第九讲 系统建模与仿真(2)

四、仿真

1. 仿真(模拟)(Simulation)概念（下棋、彩排、试验、网游）

1）定义

利用模型复现实际系统中发生的本质过程，并通过对系统模型的实验来研究存在的或设计中的系统.（优点：现实系统复杂、随机、不能解析；人机界面友好；假设条件柔性强；更易控制；快）

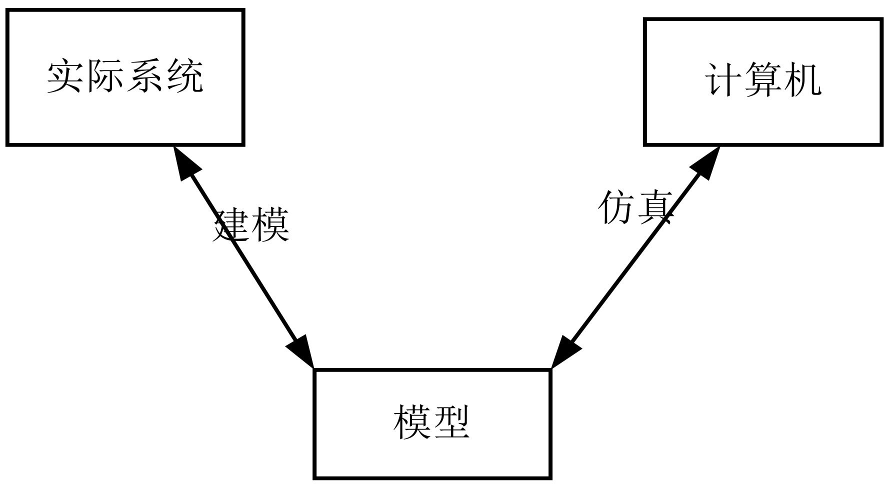
2）分类（根据模型类型划分）

物理仿真：即实物仿真，如风洞（研制飞行器的空气动力实验，真实感强、形象，但成本高、缺乏柔性）

计算机仿真(数学仿真)：模拟　数字　混合

半实物仿真：控制器(实物)+计算机上实现的控制对象（过控实验）

3）建模、仿真与计算机



建模与仿真的五个组成部分（实际系统、试验框架、基本模型、集总模型、计算机模型）

实际系统：（输入、输出、状态）行为描述（可观测变量、不可观测变量）

试验框架：假设或条件集合，同模型有效性（复制有效）之间相关

基本模型：在试验框架下，解释实际系统的行为

集总模型：基本模型的简化

计算机：复杂（仿真程序实现）

4）基本要素

- 对仿真问题的描述（模型：框架、参数；实验：环境、控制）
- 行为产生器（实验软件）
- 模型行为（指标、轨迹、结构）及其处理（分析、显示）

5）仿真的发展阶段

- 模型驱动的仿真（传统仿真）
- 含实物的仿真（物联网）
- 人在回路中的仿真（普适网格）

6）仿真的发展趋势

- 面向对象仿真（实体封装，如 UML）
- 定性仿真（关系判断、趋势分析）
- 智能仿真（知识库、专家系统）
- 分布交互仿真（P2P、网格）

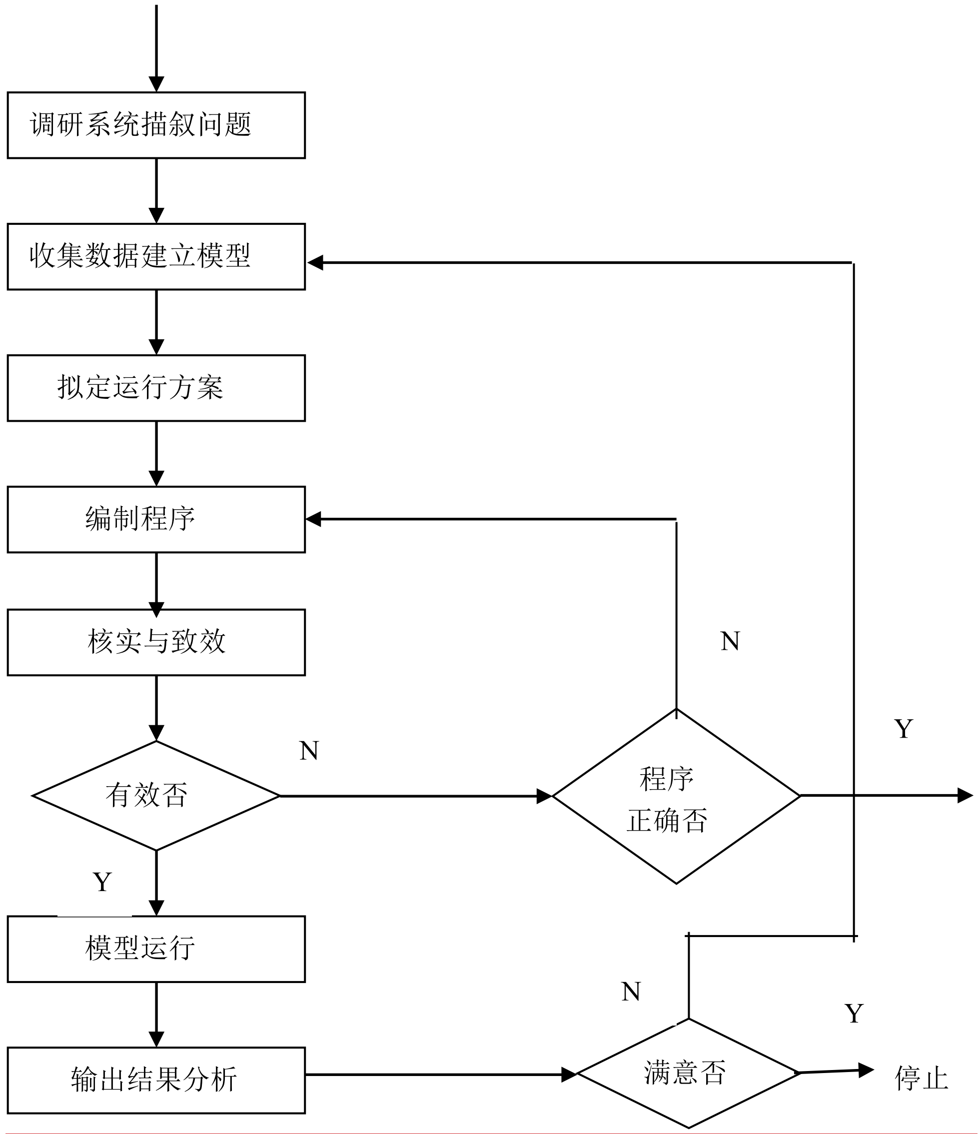
- 可视化仿真（决策剧场）
- 多媒体仿真（视觉、听觉模拟）
- 虚拟现实仿真（虚拟城市）
- Internet 网上仿真（B/S 分布式仿真）

7）仿真的对象（应用场合）

- 系统过于复杂（如存在过多的随机因素），难以采用解析法求解时，通过仿真可得到系统的动态特征。
- 系统实际运行费用过高或无法作实际运行时，借助仿真可以得到系统的有关参数。

优化设计（方案优化）、安全性和经济性（试验）、预测（天气预报）、完善系统模型（修正确认）、重复实验（收集数据、训练）

8）仿真的一般过程（面向问题：主要参数及影响、评价准则、系统边界、初始条件）



9）仿真的分类

- 物理仿真，模拟机仿真，数字仿真，数字机与模拟机混合仿真，仿真器仿真（Protel）
- 连续和离散系统仿真
- 静态和动态系统仿真
- 稳态和终态仿真（机器维修、商店营业）
- 确定性和随机性仿真

10）仿真的输出类型

- 确定型和随机型
- 连续观测值和离散观测值
- 连续分布和离散分布观测值
- 一元和多元输出
- 稳态型仿真和终止型仿真输出

11）仿真的局限性

1) 往往只能得到特解（可行解），而得不到通解（最优解）

2) 结果往往是间接的，而不是直接的（数据挖掘）

12）仿真的技术工具

连续系统仿真：DYNAMO（系统动力学）, CSMP（面向框图）

离散事件系统仿真：GPSS（通用系统模拟语言）, SIMSCRIPT（实体、属性、事件）, SIMULA（面向对象语言鼻祖）, GPSS-F

混合仿真：GASP-IV

2. 连续系统仿真（状态变量随时间连续变化）

1) 特点

- 微分方程（如果引入非线性因素，只能用仿真方法求解）

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \cdots, x_n, t)\ ; \qquad i = 1,2,\cdots\cdots,n$$

- 离散化

$$x_i(k+1) = f_i(x_1(k), x_2(k), \cdots\cdots, x_n(k), kT)\ ; \qquad i = 1,2,\cdots\cdots,n$$

- 误差和稳定性

$\Delta X = X - X_0$ 和步长 k

截断误差（近似解与准确解间的误差）和舍入误差（四舍五入带来的精度误差）

2) 仿真的主要内容

- 模型与实际系统的比较
- 系统的初态、暂态和终态
- 系统的扰动
- 系统的输入
- 求微分方程的特解或近似曲线

3) 分析的手段和工具

1）微分方程的离散化（步长 T 选择）

2）仿真计算（数值积分法）

- 欧拉法

$$x_i(k+1)=x_i(k)+Tf_i(x_1(k),x_2(k),\cdots\cdots,x_i(k),k)$$

例如：用欧拉法求下述微分方程的数值解。

$$\begin{cases}\dot{y}+y^2=0\\ y(0)=1\end{cases}$$

解：由上式可得

$$\dot{y}=-y^2$$

即

$$\dot{y}=f(y)=-y^2$$

由欧拉法递推公式，可知：

$$y_{k+1}=y_k+Tf_k$$

若取步长$T=0.1$，由$t=0$开始积分，即

$$\begin{cases}y_0=y(t=0)=1\\ f_0=f(y_0)=-y_0^2\end{cases}$$

则可得：

$$\begin{cases}y_1=y(t=0+T)=y(t=0.1)=y_0+Tf_0=1+(0.1)\times(-1^2)=0.9\\ y_2=y(t=0+2T)=y(t=0.2)=y_1+Tf_1=0.9+(0.1)\times[-(0.9)^2]=0.819\\ y_3=y(t=0+3T)=y(t=0.3)=y_2+Tf_2=0.819+(0.1)\times[-(0.819)^2]=0.7519\\ \quad\vdots\end{cases}$$

- 梯形法

$$x_i^0(k+1)=x_i(k)+Tf_i(x(k),k)$$

$$x_i^{j+1}(k+1)=x_i(k)+\frac{T}{2}\left[f_i(x(k),k)+f_i(x^j(k+1),k+1)\right]$$

其中，j=0,1,2,……

- 预报---较正法

$$x_i^0(k+1)=x_i(k)+Tf_i(x(k),k)$$

$$x_i^1(k+1)=x_i(k+1)=x_i(k)+\frac{T}{2}\left[f_i(x(k),k)+f_i\left(x^0(k+1),k+1\right)\right]$$

- 龙格---库塔法（泰勒级数展开，f 的线性组合代替 f 的高阶导数）

$$x_i(k+1)=x_i(k)+\frac{1}{6}\left[K_{1i}+2K_{2i}+2K_{3i}+K_{4i}\right] \qquad i=1,2,\cdots\cdots,n$$

$$K_{1i}=Tf_i(x(k),t_k)$$

$$K_{2i}=Tf_i(x(k)+0.5K_1,t_k+0.5T)$$

$$K_{3i}=Tf_i(x(k)+0.5K_2,t_k+0.5T)$$

$$K_{4i}=Tf_i(x(k)+0.5K_3,t_k+T)$$

- Adams 方法（线性内插和外推）
- Tustin 方法（双线性变换，对梯形法求 z 变换，$Z=e^{st}$）

$$S=\frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)$$

一般地，欧拉法、龙格---库塔法等适合于非线性系统的仿真；Adams 方法和 Tustin 方法适合于线性系统的仿真。

4) 噪声的生成(见下面“随机数发生器”)

5) 输出分析（数理统计、可视化、评价、指导决策）

6) 仿真语言或工具

CSMP（框图思想、结构语句、数据语句、控制语句）

3. 离散事件系统仿真（状态变量只在一些离散的时间点上变化）

0）问题举例

机修车间分为修理区和等待区，修理区每次只能修理一台机器。送修机器到达时，如修理区空闲，则直接进入修理区接受修理，修好后，由出口取走；如果修理区不空，则放在等待区排队待修。目前，此车间不能满足本厂的需要，据一年的统计知，机器平均等待时间为60天，平均逗留时间（等待时间加上修理时间）为75天，修理台利用率为0.98。工厂主管部门拟扩大修理区，再增加一台同样的修理台，以降低送修机器的等待时间，但又担心增加台数，会使修理台的利用率太低（如50%以下），而造成浪费。因此，想用仿真方法预测一下修理区扩大后的状况。

第一步，明确仿真目的

在机修车间问题中，仿真目的是统计计算现在系统和未来系统的平均等待时间、平均逗留时间和修理台利用率。

第二步，系统描述

（1） 系统组成成份

机修车间的系统成份可分为入口（输入过程）、等待区（排队）和修理区（服务过程）三部分。

（2） 描述变量

在入口，选用描述变量（u_i,t_i^1）表示送修机 u_i 于 t_i^1 时刻到达。

在等待区，用 $Q_1Q_2\ Q_3\dots\ \dots Q_M$ 表示排队，M 为队列的长度；（x_i, t_i^0）表示机器 x_i 于 t_i^0 时刻进入修理区，其中描述变量 t_i^0 表示机器 x_i 开始接受修理的时刻。

在修理区，用（x_i,t_i^2）表示机器 x_i 于 t_i^2 时刻修好并离去。

当有一台机器修好离去时，如队列长度 $M\neq0$，则 t_i^0 等于刚离去的那台机器的离去时刻 t_{i-1}^2；当有一台机器 u_i 到达时，如 M=0，则 t_i^0 等于这台机器的到达时刻 t_i^1。这样，描述变量(x_i, t_i^0)就是从属的，可省去。最后得到该系统的最小描述变量组为：

输入量（u_i,t_i^1）$t_i^1\in(0, 365)$ (单位：天)

状态$\begin{bmatrix} Q_1Q_2\cdots\cdots Q_M \\ (x_j,t_j^2) \end{bmatrix}$, $t_j^2\in(0, 365)$ (单位：天)，x_j 正在修理的机器

对于入口，假设在不相重叠的时区区间内机器到达数是相互独立的（无后效性），对充分小的 Δ_t,在区间[t, $t+\Delta_t$）内有一台机器到达的概率与 t 无关,而大约与区间长 Δ_t 成正比（平稳性），对于充分小的 Δ_t，在时间区间[t, $t+\Delta_t$）内有两台或两台以上机器到达的概率极小，可以忽略（普遍性），则在时间[0, t）有 n 台机器到达的概率为 $P_n(t)=(\lambda t)^n e^{-\lambda t}/n!$，即到达的机器服从泊松分布，其中 λ 表示单位时间平均到达的机器数。在上述假设下，一般机器到达的时间间隔 $T= t_{i+1}^1- t_i^1$，服从负指数分布（$F(t)=1-e^{-\lambda t}$，$t\geq0$），其密度函数为

$$f_T(t)=\begin{cases}\lambda e^{-\lambda t} & t\geq0\\ 0 & t<0\end{cases}$$

在仿真中 T 采用截尾指数分布。

在等待区，队列由 $Q_1Q_2Q_3\dots\ \dots Q_M$ 描述。采用先来先修理的排对

规则，即若又来了一台机器 u_i 要修理，队列将变成 $Q_1Q_2Q_3\dots\dots Q_Mu_i$。实际上还有按优先级修理等其他排队规则。同时为简单起见，假设等待区足够大，即队列长度不限。

在修理区，修理好一台机器所需时间 $T' = t_{i+1}^2 - t_i^2$ 也服从截尾指数分布。

（3） 参数

泊松参数分布 λ

（4） 相互关系

设当前时刻为 t，则可得 $t_{i+1}^1 = t_i^1 + T$， $t_{i+1}^2 = t_i^2 + T'$

当机器 u_{k+1} 在 t_{i+1}^1 时刻进入系统后，系统由当前状态 S_k 生成下一状态 S_{k+1}，其中 $S_k = \begin{bmatrix} Q_1Q_2\cdots\cdots Q_M \\ (x_j, t_j^2) \end{bmatrix}$

当 $t_{k+1}^1 < t_j^2$ 时， $S_{k+1} = \begin{bmatrix} Q_1Q_2\cdots\cdots Q_M u_{k+1} \\ (x_j, t_j^2) \end{bmatrix}$

当 $t_{k+1}^1 \ge t_j^2$ 时， $S_{k+1} = \begin{bmatrix} Q_2Q_3\cdots\cdots Q_M u_{k+1} \\ (x_1, t_1^2) \end{bmatrix}$， $M\ge 1$

$$S_{k+1} = \begin{bmatrix} \varnothing \\ (u_{k+1}, t_{k+1}^2) \end{bmatrix}, \quad M = 0$$

仿真目的是求平均等待时间、平均逗留时间和修理台利用率，但这些量不直接等于状态变量。据分析，这些量可由总和等待时间TWT，总和空闲时间 TFT 和到达机器总数 NT 换算出来，故尚需设置输出函数如下：

$$\begin{pmatrix} TWT_{k+1} \\ TFT_{j+1} \\ \\ NT_{k+1} \end{pmatrix} = \begin{pmatrix} TWT_k + M*(t_{k+1}^1 - t_k^1) & \text{状态每改变一次计算一次} \\ TFT_j + (t_{j+1}^1 - t_j^1) & t_j^1\text{时刻修理区为空，且在 } t_{j+1}^1 \text{ 时刻产生到达事件时计算} \\ \\ NT_k + 1 & \text{机器到达时计算} \end{pmatrix}$$

以单修理台为例，下表说明排队系统的仿真运行过程（令仿真终止时间为 240 天，平均到达速率和平均服务速率为 0.1，定义机器到达事件为 1 类事件，机器离去事件为 2 类事件，排队规则为先进先出）：

仿真时钟	事件类型	机器	到达时间	下一到达时间	修理台状态	队长	系统中机器数	修理开始时间	等待时间	修理时间	离去时间	逗留时间	已修理机器数
0	—	—	—	—	闲	0	0	—	—	—	—	—	0
0	1	1	0	7	闲→忙	0	1	0	0	10	10	10	0
7	1	2	7	25	忙	1	2	10	3	6	16	9	0
10	2	1	—	—	忙	0	1	—	—	—	10	—	1
16	2	2	—	—	忙→闲	0	0	—	—	—	16	—	2
25	1	3	25	26	闲→忙	0	1	25	0	5	30	5	2
26	1	4	26	28	忙	1	2	30	4	53	83	57	2
28	1	5	28	30	忙	2	3	83	55	34	117	89	2
30	2	3	—	—	忙	1	2	—	—	—	30	—	3
30	1	6	30	46	忙	2	3	117	87	12	129	99	3
…	…	…	…	…	…	…	…	…	…	…	…	…	…
236	2	13	—	—	忙	0	1	—	—	—	236	—	13
238	2	14	—	—	忙→闲	0	0	—	—	—	238	—	14
250	1	—	—	—	—	—	—	—	—	—	—	—	14

1）离散系统仿真

在离散系统仿真中，系统的状态只在随机的时间点上发生阶跃，而在两个时间点之间不发生变化。其特点有：

- 概率模型
- 拥挤现象和服务水平
- 统计分析

2）仿真的主要内容

- 统计参数（平均等待时间、平均队长、平均服务时间，等等）
- 稳定过程和非稳定过程
- 方案选择

3）仿真原理

蒙特卡罗 Mante – Carlo 方法论

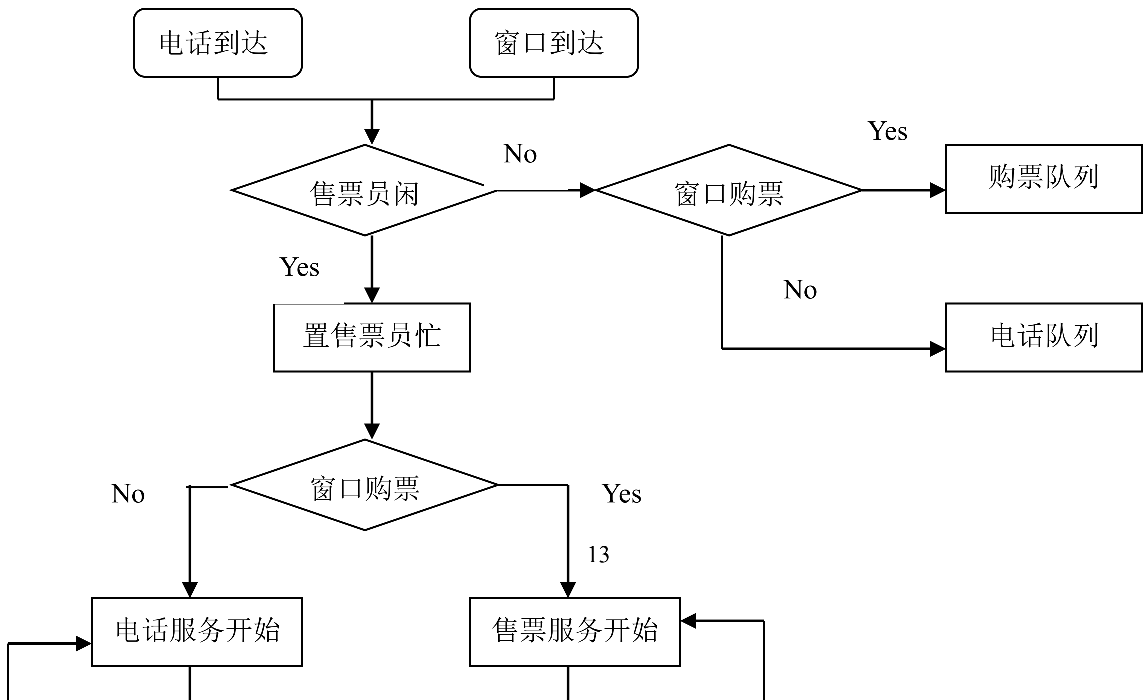
基本原理：

- 随机离散事件
- 仿真时钟及其推进方式（事件、时间间隔）
- 未来事件表
- 随机数发生器
- 采集和输出统计数据
- 事件安排/时间推进的仿真机制

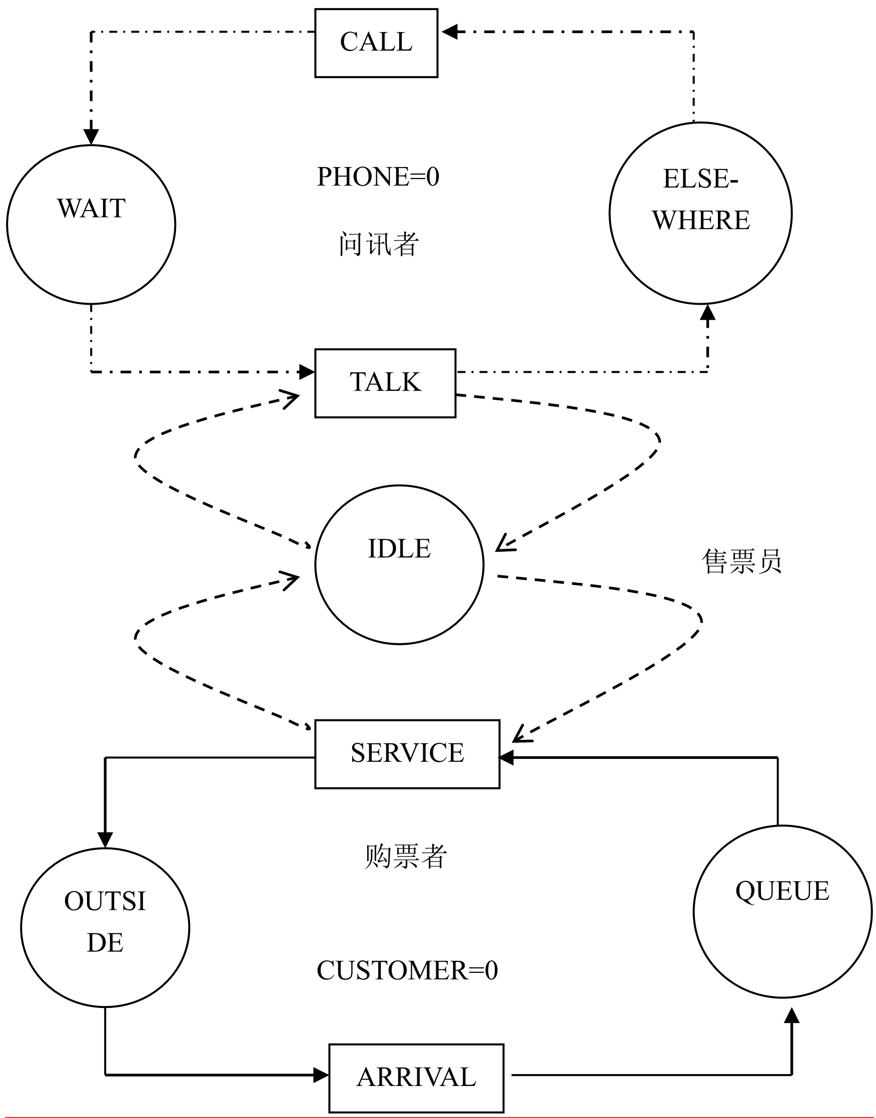
4）仿真方法

（1）建模方法：

- 实体流图法



- 活动周期图法



- Petri 网方法（条件并发、冲突诊断）

（2）仿真策略

- 事件调度法
- 活动扫描法
- 进程交互法

（3）仿真模型的计算机实现

面向事件仿真模型的实现

面向活动仿真模型的实现

5）分析的手段和工具

（1） 随机数产生器

A. **伪随机数发生器**

[0,1]均匀分布，可采用随机数表、物理方法（摇号）、数学方法。

- 中值平方法

例如，

$$x_0^2 = 76^2 = 5776, x_1 = 77, u_1 = 0.77$$
$$x_1^2 = 77^2 = 5929, x_2 = 92, u_2 = 0.92$$

……

- 中值乘积法（退化现象）

例如，种子数=5167，乘数=3729，种子数与乘数之积=19267743，产生的随机数=0.2677

- 线性同余法

$z_i = (az_{i-1} + c) \bmod m$，令 $u_i = z_i / m$

例如，取 m=16, a=5, c=3, z_0=7，则 $z_i = (5z_{i-1} + 3) \bmod 16$

于是，z_1=6, u_1=0.375

z_2=1, u_2=0.063

z_3=8, u_3=0.500

……..

B. **产生规定分布的随机变量**

离散事件仿真中常用的规定分布有负指数分布、均匀分布、正

态分布、对数正态分布、爱尔郎分布、β 分布、γ 分布、三角分布、韦伯尔分布、二项分布、泊松分布、经验分布，等。其基本原理是：

令 $F(x)$为 X 的分布函数，$G(y)$为 Y 的分布函数，且 $Y=F(X)$，则 $G(y)=p\{Y\leq y\}=p\{F(X)\leq y\}=p\{X\leq F^{-1}(y)\}=p\{X\leq x\}=F(x)=y$，所以 $g(y)=\dfrac{dG(y)}{dy}=1$，即 Y 在[0，1]区间内均匀分布。由此可根据逆变法（$X=F^{-1}(U)$）设计任意分布。（举例：负指数分布 $x=-\dfrac{1}{\lambda}\ln u$）

（2）结果分析

A. 性能测度及其估计（指标体系）

B. 终态仿真的输出分析（重复运行）

C. 稳态仿真的输出分析（分批分析）

D. 取得规定的置信区间（参数满足概率要求的取值空间）

（3）仿真语言

SLAM（混合仿真：微分方程、差分方程、事件和进程）、GPSS(面向框图)