
 - n₁=1失效严重程度为1级的失效总数
 - n₂₊=4失效严重程度在2级（含2级）以上的失效总数
 - n=14为本次测试的总失效数



要点

- 软件可靠性测试是为了分析和验证软件的可靠性定量要求而对软件进行的测试，按测试目的的不同可分为软件可靠性增长和验证测试。
- 常规的软件测试结果不能直接用来评估软件的可靠性。
- 软件操作剖面用来表示软件使用的统计规律，帮助分配开发和测试资源，提高开发和测试的效率。
- 操作剖面的构造包括确定操作模式、确定操作的发起者、选择表格或图形表示法、创建操作表、确定操作的出现率以及确定操作的发生概率。



要点

- 软件可靠性测试数据是按照软件操作剖面随机生成的。
- 输入变量是任何存在于系统外部、同时影响此系统的数据元素。
- 失效数据可分为完全失效数据和不完全失效数据，其中完全失效数据是指由每一次失效发生的时间构成的失效数据，不完全失效数据是指由各时间段内发生的失效次数构成的失效数据。



作业

- 请构造ATM模拟软件的操作剖面。涉及的概率根据经验给出。在此基础上，具体分析说明如何构造ATM模拟软件的可靠性测试数据。
- 后面进入ATM模拟软件质量PK阶段，最终将评选出若干优秀软件。

