

© Copyright by Zheng Wang  
All rights reserved.

课程编号: B008107

# 制造系统分析、控制 与设计

汪 峥

东南大学自动化学院  
2008年9月

## 教学周历

周次	教学内容
1	从控制的观念看制造系统
2	制造系统控制概述
3	基于仿真的制造系统分析与控制（1）
4	基于仿真的制造系统分析与控制（2）
5	基于 Petri 网的制造系统分析与控制（1）
6	基于 Petri 网的制造系统分析与控制（2）
7	基于 Petri 网的制造系统分析与控制（3）
8	基于排队网络与随机模型的制造系统分析与控制（1）
9	基于排队网络与随机模型的制造系统分析与控制（2）
10	基于排队网络与随机模型的制造系统分析与控制（3）
11	基于连续系统模型的制造系统分析与控制（1）
12	基于连续系统模型的制造系统分析与控制（2）
13	制造系统设计概述

# 第 1 章

## 从控制的观念看制造系统

本课程的内容包括制造系统的分析, 控制与设计三个部分, 实际上这是研究制造系统的三个基本步骤: 首先是分析一个已有的制造系统, 了解它的各种性质并对其性能进行定量的评估; 其次是对一个已有的制造系统进行控制, 以改善它的性能, 使之实现人们所期望的目标; 最后是设计并建造一个新的制造系统, 并对其进行控制, 使之实现预期的目标和要求。

控制是核心, 分析是控制的基础, 建模是分析的基础, 而设计是控制的延伸. 控制系统的一般表达形式:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$$

其意义可解释为: 控制量( $\mathbf{u}$ )施加在状态量( $\mathbf{x}$ )上, 在系统的动力学机制( $\mathbf{f}$ )的作用下, 所引起的状态的变化量( $\dot{\mathbf{x}}$ ).

从控制的观念来看, 对于制造系统也是如此, 需要了解系统的结构是怎样的? 其状态量是什么? 控制量是什么? 动力学机制是怎样的? 有哪些重要的性能指标? 应该如何控制?

### § 1.1 制造系统的结构

#### 1. 制造系统的构成元素

制造系统是一个很宽泛的概念.

狭义的制造系统: 加工设备(机床), 存储设备(缓冲区(buffer), 仓库(stock)), 物料运输设备(自动引导小车(automated guiding vehicle, AGV), 传送带(conveyer), 机器人(robot)), 零件(part), 控制系统, 工人, 刀具(cutting tools), 托盘(pallet)等. 例如: 生产流水线(flow line), 装配线(assembly line), 单元制造系统(manufacturing cell), 柔性制造系统(flexible manufacturing systems, FMS).

推广: 车间制造系统(shop floor). 增加了制造执行系统(manufacturing execution systems, MES)

再推广: 制造企业 --- 生产系统, 库存系统, 采购系统, 销售系统, ...

广义的制造系统: 敏捷制造系统(agile manufacturing systems), 虚拟企业(virtual

enterprises), 网络化制造系统(networked manufacturing systems), 物流(logistics)及供应链(supply chain)系统, ...

更广义的制造系统: 包含了产品的设计, 制造, 销售, 维护, 回收这一全生命周期(life cycle)的每个环节.

本课程着重讨论狭义的制造系统, 即车间内部的制造系统. 车间制造系统由以下主要元素构成:

1. 生产资源(生产能力, production capacity): 机器(machine), 工具(tool), 工人(operator)
2. 库存资源: 缓冲区(buffer), 仓库(storage, warehouse)
3. 物料操纵系统(material handling systems): 小车(automatically guided vehicle, AGV), 传送带(conveyer), 机器人(robot)

## 2. 制造系统的组织形式

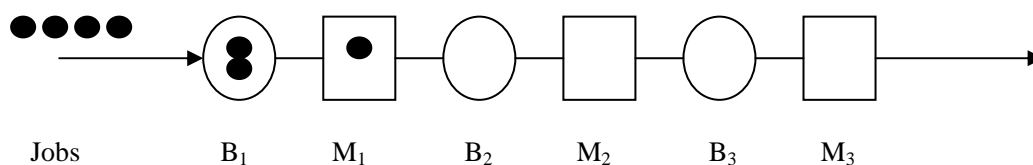
flow shop (流水线生产车间), job shop (离散作业车间), reentrant system (可重入制造系统).

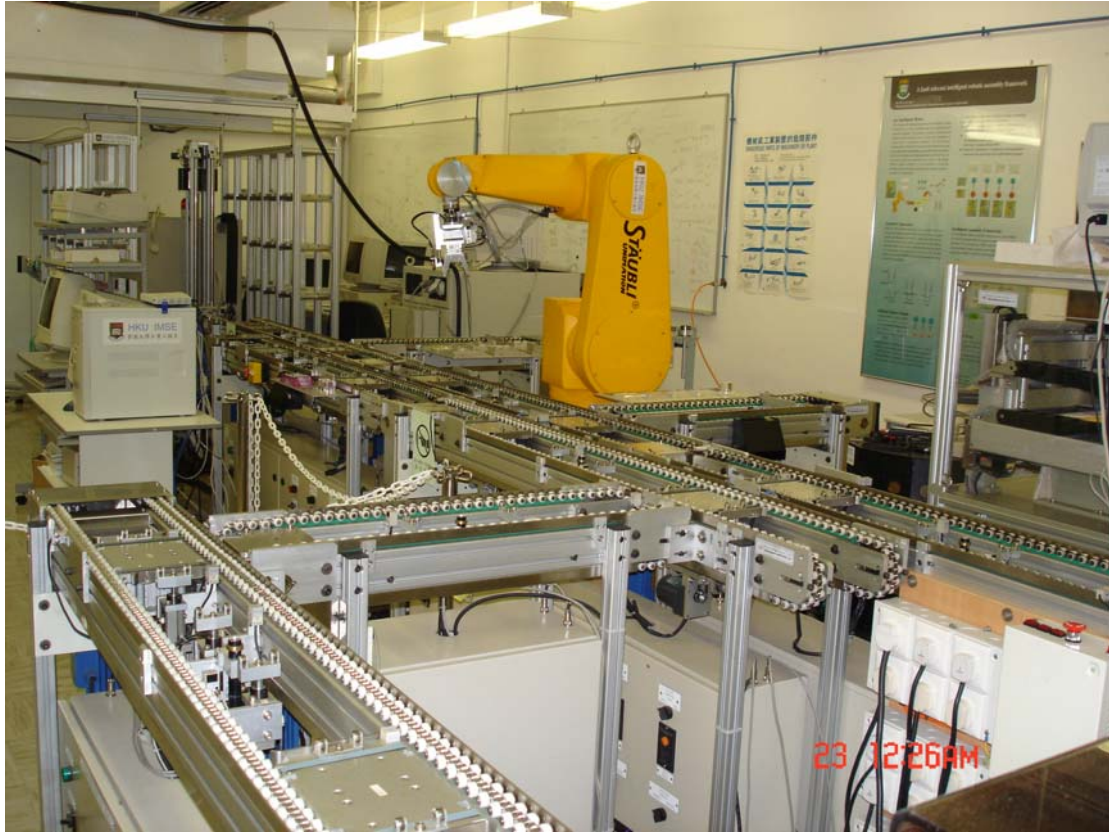
Flow Shop (流水线生产车间): 又称作 flow line, transfer line, production line

$N$  个机器或工位(workstation),  $M$  个零件(part, job), 每个零件都要经过  $N$  道工序, 每道工序在一台机器上加工, 没有重复(即多道工序在同一台机器上加工的情况), 也没有遗漏, 并且这  $M$  个零件在  $N$  个机器上的加工顺序都是一样的.

通常用于大批量(high volume), 少品种(low variety)产品的生产.

实例: 南汽总装厂跃进轻型卡车装配线, 山东临工 ZL50 装载机总装线, 上海柴油机厂某种型号的柴油发动机装配线





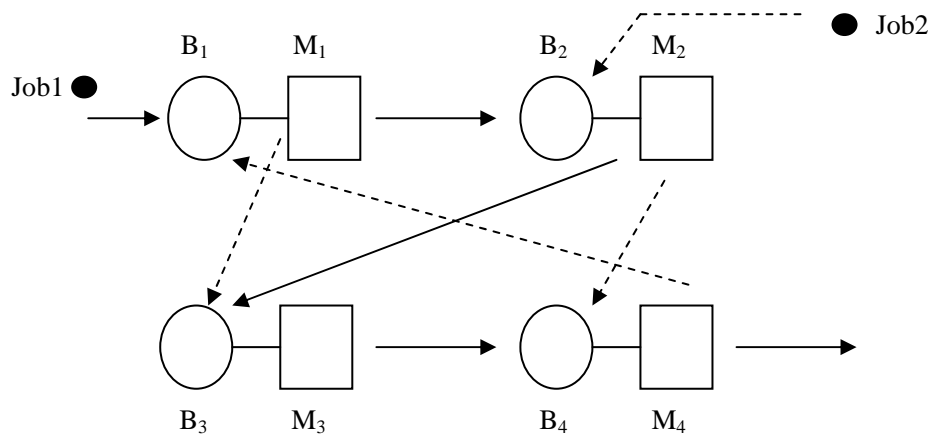
香港大学工业及制造系统工程系的柔性装配线实验装置

**Job Shop (离散作业车间):**

$N$  个机器或工位(workstation),  $M$  个零件(part, job), 每个零件都要经过不超过  $N$  道工序, 每道工序在一台机器上加工, 没有重复, 但有可能某个零件不经过某台机器, 并且这  $M$  个零件在  $N$  个机器上的加工顺序是不同的.

通常用于小批量(low volume), 多品种(high variety)产品的生产.

实例: 南汽车身厂冲压、装焊和油漆车间.





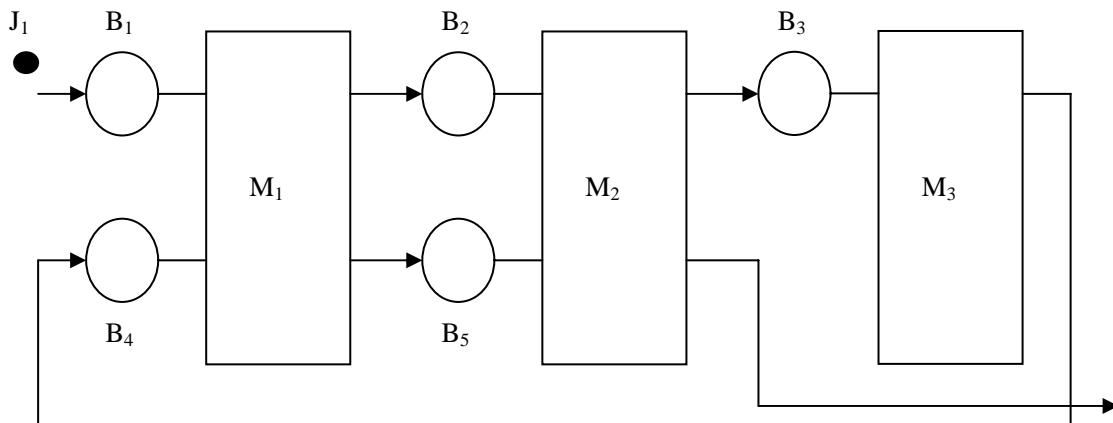
香港大学工业及制造系统工程系柔性制造系统实验装置

reentrant system (可重入制造系统):

$N$  个机器或工位(workstation),  $M$  个零件(part, job), 每个零件要经过可能超过  $N$  道工序, 每个零件的不同工序可能在一台机器上加工, 形成回流现象, 但也有可能某个零件不经过某台机器, 并且这  $M$  个零件在  $N$  个机器上的加工顺序是不完全相同的.

例如: 集成电路芯片晶圆制造系统(IC wafer fab).

实例: 上海中芯国际, 台积电, 华虹 NEC 的 6 英寸, 8 英寸, 12 英寸晶圆制造车间.

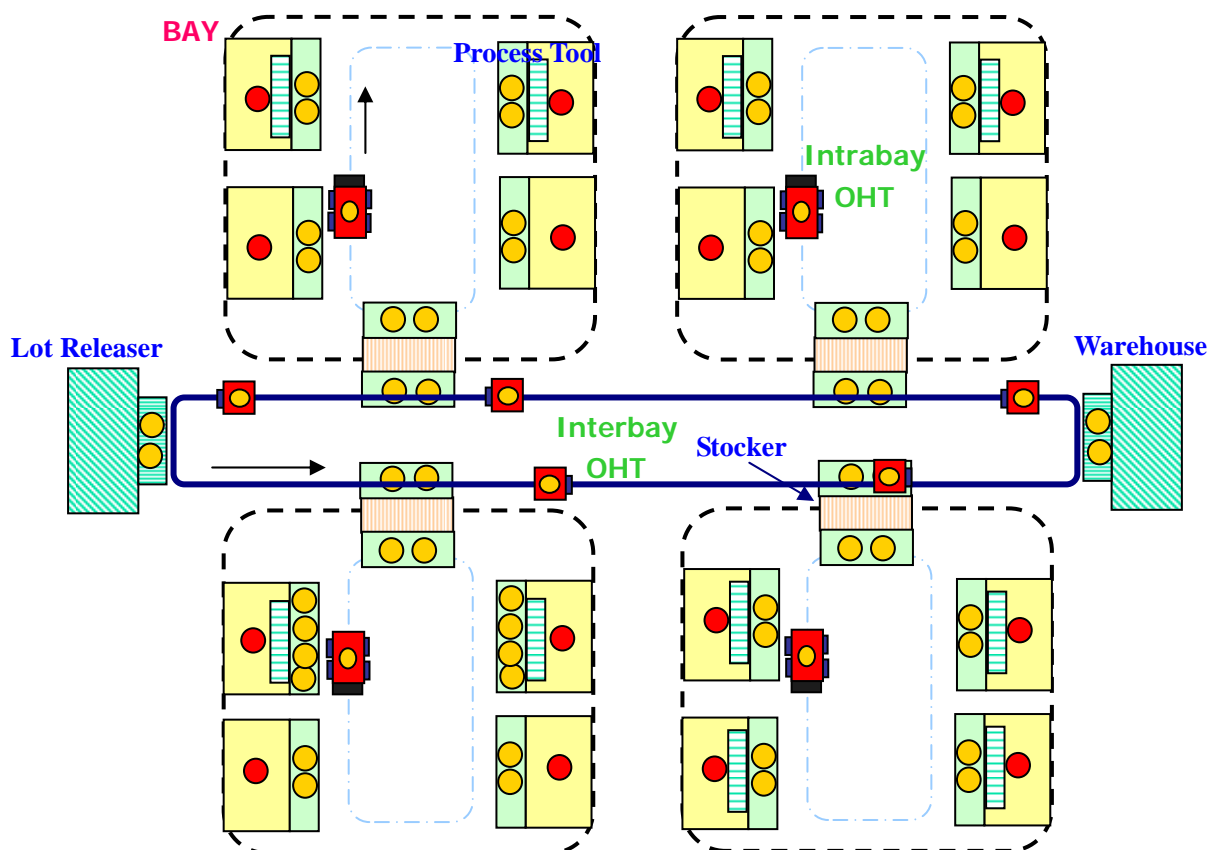


### IC wafer 的制造过程.

集成电路芯片的制造是制造业中的一个新兴行业，尤其在我国的内地，是1990年代后期才发展起来的，主要生产工厂集中于上海，例如：中芯国际，台积电，华虹 NEC 等。集成电路芯片的制造通常分为两部分：前道工序(front end)和后道工序(back end)，前道工序为晶圆的制造(wafer fabrication)，后道工序包括晶圆的切割，封装与测试。其中，wafer 的制造最为复杂。按照尺寸的不同，wafer 可分为6英寸，8英寸(200mm)，12英寸(300mm)，相应的制造系统有所不同。300mm wafer fab 的自动化程度最高。

集成电路晶圆制造系统有不同于传统机械制造系统的特点：

- 工序多---上百道工序，主要分为四类；沉积与扩散(deposition and diffusion) (or thin film)，照相(lithography)，蚀刻(etching)，离子注入(ion implanting)
- 每道工序时间差异大---从一、二十分钟到十几小时。
- 每道工序的工艺要求和生产方式不同---有的以单个 wafer 为单位，有的以 lot 为单位(一个 lot 由 12 或 24 个 wafer 构成)，有的以 batch 为单位(batch size 可能大于 lot size，也可能小于 lot size)。
- 设备多且价格昂贵---分为生产设备，存储设备，物料运输设备。一个 300mm wafer fab 有上百台的生产设备。
- 工艺路线复杂，一个显著的特点是可重入(re-entrant)的物流，即一种 wafer 的不同层的工序在同一台设备上加工，这给生产过程调度带来了很大困难。
- 经常有急件(hot lot)和返工(rework)出现。
- 设备容易发生故障，要安排预防性维护(preventive maintenance)



上图引自 'HiFiVE v3.0 --- A High-Fidelity Virtual Environment for 300mm Wafer Fabrication. by HiFiVE Project Team, Keck Virtual Factory Laboratory, Georgia Institute of Technology.

wafer fab 的布局: 一个 wafer fab 通常由多个 bay 构成. bay 相当于一个工段或制造单元, 通常工艺相近的设备(tool)构成一个 bay. 一个 bay 有一个或几个 stocker, 用于存放 wafer 或工具. 一个 bay 的内部有 intra-bay 的物料运输系统(material handling systems, e.g. AGV), 整个 fab 有 inter-bay 的物料运输系统.

## § 1.2 制造系统的状态量、控制量、动力学机制与性能指标

### 1. 状态量(states)

一个制造系统的状态量通常可分为如下几类: 机器状态, 库存状态, 工件状态.

机器状态: 机器可能处于正常工作状态(operational, up)或故障状态(failure, down); 在正常时可能处于工作状态(或忙状态 busy), 也可能处于空闲状态(idle). 一般地, 经常假定机器的正常工作时间是服从负指数分布的随机变量, 概率密度函数为

$$f(t) = \begin{cases} p \exp\{-pt\}, & \text{if } t \geq 0 \\ 0, & \text{if } t < 0 \end{cases}$$

数学期望为  $1/p$ , 其意义为机器从开始正常工作到发生故障的时间的平均值, 通常称作 MTTF(mean time to failure). 通常假定机器的故障修复时间也是服从负指数分布的随机变量, 只是参数为  $r$ , 数学期望为  $1/r$ , 其意义为机器从发生故障到故障排除之间的时间的平均值, 通常称作 MTTR(mean time to repair).

库存状态: 指的是每个机器上正在加工的, 或机器旁的缓冲区中存储的, 或物料运输系统正在搬运的, 或仓库中存储的各种工件的数量. 这些工件通常称作“在制品”(work-in-process, WIP). 而在制品的数量通常称作“在制品水平”(WIP level). 例如: 在上文的 job shop 的例子中, 可用  $n_{ij}$  表示在缓冲区  $B_i$  中存储的工件种类  $J_i$  的数量.

工件状态: 指的是每个工件当前处于哪一道工序, 并且处于制造系统中的什么位置, 例如: 在哪个机器上加工, 存储在哪个缓冲区里, 或是在哪个小车上或传送带上运输.

### 2. 控制量(controls)

控制量分为两类: 关于工件的控制量, 关于机器的控制量

关于工件的控制量: 主要指生产计划和生产调度.

生产计划(planning)是指在一定的时间范围之内, 每种工件该生产多少. 例如: 可以用  $u_{ij}(t)$  表示在时间周期  $t$  内, 机器  $i$  生产工件  $j$  的数量. 这里时间周期可以是 1



个班次, 1 天, 1 周或 1 个月, 也可以更短或更长. 当时间周期的长度尽可能短时,  $u_{ij}(t)$  也可称做生产率(production rate).

生产调度(scheduling)是指在一个时间周期之内, 在每台机器上各种工件的加工顺序. 例如: 某台机器旁边的缓冲区内有工件  $J_1, J_2, J_3$  和  $J_4$  等待加工, 该机器按照  $J_2 \rightarrow J_4 \rightarrow J_3 \rightarrow J_1$  的顺序进行加工, 则称  $J_2 J_4 J_3 J_1$  为一个调度. 如果这个顺序是事先排好的, 称为静态调度; 如果按照这个顺序是根据生产过程中的当前状态来决定的, 则称为动态调度, 或实时调度(real-time scheduling). 实时调度更具有控制量的意义. 实时调度通常根据一些规则来进行, 例如: 先到先服务(FCFS, first come first service), 最小交货期优先(EDD, earliest due date), 最短加工时间优先(SPT, shortest processing time)等.

注意: 计划解决的是“生产数量”问题, 调度解决的是“生产顺序”问题, 这是两者的区别. 然而, 当生产计划的时间周期短到与生产一个工件的时间接近时, 从一定程度上说, 计划与调度就统一了.

关于机器的控制量: 主要是指机器加工任务的实时分配以及机器维修的安排.

机器加工任务的实时分配是指当多个机器都可以加工某个工件时, 可以根据机器的状态把这个工件分配给其中 1 台机器. 例如: 可用  $v_{ij}(t) = 1$  表示在某个时刻  $t$  安排机器  $i$  加工工件  $j$ ;  $v_{ij}(t) = 0$  表示在某个时刻  $t$  不安排机器  $i$  加工工件  $j$ . 机器维修的安排是指在机器为发生故障时安排维修, 以避免以后可能发生的故障对生产造成的损失, 这种维修通常称为“预防性维修(preventive maintenance)”

### 3. 动力学机制(dynamics)

可以用连续时间系统模型, 离散时间系统模型和离散事件系统模型来刻画制造系统的动力学机制(dynamics).

实际上, 这是对制造系统的三类不同的描述方式, 是为了解决不同的问题而建立不同层面上的数学模型. 从微观层面上说, 制造系统是一个离散事件动态系统(discrete event dynamic systems, DEDS), 其状态的改变是事件驱动的(event-driven); 但是从宏观层面上看, 为了分析制造系统的一些总体特性, 并在总体上对其加以控制, 制造系统也可以看作是连续变量动态系统(continuous variables dynamic systems, CVDS), 包括连续时间系统和离散时间系统, 其共同特点是系统的状态变化是时间驱动的. 为制造系统建立什么样的模型, 完全取决于要解决什么样的问题, 根据问题来决定采取合适的办法描述与之相关的 dynamics.

例 1. 制造系统的离散时间动态系统模型

$$x_i(t+1) = x_i(t) + u_i(t) - d_i(t)$$

这里  $x_i(t)$  表示生产周期  $t$  开始时产品  $i$  的在制品数量;  $u_i(t)$  表示生产周期  $t$  内产品  $i$  的生产数量;  $d_i(t)$  表示生产周期  $t$  内对产品  $i$  的需求数量.

例2. 制造系统的连续时间动态系统模型

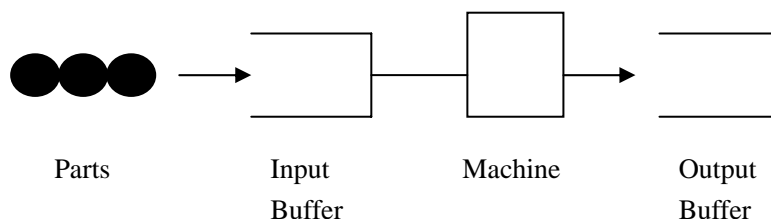
$$\frac{dx_i(t)}{dt} = u_i(t) - d_i(t)$$

这里  $x_i(t)$  表示  $t$  时刻产品  $i$  的在制品数量;  $u_i(t)$  表示  $t$  时刻产品  $i$  的生产率;  $d_i(t)$  表示  $t$  时刻对产品  $i$  的需求率.

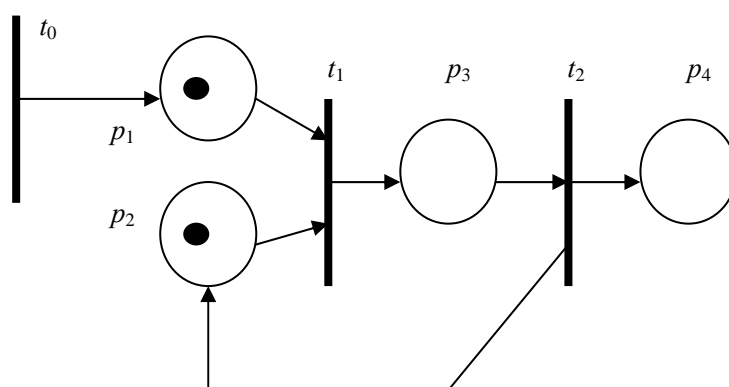
例3. 制造系统的离散事件动态系统模型---Petri 网

Petri Nets --- 在逻辑层面上描述离散事件动态系统的理论与方法.

如一个单机器制造系统:



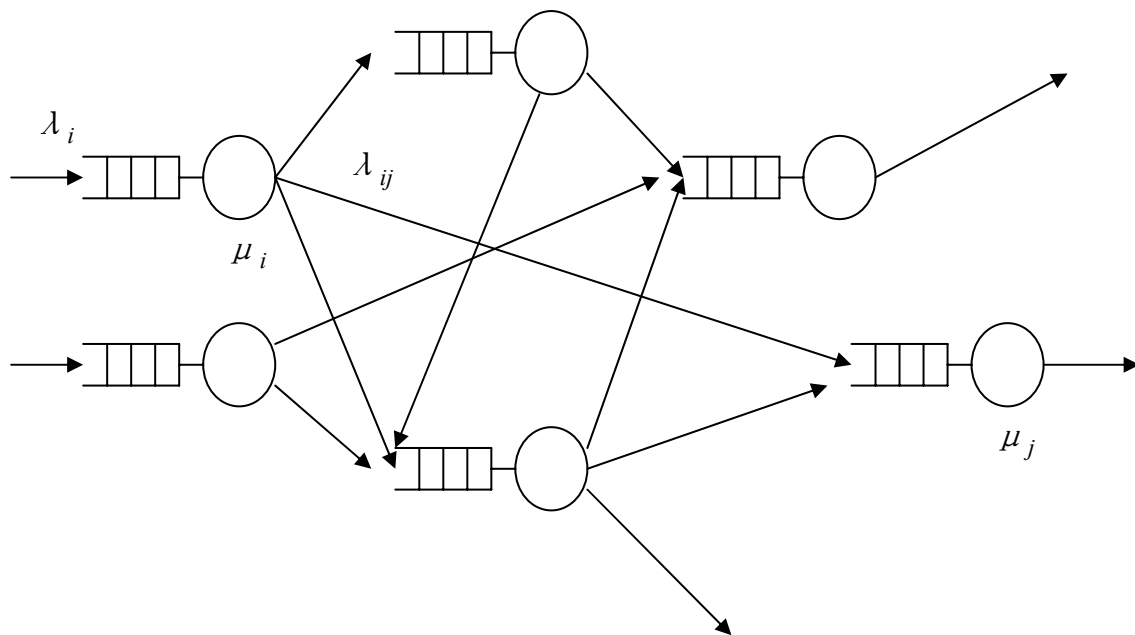
可建立 Petri 网模型如下:



其中, 库所  $p_1$  表示零件处于等待加工的状态; 库所  $p_2$  表示机器处于空闲状态; 变迁  $t_0$  表示产生一待加工的零件; 变迁  $t_1$  表示零件加工开始这一事件; 库所  $p_3$  表示机器处于正在加工零件的状态; 变迁  $t_2$  表示零件完成加工这一事件; 库所  $p_4$  表示零件处于正在加工已经完成的状态. 黑点(token)表示系统当前状态.

例4. 制造系统的离散事件动态系统模型---排队网络

令  $\lambda_i$  表示零件的到达率;  $\mu_i$  表示机器的生产率;  $\lambda_{ij}$  表示零件在机器 I 上加工完成一道工序后转移到机器 j 进行下一道工序加工的转移率. 则一个 job shop 可用下图所示的排队网络表示. 对于一些特定形式的排队网络, 例如 Jackson 开环和闭环排队网络, 其稳态概率分布的解析形式的表达已经获得, 并且据此获得了一些基本的性能指标, 例如平均队长, 平均等待时间, 平均忙期, 零件在系统内的平均停留时间等; 对于一般形式的排队网络, 可用近似计算的方法获得其稳态概率分布和性能指标.



#### 4. 性能指标(performance criteria)

制造系统的主要性能指标包括：生产率，在制品水平，设备利用率或生产负荷，生产节拍，生产周期，...

**生产率(production rate, throughput):** 制造系统单位时间内生产某种产品的平均数量，通常用于描述 flow shop 型的装配线。

**在制品水平(work-in-process, WIP):** 每个缓冲区中在制品的平均数，以及相应的存储费用或欠产惩罚费用。

**设备利用率(utilization)或负荷(workload):** 每个机器、工具或工位用于生产时间和总的可用时间(available capacity)之比。

**生产节拍(cycle time):** 相邻两个产品完成生产并离开生产线的的时间间隔。

**生产周期(sojourn time):** 某个产品从进入制造系统到完成生产离开该系统所需要的时间。

确定了制造系统的状态量，控制量，动力学机制和性能指标后，就可以建立制造系统的模型，并在模型的基础上对制造系统进行分析与控制。

分析的方法可分为两类：解析法，仿真法。

1. 解析法通常是指在考虑制造系统生产过程中的随机因素(如机器故障，需求波动)的情况下，利用排队论等数学理论，获得其性能指标的封闭形式的解(closed-form solution)。

2. 仿真法通常指在一些专门的制造系统仿真软件的支持下，建立所要分析的制造系统的模型，运行这一模型，并利用仿真软件提供的统计分析功能获得对制造系统的性能指标。

注意：在仿真软件提供的环境下，对制造系统建模比较直接；但另一方面，

由于建模者没有在更抽象的层面上刻画制造系统的动力学机制，因此对于动力学机制比较复杂的制造系统，在仿真软件提供的环境下建模未必容易和准确。所以，在这些情况下，应先用 Petri 网建立逻辑层面的模型，再转换成仿真软件环境下的模型。

对制造系统的控制可分为三类：

1. 在连续时间动态系统模型的基础上，获得制造系统的最优控制，这是一个状态反馈控制律。
2. 在离散时间动态系统模型的基础上，获得生产计划，一般是根据月计划或周计划来制订天计划或班计划。生产计划是一个开环控制问题，获得的结果通常是静态的，而不是上述动态的反馈控制律。
3. 在离散事件动态系统模型的基础上，获得实时的生产调度。

制造系统的设计：

设计就是要构造一个制造系统，使得它在满足生产的需求的前提下成本最低。制造系统的设计通常包含以下内容：

生产能力规划(capacity planning)，即选择合适的机器类型和数量；

单元成组(cell formation)，即选择合适的机器构成加工单元；

设备布局(facility layout)，即把设备在车间内合理摆放，使得物流顺畅，花在零件运输上的时间少，而且充分利用车间场地面积；

装配线平衡(assembly line balancing)，即通过把装配工序分配到装配工位上，使得每个工位上的装配工作量（通常用时间来度量）尽量接近。

工艺路线规划(routes planning)，即对于某种产品有多个可选加工路线的情况，选择合理的加工路线，使得生产率能够满足，设备的生产能力能够得到充分利用。

缓冲区分配(buffer allocation)，即在各机器或工位之间合理地设置缓冲区的容量，既保证生产率，同时在制品库存也不至于过大。

## § 1.3 计算机集成制造

计算机集成制造(computer integrated manufacturing, CIM)

概念的提出: Joseph Harrington 1973. ---- 两个基本观念

1. 信息的观念. 产品的制造过程可看作信息的演化过程，包括信息的产生，收集，存储，处理，分析，... 产品是信息的载体，是信息的物化。
2. 集成的观念. 大量的信息不是孤立的，对信息处理也不是孤立的，而是要把它当作一个整体。并且这个信息的整体不是静止的，而是在不断演化。

在这一概念的基础上，人们在企业内部建立了计算机集成制造系统(computer integrated manufacturing systems, CIMS) --- 是主要用于制造业的综合自动化系统。它在计算机网络和分布式数据库的支持下，把产品的设计、制造和管理等方面的信息和功能集成起来，优化和利用企业内部及外部的经营、生产过程和各种资源，从而缩短产品开发周期，提高质量，降低成本改善服务，提高企业的生产能力、

应变能力和市场竞争能力.

机床制造企业 CIMS 体系结构:

(1) 管理信息分系统

机床制造业产品结构复杂, 制造工艺复杂, 所需设备和工艺装备繁多, 因此需要以企业资源计划(enterprise resource planning, ERP)为核心, 对企业经营决策、生产计划、生产技术准备、销售、供应、财务、成本、设备、工具、人力资源等进行有效的管理, 以达到缩短产品生产周期(加快对用户或市场的响应速度)、提高产品质量、降低生产成本、改善服务水平的目标(TQCS).

TQCS 同时也是下述各子系统的目标, 实际上, 它是企业实施 CIMS 的目标.

(2) 产品设计与工艺设计自动化分系统

产品的概念设计, 方案设计, 详细结构设计, 工程分析, 工艺设计, 数控编程, 以及产品数据管理和产品生命周期管理 ---CAD/CAE/CAPP/CAM ---PDM/PLM. 其中, PDM/PLM 是 CAD/CAE/CAPP/CAM 的集成平台. 常用的开发平台: UG, ProE, Catia, SolidWorks, Metaphase, TeamCenter.

(3) 制造自动化分系统

数控机床, 加工中心, 自动引导小车, 立体仓库等制造系统底层设备的控制系统. 他们通过计算机网络与管理信息分系统和产品及工艺设计分系统相连, 进行数据的采集, 传输, 以及生产指令和加工程序的传达.

(4) 质量与可靠性分系统

质量数据采集, 质量评价, 质量控制与管理. 方法: 统计质量控制(statistical quantity control, SQC), 统计过程控制(statistical process control, SPC), ...

(5) 计算机网络分系统

CIMS 各子系统相互通讯的平台---- Lan/Intranet/Internet.

(6) 数据库分系统

CIMS 中数据共享的基础. 数据库的设计要在整个企业的范围内从逻辑上统一考虑, 设计基表, 保证数据的唯一性, 安排合理的冗余度, 要保证系统的安全性, 设定用户的不同权限. 建立统一、合理、安全、可靠的数据库系统是成功实现 CIMS 信息共享的技术保证.

北京第一机床厂计算机集成制造系统的体系结构  
(BYJC-CIMS)

用户接口		通信接口	
工程设计分系统 EDIS	管理信息分系统 MIS	制造自动化分系统 MAS	质量管理控制分系统 CAQ
操作系统(OS), 网络软件(通讯协议), 数据库管理系统(DBMS)			
计算机, 网络			

### 工程设计分系统(EDIS)

零件设计	装配设计	夹具设计	运动仿真	动静刚度模拟分析	产品目录生成	加工工艺	装配工艺	工具表	数控程序生成	刀具清单生成
CAD			CAE		PDM	CAPP		CAM		
数据库管理系统 Oracle										
网络通讯软件										
操作系统 UNIX										
服务器, 工作站, 网络										

### 管理信息分系统(MIS)

经营决策子系统	销售管理	产品数据管理	库存管理	主生产计划	物料需求计划	车间作业管理	采购管理	成本管理	财务总帐	
经营子系统		生产计划与物资管理子系统					财务子系统			
数据库管理系统 Oracle										
操作系统, 网络通讯软件										
服务器, 工作站, 网络										

### 制造自动化分系统(MAS)

作业调度	刀具管理	NC 程序管理	主轴单元	套筒单元	丝杆单元	.....	NC 程序加载和反馈	加工过程仿真
FMS 车间管理系统			轴套零件 GT 生产系统			数控机床 DNC 系统		
自行开发车间控制器软件			自行开发管理软件			自行开发控制软件		
数据库管理系统 Oracle, 操作系统								
计算机, 网络								

## 质量管理与控制分系统(CAQ)

设计 质量 标准	设计 质量 评审	检验 规程 编制	现场 数据 采集	统计 分析	质量成本 规划	质量成本 核算	台帐	检修 计划
设计质量管理		制造质量管理		质量成本管理		计量器具管理		
数据库管理系统 Oracle								
操作系统, 网络通讯软件								
计算机, 网络								

北京第一机床厂 CIMS 应用工程是国家 863/CIMS 主题的重点项目, 由北京第一机床厂和东南大学合作完成, 1995 年度获颁“制造工程师学会工业领先奖(SME Industry Leading Awards)”。

关于先进制造技术的一些观念:

柔性制造系统(flexible manufacturing systems, FMS)

敏捷制造(agile manufacturing, AM)

虚拟企业(virtual enterprise, VM)

准时生产(just in time, JIT): TOYOTA

精益生产(lean production, LP): MIT, “*The Machine That Changed the World*”

虚拟制造(virtual manufacturing)

并行工程(concurrent engineering, CE)

网络化制造(networked manufacturing)

可重构制造(re-configurable manufacturing)

绿色制造(re-manufacturing)

经营过程重构(或业务流程再造)(business process reengineering, BPR)

## § 1.4 一些新技术在制造系统中的应用: 虚拟现实技术与射频识别技术

### ● 虚拟现实技术 (Virtual Reality)

虚拟现实设备:

1. 高性能的计算机。应具备强大的数值计算能力和图形图像处理能力, 例如 SGI 的专用图形工作站。
2. 虚拟现实软件系统。主要包括图形图像建模系统, 科学与工程计算可视化系统。
3. 虚拟现实外部设备。主要用于提供便捷直接的人机接口, 便于人对虚拟系统中的对象进行直接的操纵, 以提供给用户很强的沉浸感. 例如: 立体

眼镜、头盔显示器、数据手套等。

例：香港理工大学工业及系统工程系数码工厂实验室

VR 系统：



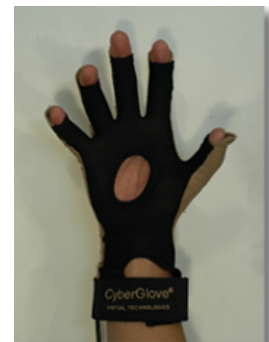
VR 设备：



SGI 图形工作站



立体眼镜



数据手套



Baron 投影仪



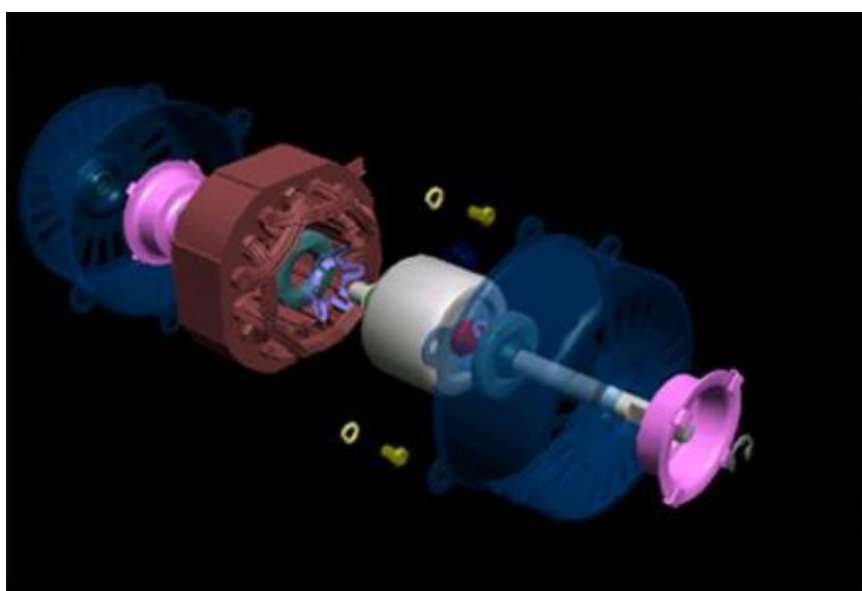
头盔显示器



研究工作:



制造系统仿真



虚拟装配

● 射频识别技术 (Radio Frequency Identification, RFID)

射频识别技术是一种通过附着在产品上的带有天线的芯片来发射包含产品信息的无线电信号, 用安装在一定范围内的解读器接收这些信号, 读取产品的信息, 并通过计算机网络来查询、跟踪和处理这些产品信息的技术. 在国际上, 1999 年, 剑桥大学、麻省理工学院等研究机构联合组建了“自动识别中心(Auto-ID Center)”, 对射频识别技术及其应用进行了广泛而深入的研究. 在国内, 复旦大学作为 Auto-ID 中心的成员之一, 在射频识别芯片的设计和制造方面做了大量工作. 香港科技大学浙江先进制造技术研究所也在从事这一技术在物流和供应链方面的应用研究. 射频识别技术的特点包括:

(1) 用解读器在一定范围内读取射频芯片发射的包含产品信息的无线电信号,大大提高了信息读取速度.

(2) 以电子产品编码(electronic product code, EPC)来标识每一个产品个体,射频芯片把 EPC 发射给解读器,解读器把 EPC 传输给客户端计算机.客户端计算机可通过 Internet 查询该产品的信息.

(3) 用物理标识语言(physical markup language, PML)描述产品的详细信息,存储并运行于 PML 服务器上.

(4) 用对象名服务器(object name server, ONS, 相当于通常的域名服务器 DNS)来解析电子产品编码,引导客户端计算机定位于存储产品信息的 PML 服务器(相当于通常的 Web 服务器).

(5) Auto-ID 中心开发了用于射频识别信息处理的软件系统 Savant, 它是一种中间层的分布式产品信息处理软件,可用于数据校正、解读器协调、数据传输、数据存储和任务管理,并且可以实现与其他生产管理系统的集成.

(6) 射频识别技术的发展目标是实现与“信息的互联网(Internet of Bits)”相对应的“物质的互联网(Internet of Things)”.

射频识别技术已经在零售、物流等行业得到了应用.在制造行业,Auto-ID 中心的科学家们也已经实现了应用射频识别技术对车间制造系统的控制.有关这方面的研究,参见剑桥大学 Auto-ID 实验室的工作:

Hodges, S., Thorne, A., Garcia, A., Chirn, J. L., Harrison, M., and McFarlane, D., 2002. Auto-IDBased Control Demonstration Phase 1: Pickand Place Packing with Conventional Control. *CAM-AUTOID-WH-006*.

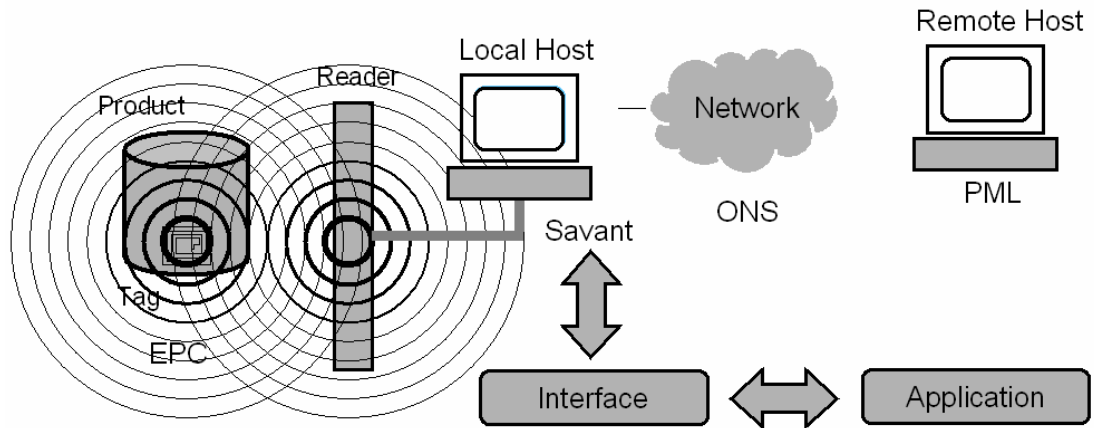
亦可参考香港大学与西安交通大学合作的工作:

Huang, G. Q., Zhang, Y. F. and Jiang, P. Y., 2007. RFID-based wireless manufacturing for real-time management of job shop WIP inventories. *International Journal of Advanced Manufacturing Systems*.

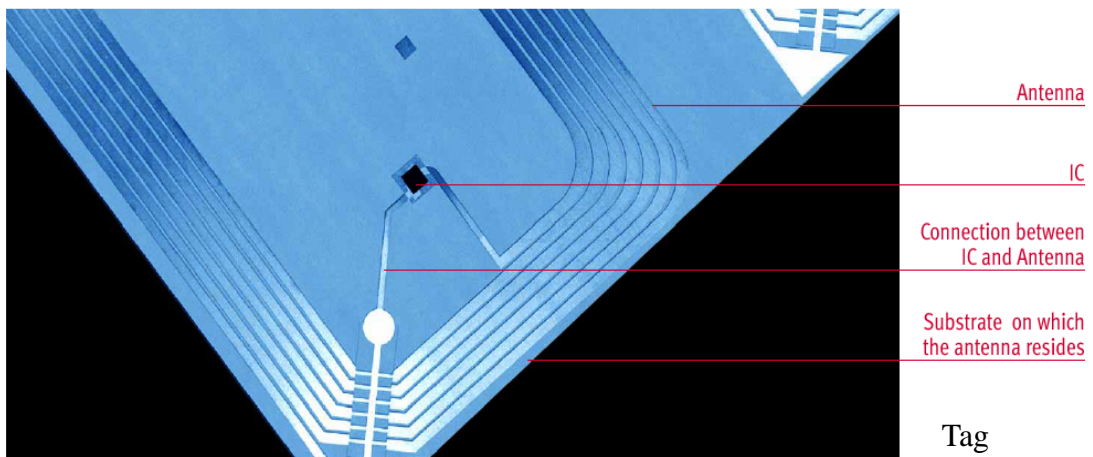
Huang, G. Q., Zhang, Y. F. and Jiang, P. Y., 2007. RFID-based wireless manufacturing for walking-worker assembly islands with fixed-position layouts. *Robotics and Computer-Integrated Manufacturing*, Vol.23, pp.469–477.

国家高技术研究发展计划(863 计划)在 2006 年开始资助关于射频识别技术的重大专项的研究工作。

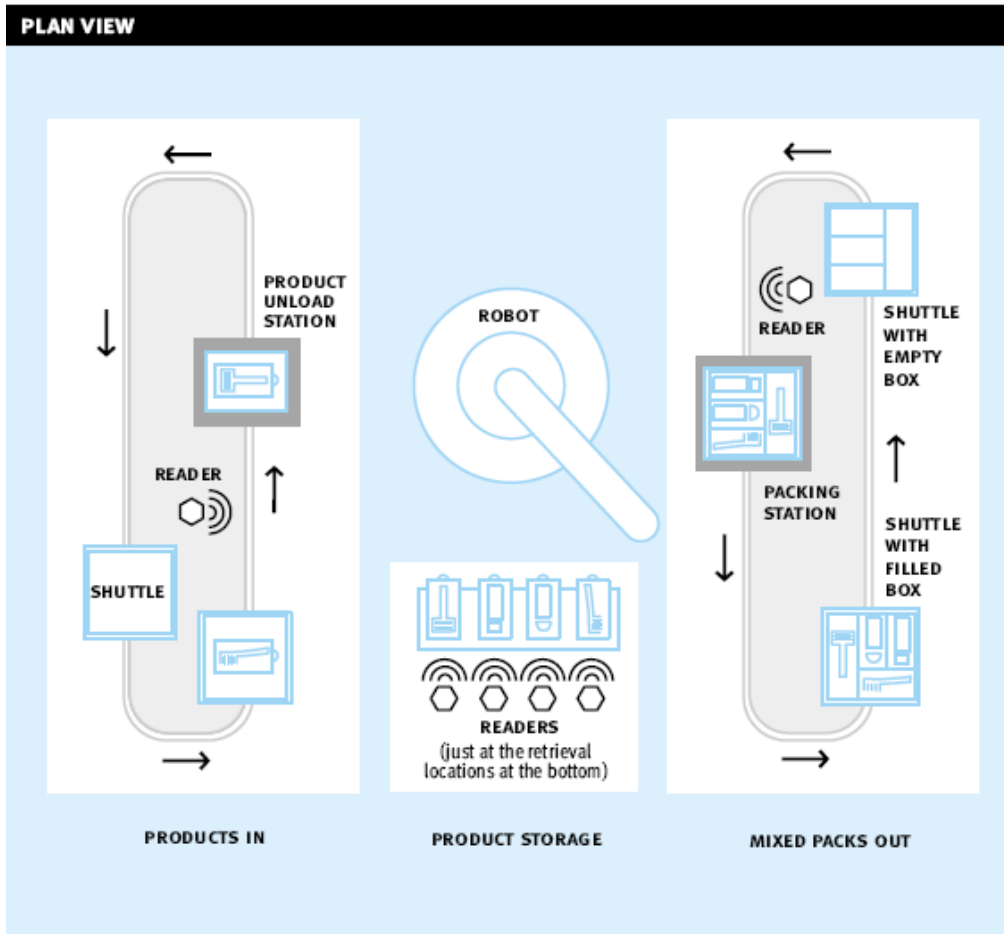
- **TAGS** RFID
- **EPC** Electronic Product Code
- **ONS** Object Name Service
- **PML** Physical Markup Language
- **Savant™** Distributed Operating System



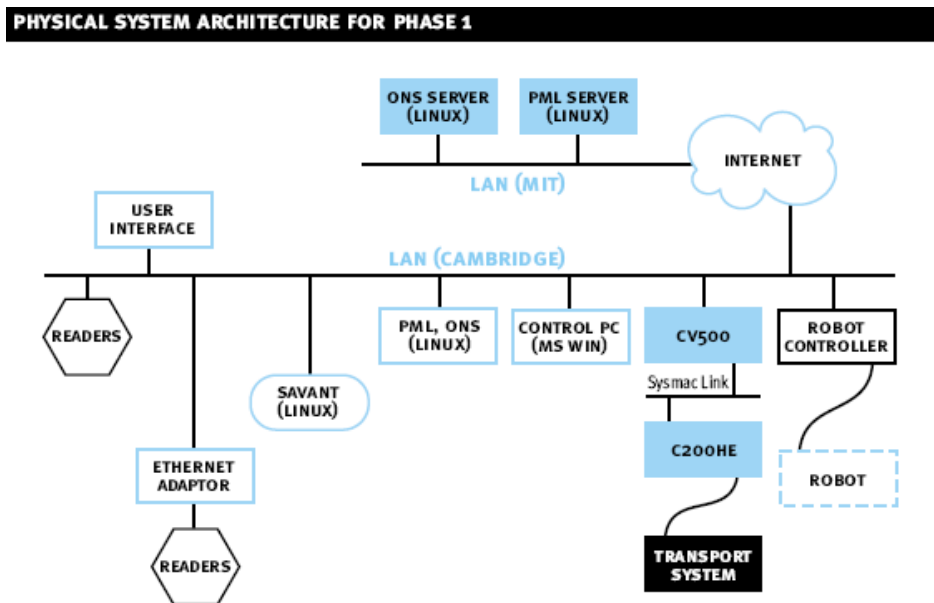
RFID 系统的结构



Readers

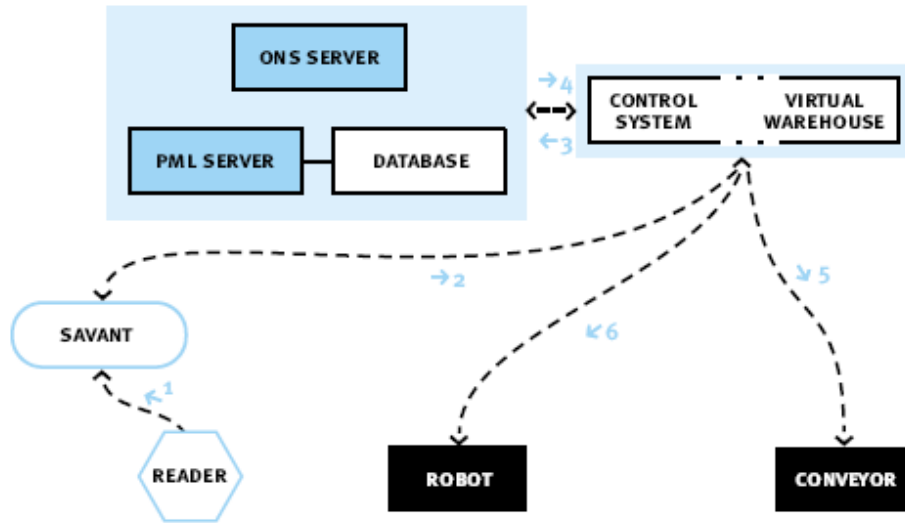


射频识别技术在制造系统控制中的应用：零件的识别、存储与装配



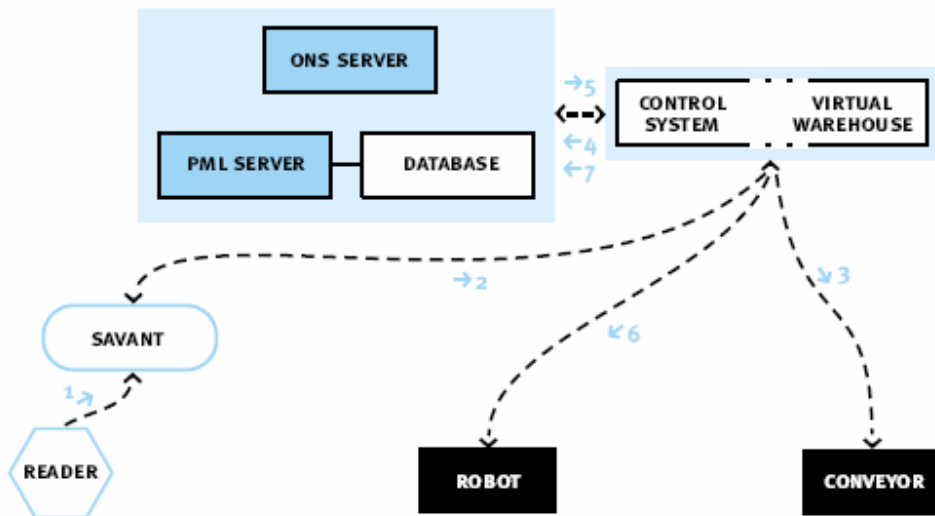
基于 RFID 的零件识别、存储与装配控制系统结构

**INFORMATION FLOW FOLLOWING THE ARRIVAL OF AN ITEM CARRIER**



到达系统的零件的控制信息流程

**INFORMATION FLOW FOLLOWING THE ARRIVAL OF A GIFT BOX**



零件装配过程控制信息流程

上述各图引自：

1. Brock, D., 2004. Beyond ebXML: Real-time logistics and telemetry. *Presentation Document, Auto-ID Center, Massachusetts Institute of Technology.*
2. Hodges, S., Thorne, A., Garcia, A., Chirn, J. L., Harrison, M., and McFarlane, D., 2002. Auto-IDBased Control Demonstration Phase 1: Pickand Place Packing with Conventional Control. *CAM-AUTOID-WH-006.*

## 附录: 主要著作, 期刊, 研究机构和研究者

### 重要著作:

1. Hopp, W. J. and Spearman, M. L. *Factory Physics*. (工厂物理学(影印版)). 北京: 清华大学出版社, 2002.
2. Gershwin, S. B. *Manufacturing Systems Engineering*. Prentice Hall, Inc., 2002.
3. 郑大钟, 赵千川. *离散事件动态系统*. 北京: 清华大学出版社, 2002
4. Womack, J. et al. *改变世界的机器(The Machine That Changed the World)*. 北京: 商务印书馆, 1999.
5. Womack, J. et al. *精益思想(Lean Thinking)*. 北京: 商务印书馆, 2000.

### 国际期刊:

1. IEEE Transactions on Automatic Control
2. IEEE Transactions on Systems, Man and Cybernetics (Part A, Part B, and Part C)
3. IEEE Transactions on Engineering Management
4. IEEE Transactions on Semiconductor Manufacturing
5. IEEE Transactions on Automation Science and Engineering
6. Operations Research
7. Management Science
8. Automatica
9. IIE Transactions
10. International Journal of Production Research
11. European Journal of Operational Research
12. Computers and Industrial Engineering
13. Computers and Operations Research
14. Annals of Operations Research
15. Robotics and Computer Integrated Manufacturing
16. International Journal of Production Economics
17. International Journal of Computer Integrated Manufacturing
18. Computers in Industry
19. International Journal of Advanced Manufacturing Technology
20. Concurrent Engineering: Research and Applications
21. Production and Operations Management

### 国际学会:

IEEE: Institute of Electrical and Electronic Engineers.

IIE: Institute of Industrial Engineers.

INFORMS: Institution for Operations Research and Management Sciences.

ASME: American Society of Mechanical Engineers

工业工程, 机械工程, 系统工程与运筹学领域的重要研究机构:

Massachusetts Institute of Technology (MIT):

Center of Technology, Policy and Industry Development

Center of Innovation and Product Development

Engineering Systems Division

Sloan School of Management

Department of Mechanical Engineering

Georgia Institute of Technology (Georgia Tech)

Purdue University

Northwestern University

University of California at Berkeley

University of Michigan at Ann Arbor

University of Cambridge

London School of Economics and Political Science (LSE)

Imperial College of Science, Technology and Medicine

Loughbrough University

The Hong Kong University of Science and Technology (HKUST)

The University of Hong Kong

The Chinese University of Hong Kong

The Hong Kong Polytechnic University

控制科学, 运筹学, 工业工程, 机械工程领域的重要学者:

Ho, Yu-Chi (何毓琦): Harvard University

Cao, Xi-Ren (曹希仁): HKUST

Gong, Weibo (龚维博): University of Massachusetts

Luh, Perter B. (陆宝森): University of Connecticut

Lee, Chung-Yee (李忠义): HKUST

Lee, Hau L. (李效良): Stanford University

Tseng, Mitchell M. (曾明哲): HKUST

Kuo, Way (郭位): University of Tennessee

Zhou, MengChu (周孟初): New Jersey Institute of Technology

Jin, Jionghua (金炯华): University of Arizona

Dai, J. G. (戴建刚): Georgia Tech

Yao, David D. (姚大卫): Columbia University

Gershwin, Stanley B.: MIT

Hopp, Wallace J.: University of Michigan, Ann Arbor

Spearman, Mark L.: Georgia Tech

Kumar, P. R.: University of Illinois at Urbana Champaign

Buzacott: York University

## 第 2 章

# 制造系统控制概述

根据不同的制造系统模型,可以获得不同的制造系统控制方法.一般地,对于制造系统的控制可分为三种:

1. 基于离散时间系统模型的生产控制(静态的生产计划);
2. 基于连续时间系统模型的生产控制(实时反馈控制律);
3. 基于离散事件系统模型的生产控制(车间生产调度);

本章将对这三种控制作一概述.

### § 2.1 基于离散时间系统模型的生产控制

#### 1. 生产计划问题的一般提法

生产计划问题的一般提法是:在一定的时间范围(horizon)内,确定每个周期(period)内每种产品的生产数量,使得总的生产成本最小(包括在制品库存成本(WIP inventory, 或超产 overproduction), 最终产品库存成本(finished goods inventory, FGI), 最终产品欠产成本(backlog, underproduction), 设备使用成本, 加班成本(overtime), 生产准备成本(setup)等), 同时满足生产系统动态方程, 生产能力约束, 递阶关系约束等约束条件.

当制订计划时考虑的时间周期数, 产品种类数, 工序数很多时, 问题的规模就会很大, 使得求解困难. 如果把所有应该考虑的因素, 例如生产准备时间, 产品的结构等, 都反映到一个模型中, 那么这个模型就会非常复杂而难以分析. 为解决这个问题, 研究者提出建立生产计划问题的递阶模型, 于是就产生了这样一个研究领域 --- “递阶生产计划”.

#### 2. 递阶生产计划(hierarchical production planning)

“递阶”就是分层. 对于生产计划问题而言, 从不同的角度出发, 有不同的递阶方法. 到目前为止, 递阶分解的方法主要可分为如下几类:

- (1) 按产品族分解.
- (2) 按产品结构分解.
- (3) 按工艺路线和生产组织结构分解.
- (4) 按时间分解.

这里, 简单介绍按产品结构分解的递解生产计划问题.

##### 按产品结构分解

按产品结构分解的递阶生产计划方法是最为常见, 例如下面将要介绍的“物料需



求计划(MRP)”的制订. 如果要对生产成本进行优化, 则可建立多阶段生产计划模型. 这里仅考虑两阶段的特殊情形, 即先在机加工车间生产零件, 然后在装配车间生产产品.

首先建立装配车间的产品生产计划问题的模型:

$$(P1) \quad \text{Min} \quad \sum_{i=1}^I \sum_{t=1}^T [c_i^+ x_i^+(t) + c_i^- x_i^-(t) + p_i u_i(t)] + \sum_{t=1}^T [\lambda R(t) + \theta O(t)] \quad (1)$$

$$\text{s.t.} \quad x_i(t) = x_i(t-1) + u_i(t) - d_i(t), \quad i = 1, \dots, I; t = 1, \dots, T; \quad (2)$$

$$\sum_{i=1}^I m_i u_i(t) \leq R(t) + O(t), \quad t = 1, \dots, T; \quad (3)$$

$$x_i(t) \geq 0, u_i(t) \geq 0, R(t) \geq 0, O(t) \geq 0$$

这里,

$T$ : 生产计划的周期数, 即时间范围(horizon)的长度;

$I$ : 产品的种类数;

$t$ : 时间标识;

$i$ : 产品种类标识;

$x_i(t)$ :  $t$ 周期开始时产品 $i$ 的在制品数量;

$x_i^+(t)$ :  $t$ 周期开始时产品 $i$ 的超产数量,  $x_i^+(t) = \max\{x_i(t), 0\}$ ;

$x_i^-(t)$ :  $t$ 周期开始时产品 $i$ 的欠产数量,  $x_i^-(t) = \max\{-x_i(t), 0\}$ ;

$c_i^+$ : 产品 $i$ 的单位超产成本;

$c_i^-$ : 产品 $i$ 的单位欠产成本;

$u_i(t)$ :  $t$ 周期内产品 $i$ 的生产数量;

$p_i$ : 产品 $i$ 的单位生产成本(不含人工);

$d_i(t)$ : 周期 $t$ 内对产品 $i$ 的需求.

$m_i$ : 生产一个单位的产品 $i$ 所需要的时间;

$\lambda$ : 装配车间单位正常工作时间成本;

$\theta$ : 装配车间单位加班工作时间成本;

$R(t)$ :  $t$ 周期可用(available)的正常工作时间;

$O(t)$ :  $t$ 周期可用的加班时间;

然后建立机加工车间的零件生产计划问题的模型.

$$(P2) \quad \text{Min} \quad \sum_{j=1}^J \sum_{t=1}^T [\hat{c}_j^+(t) y_j^+(t) + \hat{c}_j^-(t) y_j^-(t) + \hat{p}_j(t) v_j(t)] + \sum_{t=1}^T [\hat{\lambda} \hat{R}(t) + \hat{\theta} \hat{O}(t)] \quad (4)$$

$$\text{s.t.} \quad y_j(t) = y_j(t-1) + v_j(t) - \sum_{i=1}^I n_{ij} u_i(t + L_{ij}), \quad j = 1, \dots, J; t = 1, \dots, T; \quad (5)$$

$$\sum_{j=1}^J \hat{\tau}_j v_j(t) \leq \hat{R}(t) + \hat{O}(t), \quad t = 1, \dots, T; \quad (6)$$

$$y_i(t) \geq 0, v_i(t) \geq 0, \hat{R}(t) \geq 0, \hat{O}(t) \geq 0$$

这里,

$J$ : 零件类型数;

$n_{ij}$ : 产品  $i$  对零件  $j$  的平均需求数量;

$L_{ij}$ : 产品  $i$  对零件  $j$  的加工提前期的要求;

$y_j(t)$ : 零件  $j$  在周期  $t$  结束时的库存;

$y_j^+(t)$ :  $t$  周期开始时零件  $j$  的超产数量,  $y_j^+(t) = \max\{y_j(t), 0\}$ ;

$y_j^-(t)$ :  $t$  周期开始时零件  $j$  的欠产数量,  $y_j^-(t) = \max\{-y_j(t), 0\}$ ;

$\hat{c}_j^+$ : 零件  $j$  的单位超产成本;

$\hat{c}_j^-$ : 零件  $j$  的单位欠产成本;

$v_j(t)$ : 零件  $j$  在周期  $t$  的生产量;

$\hat{p}_j(t)$ : 零件  $j$  的单位生产成本(不含人工);

$\hat{\tau}_j$ : 生产每个单位零件  $j$  的时间;

$\hat{\lambda}$ : 零件生产车间中正常工作时间单位人工成本;

$\hat{\theta}$ : 零件生产车间中加班时间单位人工成本;

$\hat{R}(t)$ : 零件生产车间中周期  $t$  内正常工作时间;

$\hat{O}(t)$ : 零件生产车间中周期  $t$  内的加班时间;

注意: 上述两个模型的关联体现在公式(5)中的  $\sum_{i=1}^I n_{ij} u_i(t + L_{ij})$ .

### 3. 物料需求计划(material requirements planning, MRP)

上文提到, 物料需求计划(MRP)实际上是一种根据产品结构对生产计划进行分解的方法. MRP 要完成的工作包括: 每种零部件的生产数量(对于自制件)、采购数量(对于外购件), 以及生产或定货时间. 本节简单介绍 MRP 方法.

MRP: Material Requirements Planning 物料需求计划

MRPII: Manufacturing Resource Planning 制造资源计划

ERP: Enterprise Resource Planning 企业资源计划

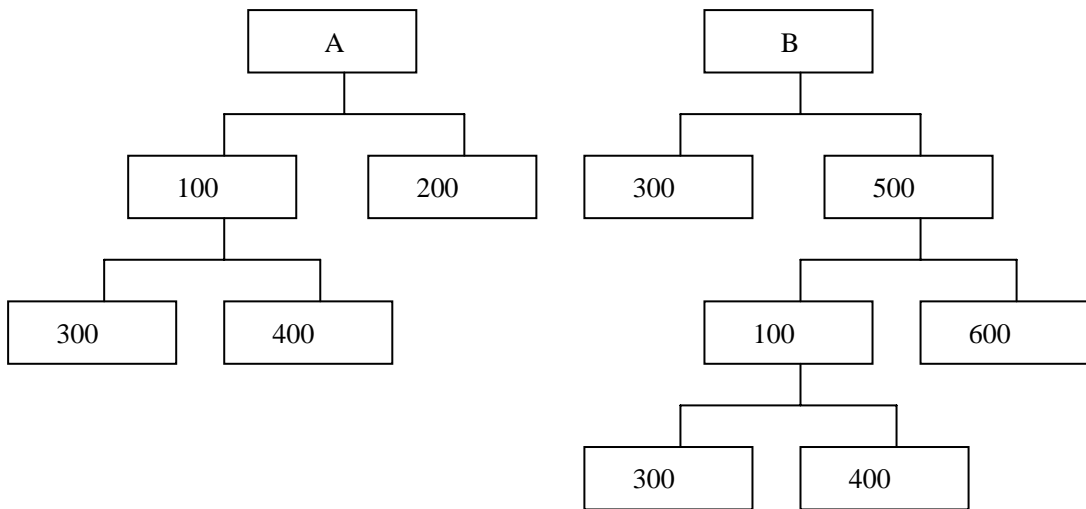
在 MRP 出现之前, 生产过程的控制是在库存控制的基础上实现的, 但比较适合于最终产品的生产过程控制, 但对于组成这些产品的零部件的生产过程的控制则不容易做到很精确. 而 MRP 则解决了这个问题.

一些基本概念:

独立需求(Independent Demand): 来自于制造系统外部的需求, 包括最终的产品和一些对零部件的单独需求(比如备件).

非独立需求(Dependent Demand): 构成独立需求产品的零部件的需求.

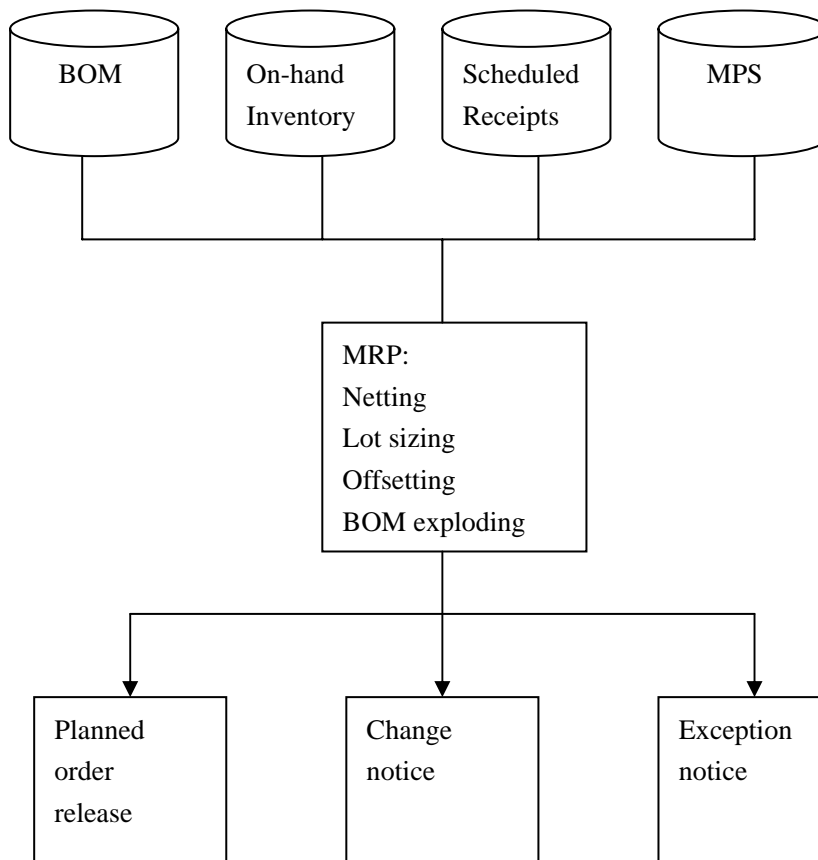
物料清单(Bill of Material, BOM): 独立需求与非独立需求之间的树状结构关系.



MRP 的输入: 物料清单(BOM), 主生产计划(MPS), 当前库存, 已安排的计划

MRP 的输出: 生产计划, 变更通知, 例外事件的处理

MRP 的流程: 由毛需求计算净需求, 确定生产批量, 确定投入生产时间, BOM 的展开



MRP 的输入:

1. 主生产计划(MPS, master production planning)

主生产计划指所有独立需求的产品, 部件或零件的数量和交付期(due date). 这里, 独立需求包括对物料清单中所有根节点(end item)所代表的产品的需求, 也包括对那些位于物料清单的较低层次的零件的来自系统外部的需求. 主生产计划给出的是毛需求(gross requirements), MRP 的一个重要功能是通过毛需求获得净需求(net requirements).

2. 产品信息(item master file)

产品信息包括物料清单(BOM), 批量划分规则(LSR, lot-sizing rule)和提前期规划(PLT, planning lead time). 其中, 批量划分规则解决的是如何确定零件的生产批量以平衡库存与生产准备(setup)时间之间的矛盾, 因为批量小固然可以降低库存, 但同时批次增多, 相应地生产准备次数与时间会增加; 另一方面, 批量大可以减少生产准备次数, 但同时会导致库存增加. 这里面有一个 tradeoff (权衡). 提前期规划是用来确定对某个零件的生产何时开始, 因为零件的生产需要消耗一定的时间, 因此在交付之前必须有一个提前期(lead time).

3. 当前库存(on-hand inventory)

当前库存产品或零部件的类型和数量.

4. 已安排的计划(scheduled receipts)

已安排的计划实际上是已经下达的计划中的生产任务或订单(planned order release). 因此它是 MRP 上一次运行的结果, 或者正在生产中(对生产任务而言), 或者正在采购中(对采购订单中的订货而言). 一旦生产完成或外购货物到达, 则已安排的计划就变成了当前库存的一部分了.

MRP 的输出:

1. 生产或采购计划(POR, planned order release).

包括要生产或采购的产品或零部件的类型、数量和交付期.

2. 计划变更通知(change notice)

对已下达的生产或采购计划的修改, 例如更改交付期, 或提前(expediting)或推后(deferring).

3. 例外报告(exception report)

向用户报告不能按计划交货.

MRP 的计算过程:  $MRP = MPS * BOM - OHI$

1. 计算净需求(netting)

令  $D_t$ : 时间阶段  $t$  的毛需求(gross requirements);

$S_t$ : 计划在时间阶段  $t$  完成的生产或采购计划(scheduled receipts).

$I_t$ : 在时间阶段  $t$  结束时的当前库存(on-hand inventory).

$N_t$ : 时间阶段  $t$  的净需求(net requirements).

于是  $I_t = I_{t-1} + S_t - D_t$ . 若  $I_t > 0$ , 则净需求为 0; 否则, 净需求为其绝对值. 令  $t^*$  表示  $I_t$  由正变负的时刻, 则净需求为

$$N_t = \begin{cases} 0, & \text{if } t < t^* \\ -I_t, & \text{if } t = t^* \\ D_t, & \text{if } t > t^* \end{cases}$$

## 2. 确定批量(lot sizing)

最常用的两种确定生产批量的方法:

- (1) 最小批量生产(lot for lot): 每个周期的净需求量是多少就生产多少.
- (2) 固定定货周期(fixed order period): 把  $m$  个周期的净需求量累积在一起生产, 以减少生产准备(setup)的次数.

显然,  $m=1$  时, 以上两种方法是等价的.

## 3. 确定生产任务或采购订单下达时间(time phasing)

根据交付期和提前期确定生产或采购任务下达时间: 生产或采购任务下达时间 (Start time) = 交付期(Due date) - 提前期(Lead time). 一般地, 在 MRP 中, 假定 lead time 是固定的.

## 4. BOM 展开(BOM explosion)

首先计算 LLC=0 的产品或零部件的生产计划, 然后按照 LLC 由小到大的顺序, 根据 BOM 计算各零部件的生产计划.

注: LLC 编码方法(Low-level Code). 为 BOM 中每个节点上的零部件建立一个标识其在 BOM 中所处层次的代码, 其编码规则是: 根节点的 LLC=0, 对于处于其他节点位置上的某类零部件  $P_i$ , 若其所有可能的父节点的集合为  $F_i$ , 则  $P_i$  的代码为:  $LLC_i = \max\{LLC_j | j \in F_i\} + 1$ .

例: 零件 A 的 MRP 计算流程

Part A		1	2	3	4	5	6	7	8
Gross requirements		15	20	50	10	30	30	30	30
Scheduled receipts		10	10		100				
Adjusted SRs			20	100					
On-hand inventory	20	5	5	55	45	15	-15	---	---
Net requirements							15	30	30
Planned order receipts							45		30
Planned order releases					45		30		

零件 B, 100, 300, 500 的当前库存, 已安排生产任务的交付期和产量, 批量划分规则和提前期

Part Number	Current On-Hand	Lot-Sizing Rule	Lead Time
B	40	2 weeks	2 weeks
100	40	Lot-for-lot	2 weeks
300	40	Lot-for-lot	1 week
500	40	Lot-for-lot	4 weeks

零件 B 的 MRP 计算流程

Part B		1	2	3	4	5	6	7	8
Gross requirements		10	15	10	20	20	15	15	15
Scheduled receipts									
Adjusted SRs									
On-hand inventory	40	30	15	5	-15	---	---	---	---
Net requirements					15	20	15	15	15
Planned order receipts					35		30		15
Planned order releases			35		30		15		

零件 500 的 MRP 计算流程(其毛需求量即为零件 B 的净需求量乘以每份零件 B 中零件 500 的数量)

Part 500		1	2	3	4	5	6	7	8
Gross requirements			35		30		15		
Scheduled receipts									
Adjusted SRs									
On-hand inventory	40	40	5	5	-25	---	---	---	---
Net requirements					25		15		
Planned order receipts					25		15		
Planned order releases		25 *	15						

\* 例外报告: 延迟下达生产指令

零件 100 的 MRP 计算流程(其毛需求量=零件 A 的净需求量乘以每份零件 A 中零件 100 的数量+零件 500 的净需求量乘以每份零件 500 中零件 100 的数量)

Part 100		1	2	3	4	5	6	7	8
Required from A					90		60		
Required from 500		25	15						
Gross requirements		25	15		90		60		
Scheduled receipts									
Adjusted SRs									

On-hand inventory	40	15	0	0	-90	---	---	---	---
Net requirements					90		60		
Planned order receipts					90		60		
Planned order releases			90		60				

零件 300 的 MRP 计算流程(其毛需求量=零件 B 的净需求量乘以每份零件 B 中零件 300 的数量+零件 100 的净需求量乘以每份零件 100 中零件 300 的数量)

Part 300		1	2	3	4	5	6	7	8
Required from B			35		30		15		
Required from 100			90		60				
Gross requirements			125		90		15		
Scheduled receipts			100						
Adjusted SRs			100						
On-hand inventory	50	50	25	25	-65	---	---	---	---
Net requirements					65		15		
Planned order receipts					65		15		
Planned order releases				65		15			

### 制造资源计划(Manufacturing Resource Planning, MRPII):

MRPII 是对 MRP 的拓展, 它超越了物料的范畴, 从制造资源的角度分析企业的效益. 它以 BOM 为基础, 自下而上地计算每件物料的材料成本、人工成本、制造成本, 从而得出产品的最终成本, 实现物料信息与资金信息的集成. 把帐务处理与事务处理相结合, 追溯资金的流动过程. MRPII 分为长期计划、中期计划和短期生产控制这三个层次, 且增加了需求预测, 能力需求计划等功能.

长期计划: 长期需求预测, 资源计划, 长期的总体生产计划.

中期计划: 主生产计划, 粗能力需求计划, 能力需求计划.

短期生产控制: 生产任务下达, 生产调度, 输入输出控制.

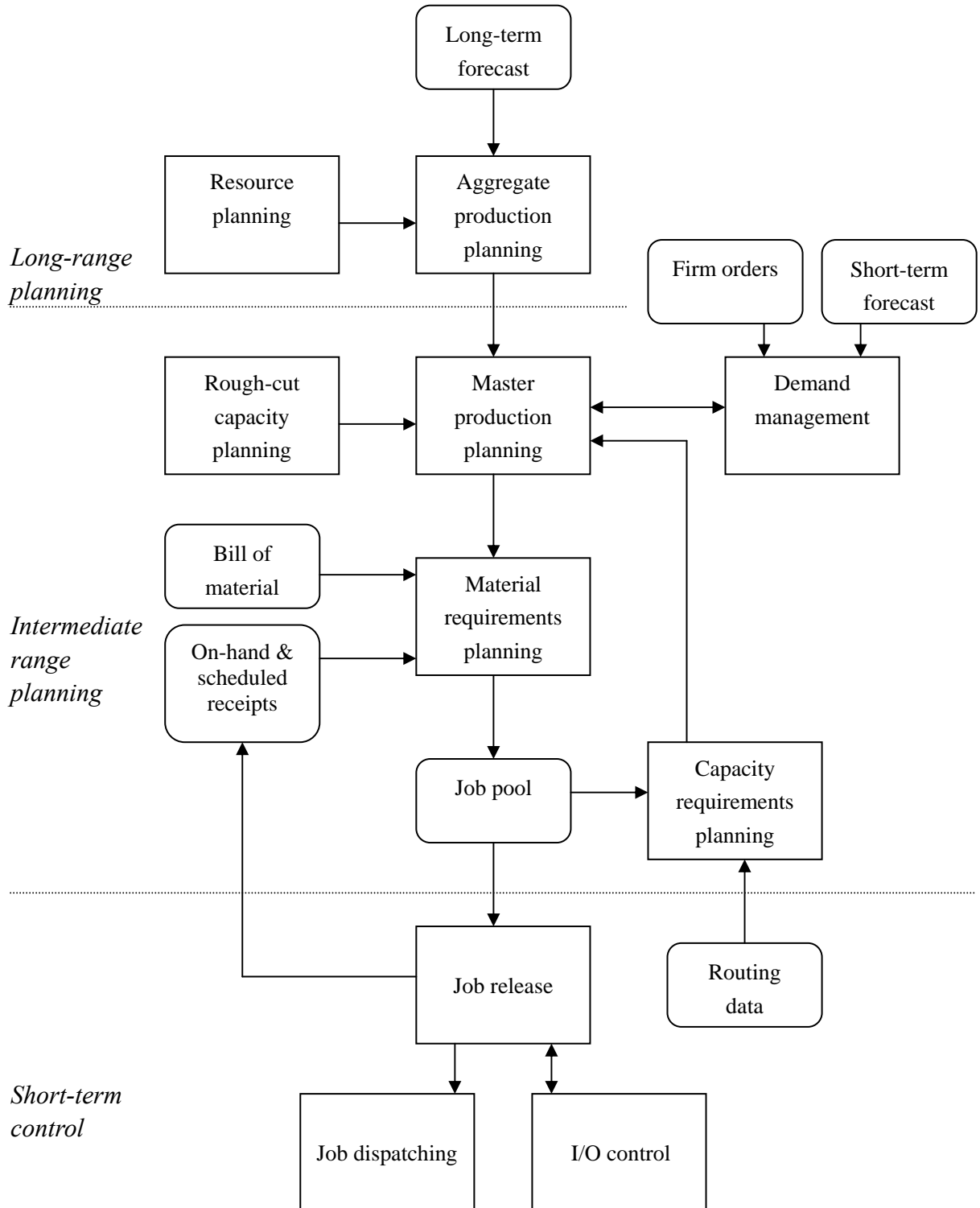
主要模块: 产品数据管理, MPS, MRP, CRP, 库存管理, 车间作业管理, 质量管理, 销售管理, 采购管理, 财务管理, 成本管理等.

### 企业资源计划(Enterprise Resource Planning, ERP):

把 MRPII 拓展到企业的产、供、销、人、财、物、质量管理、库存管理和供应链管理等各个方面. ERP 把企业各方面的信息集成起来, 有助于企业管理者做出正确的决策. 它超越了企业内部的范畴, 把客户需求, 企业内部的经营活动以及供应商的资源融合起来. 增加了供应链管理(supply chain management, SCM)、客户关系管理(CRM, customer relationship management)等模块. 它超越了

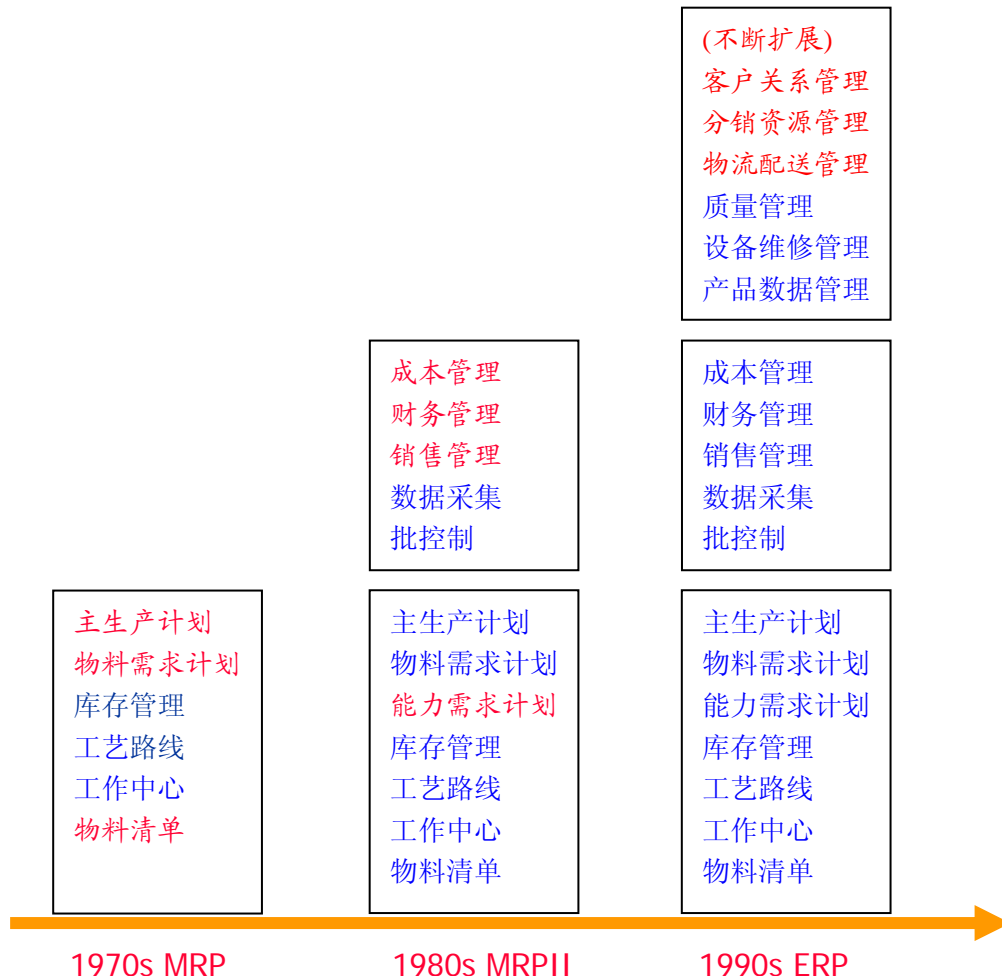
制造业的范畴，应用于金融、IT、零售等行业。其优点包括：各部门信息与功能的集成，统一的用户界面，统一的数据库，统一的系统结构和工具集等。其缺点包括：与现存系统的不完全兼容，实施的长期性和高成本，回报周期长等。

主流的 ERP 系统: SAP/R3 --- 面向事务处理的 ERP 系统，包括财务管理，人力资源管理，制造与物流管理，销售与分销管理等四个方面的应用软件包。



MRPII 的递阶结构.

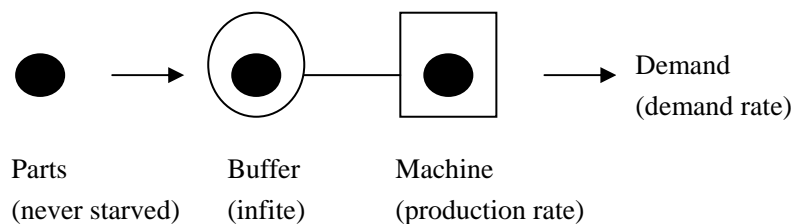




## § 2.2 基于连续时间系统模型的生产控制

如果把制造系统看作一个连续变量动态系统，考虑随机因素(例如机器的故障或需求的不确定性)与生产能力的约束，并以生产成本最小化为性能指标，则可以建立生产系统最优控制问题的数学模型。本节仅简单介绍单机器单品种制造系统的最优控制策略。

单机器单品种(single machine and single part-type)制造系统:



这样一个单机器单品种制造系统, 假设不会发生缺货; 缓冲区容量无限; 机器在正常状态下的最大生产率(production rate)是  $\mu$ , 为一常数; 需求率为  $d$ , 也为一常数; 机器可能发生故障, 以  $\alpha(t)$  表示  $t$  时刻机器的状态,  $\alpha(t)=1$  表示机器是正常的(operational 或 up),  $\alpha(t)=0$  表示机器发生故障(failure 或 down), 单位时间内的故障率为  $p$ , 修复率为  $q$ ; 以  $x(t)$  表示  $t$  时刻系统内的在制品状态, 若  $x(t) \geq 0$ , 则  $x(t)$  表示实际的在制品数量, 若  $x(t) < 0$ , 则  $x(t)$  表示欠产数量; 以  $u(t)$  表示  $t$  时刻机器的实际生产率, 当然  $u(t) \leq \mu$ .

对于这样一个简单的制造系统, 可以容易地写出它作为一个连续时间系统的动态方程:

$$\frac{dx(t)}{dt} = u(t) - d. \quad (1)$$

其中, 控制变量  $u(t)$  是有约束的:

$$u(t) \leq \alpha(t)\mu. \quad (2)$$

现在考虑控制的目标: 最小化生产成本. 这里仅考虑两类成本: 在制品库存成本和欠产惩罚. 以  $c^+$  表示单位时间单位数量的在制品的成本, 以  $c^-$  表示单位时间单位数量欠产惩罚, 另  $x^+(t) = \max\{x(t), 0\}$  表示实际在制品数量, 以  $x^-(t) = \max\{-x(t), 0\}$  表示实际欠产数量. 于是在  $[0, T]$  这段时间内的生产成本的均值为:

$$\frac{1}{T} E \int_0^T [c^+ x^+(t) + c^- x^-(t)] dt \quad (3)$$

于是这样一类单机器单品种制造系统的最优控制问题可以表述为:

$$\begin{aligned} \min \quad & \frac{1}{T} E \int_0^T [c^+ x^+(t) + c^- x^-(t)] dt \\ \text{s.t.} \quad & \frac{dx(t)}{dt} = u(t) - d \\ & u(t) \leq \alpha(t)\mu \\ & u(t) \geq 0 \end{aligned}$$

用动态规划法来解这一最优控制问题, 可以获得这是一种根据不同状态来切换控制变量值的控制策略, 称为 bang-bang 控制:

$$u = \begin{cases} 0, & \text{if } x > Z \text{ and } \alpha = 1 \\ d, & \text{if } x = Z \text{ and } \alpha = 1 \\ \mu, & \text{if } x < Z \text{ and } \alpha = 1 \\ 0, & \text{if } \alpha = 0 \end{cases} \quad (4)$$

这里,  $Z$  为一常数, 称作 Hedging Point, 因此, 这种控制策略也称作 Hedging Point Control Policy.

## § 2.3 基于离散事件系统模型的生产控制

### 1. 车间生产控制(shop floor control)

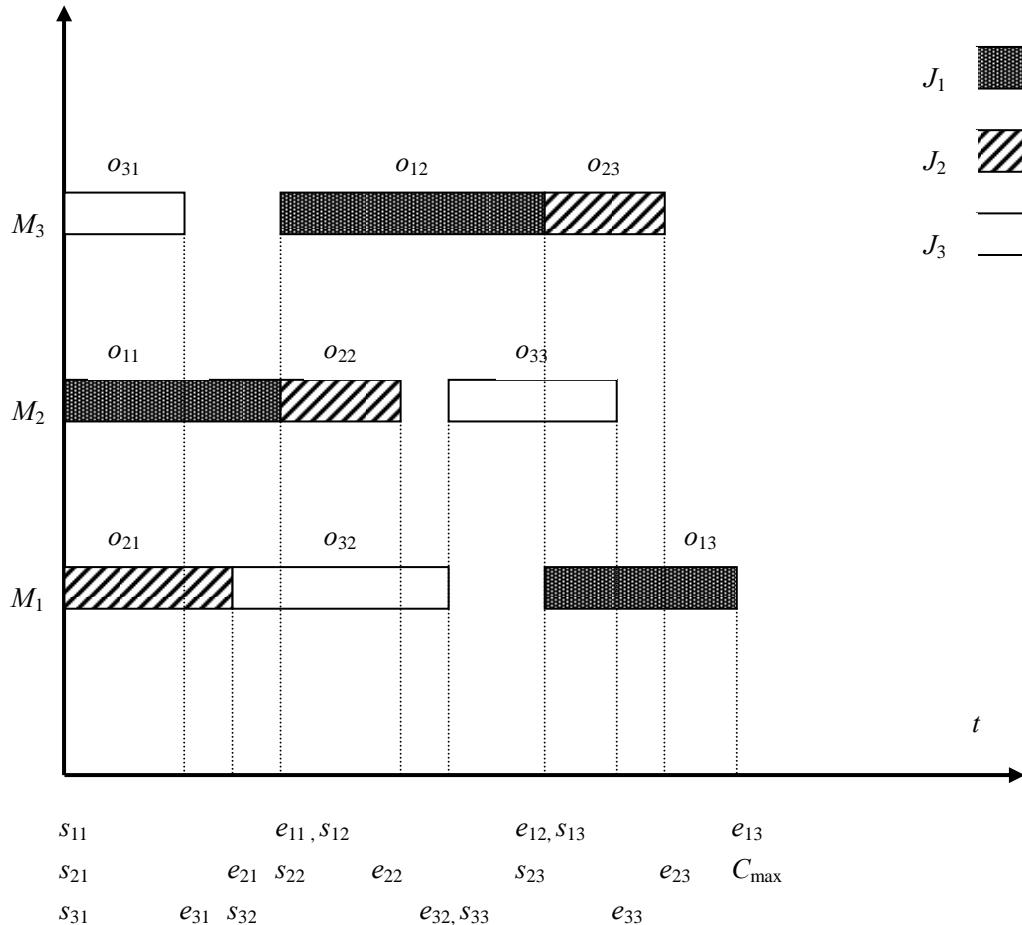
车间生产控制的实现有“推(push)”和“拉(pull)”两种方式. “推”生产方式根据需求来调度生产作业, “拉”生产方式根据系统状态来授权(authorize)生产作业的下达. 推生产方式能够保证满足生产率的要求, 但 **WIP** 可能较高; 拉生产方式能够保证 **WIP** 控制在一定水平上, 但不能保证生产率一定满足要求.

#### “推(push)”生产方式

在 § 2.1 中介绍的 **MRP** 生产方式是一种“推(push)”生产方式. **MRP** 系统生成的每天或每班的物料需求计划, 下达给车间; 车间生产控制系统则通过对生产过程进行调度来满足物料需求计划. 生产调度问题是一个很难的问题, 其一般提法如下:

给定  $n$  个工件(job),  $m$  台机器, 其中  $n$  个工件的加工工艺顺序(即所经过的机器的顺序)是已经确定的, 要确定每台机器上每个工件的加工顺序, 使得某个性能指标(例如所有工件加工完成的时间)最优.

例如: 下图(Gantt 图)所示的一个 3 零件 3 机器的调度方案.



由这个简单的例子可以看出:

1. 生产能力没有被全部利用, 机器有空闲, 这就可能造成生产计划可能完不成;
2. 正因为机器有空闲时间, 所以调度存在优化的余地;
3. 每个零件的加工必须满足工艺顺序约束;
4. 每台机器同时只能加工一个零件.
5. 每道工序的加工一旦开始就不能中断, 直到加工结束为止.

上述 3-5 为一个可行调度的基本约束条件.

调度问题的常用性能指标(performance measurements):

1. 完成所有工件加工的最大时间(时间跨度, makespan):

$$C_{\max} = \max\{C_i, i = 1, \dots, n\}, \text{ 这里 } C_i \text{ 为工件 } i \text{ 的完成时间,}$$

2. 完成所有加工作业的时间:  $\bar{C} = \sum_{i=1}^n C_i / n$ ;

3. 所有工件的最大生产周期(sojourn time):

$$F_{\max} = \max\{F_i, i = 1, \dots, n\}, \text{ 这里 } F_i \text{ 为工件 } i \text{ 的停留时间, } F_i = C_i - b_i;$$

4. 所有工件的平均生产周期:  $\bar{F} = \sum_{i=1}^n F_i / n$ ;

5. 工件加工的最大推迟完成时间(lateness):  $L_{\max} = \max\{L_i, i = 1, \dots, n\}$ , 这里  $L_i$  为工件  $i$  的推迟时间,  $L_i = C_i - d_i$ , 其中  $d_i$  为工件  $i$  的交货期(due date);

6. 工件加工的平均推迟完成时间:  $\bar{L} = \sum_{i=1}^n L_i / n$ ;

7. 工件加工的最大拖后完成时间(tardiness):  $T_{\max} = \max\{T_i, i = 1, \dots, n\}$ , 这里  $T_i$  为工件  $i$  的拖后时间,  $T_i = \max\{L_i, 0\}$ ;

8. 工件加工的平均拖后完成时间:  $\bar{T} = \sum_{i=1}^n T_i / n$ ;

单机调度问题:

**Minimize  $\bar{F}$** :把工件按加工时间从小到大排序, SPT(shortest processing time)调度.

**Minimize  $L_{\max}$** :把工件按交货期从小到大排序, EDD(earliest due date)调度.

注意: 目前很多关于单机调度的研究考虑的情况更为复杂一些, 例如考虑批量生产的情形, 考虑准备时间和排序有关的情形.

双机调度问题:

**Minimize  $C_{\max}$** : Johnson 算法. 把所有工件分为两个集合 A 与 B. 凡是在机器 1 上的加工时间小于或等于在机器 2 上的加工时间的工件属于集合 A, 而其余的工件属于集合 B. 先按照最短加工时间优先的规则调度集合 A 中的工件, 然后按照最长加工时间优先的原则调度集合 B 中的工件, 则获得可以最小化加工时间跨度的最优调度.

研究表明, 机器数  $m \geq 3$  的  $n$  个工件的调度问题是 NP 困难(NP-hard)的, 即至今尚未找到多项式复杂程度的算法来解决此问题. 在 NP 困难问题中, 最困难的一类

称作 NP 完全问题(NP-complete), 一旦 NP 完全问题被解决, 则所有的 NP 问题都被解决. 对于 NP 难题, 既然难以获得其最优解, 那么就会有求取满意解的方法, 例如: 基于规则的调度, 启发式算法, 智能优化算法等.

常用的调度规则:

1. FCFS: 先到先服务.
2. LCFS: 后到先服务.
3. EDD: 交货期最早优先.
4. SPT: 最短加工时间优先.
5. LPT: 最长加工时间优先.
6. LWKR: 最少剩余加工时间优先.
7. MWKR: 最多剩余加工时间优先.
8. SLK: 最小松弛时间优先.  $SLACK = d_i - t - \sum_{q=j}^{n_i} p_{iq}$ .
9. FOPNR: 剩余工序数  $(n_i - j + 1)$  最少优先.
10. MOPNR: 剩余工序最多优先.

这些规则的有效性要通过仿真来验证和比较. 规则不是一成不变的, 应当针对具体的问题应用现有的规则, 或设计新的规则.

解生产调度问题的一些常用算法:

分支定界算法(Branch and Bound), 拉格朗日松弛法(Lagrange Relaxation)  
禁忌搜索算法(Tabu Searching), 模拟退火算法(Simulated Annealing),  
遗传算法(Genetic Algorithm)

### “拉(pull)”生产方式

“拉”生产方式起源于丰田汽车的“准时生产(JIT, just in time)”, 即只有下游在需要的时候才生产, 是一种旨在尽可能减少浪费的生产管理思想和模式. 其总的目标是: 在生产线上, 物流应该是连续的, 顺畅的. 每个工位只有在它需要时才从其上游工位获得所需的物料, 且数量是精确的. 实际上 JIT 的内涵不仅仅是准时, JIT 的目标是实现“七个零”: 零缺陷(zero defects), 零批量(zero lot size), 零准备时间(zero setups), 零故障(zero breakdowns), 零搬运(zero handling), 零提前期(zero lead time), 零波动(zero surging)

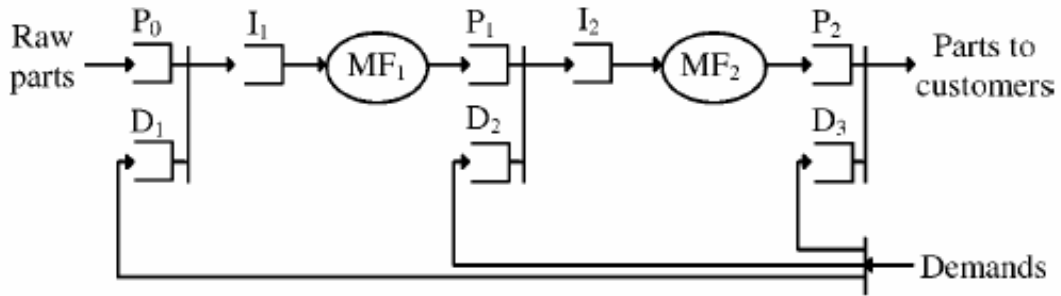
实现 JIT 的方法:

1. 平滑生产(production smoothing): 生产线各工位负荷尽量平衡;
2. 能力缓冲(capacity buffer): MRP 是通过 WIP 来缓冲生产过程中出现的波动, 而 JIT 追求零在制品, 因此采用增加生产能力(时间)的方法来平滑生产过程.
3. 减少生产准备时间(setup reduction): MRP 采用增大批量的方法减少准备时间, 这样会导致 WIP 增加. 而 JIT 正相反, 它要求减少生产准备绝对时间.
4. 培训多面手工人(cross-training).
5. 改善车间布局(plant layout).
6. 全面质量管理(total quality management): 重视解决生产过程中每个环节的质量问题, 一个工人甚至可以停止整个生产线来解决质量问题, 不把质量问题

传播到下游生产环节,也不接受来自上游生产环节的有质量问题的零件。  
1990年,MIT的科学家在JIT的基础上提出了“精益生产(lean production)”。

几种“拉”生产方式的实现方法

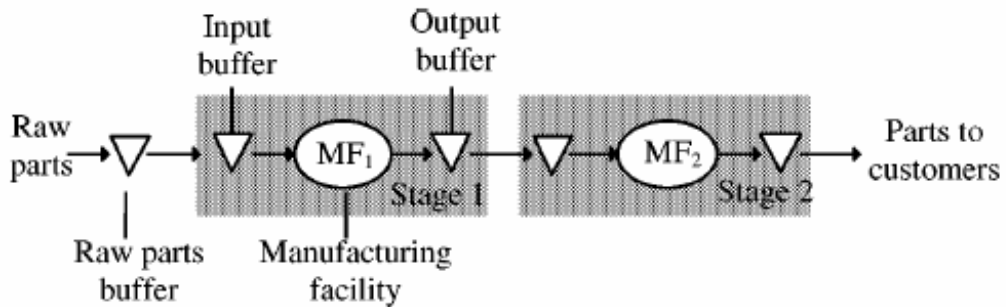
1. 基本库存控制系统(Base Stock Control System, BSCS)



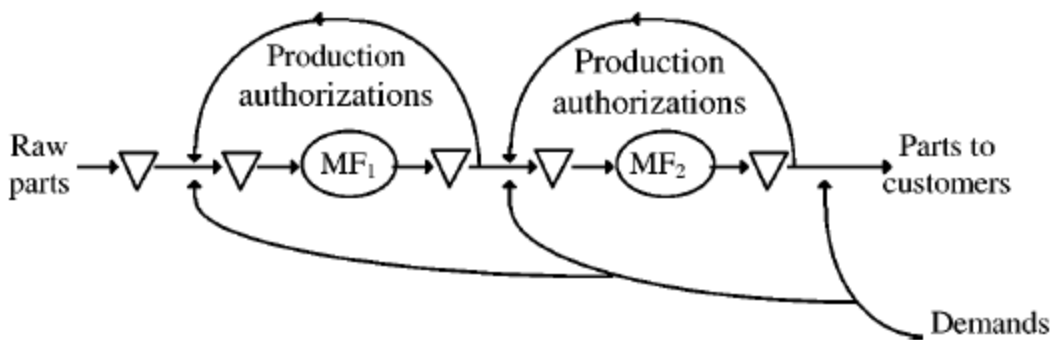
上图中,  $P_i$  中的在制品数量为基本库存量, 每到达一个需求, 则上游机器释放一个零件给下游机器, 下游机器生产一个零件, 补充库存, 使库存量始终保持在基本库存量的水平上。

2. 看板控制系统(kanban control system, KCS)

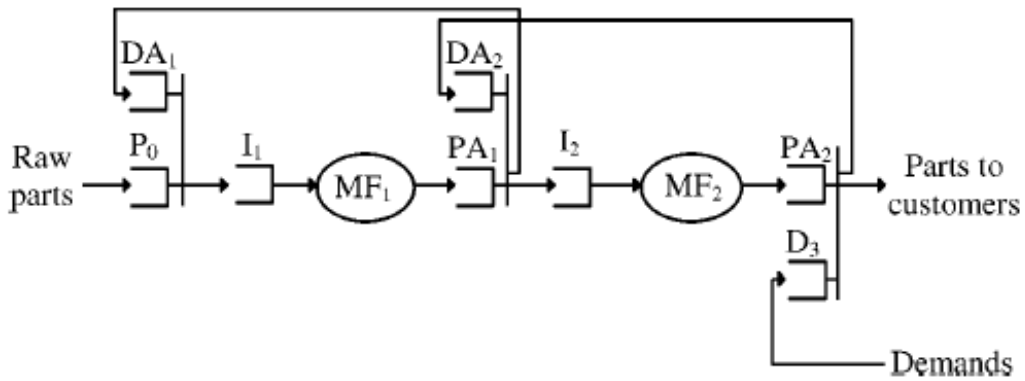
“看板”又称“生产授权卡(production authorization card, PAC)”。设有一如下图所示的串行生产线:



在看板控制系统作用下, 其生产过程如下图所示:



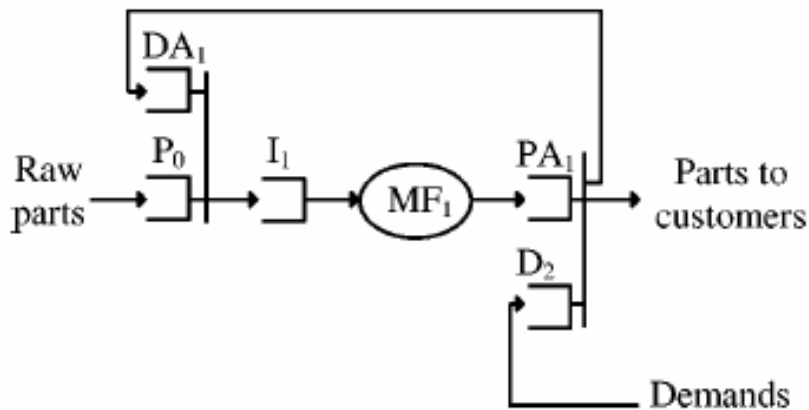
其逻辑关系如下图所示:



上图中， $DA_i$  中保存的是“看板”，看板的数量是确定的，只有在有看板授权 (authorize)的情况下，上游机器才能释放一个零件给下游机器。

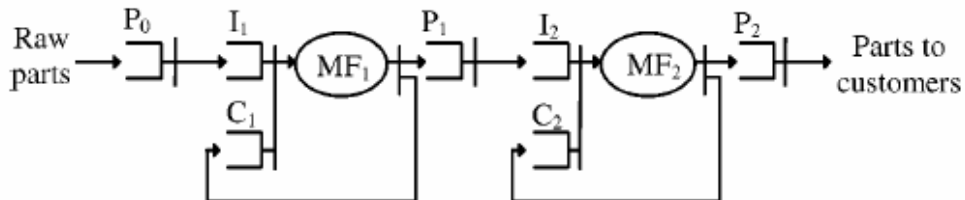
### 3. 定量在制品(constant work-in-process, CONWIP)控制系统:

生产线中的在制品总量是常数，由 kanban 来控制，每个 kanban 附着于一个工件。当一个工件完成生产离开生产线时，释放出其 kanban，该 kanban 返回生产线起始端，授权一个新的工件进入生产线。

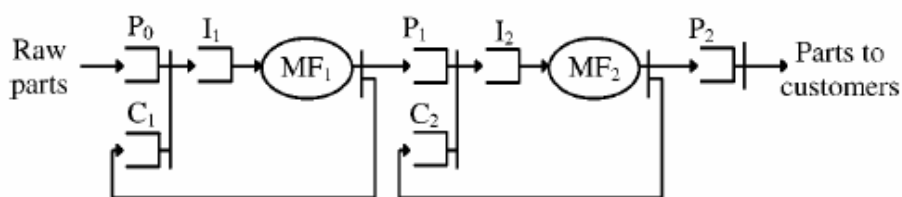


### 4. 分阶段在制品控制系统(stage-WIP-control system)

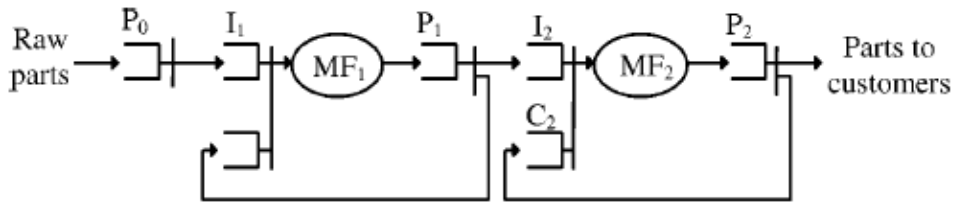
以看板控制机器的生产:



以看板控制机器及其输入缓冲区:

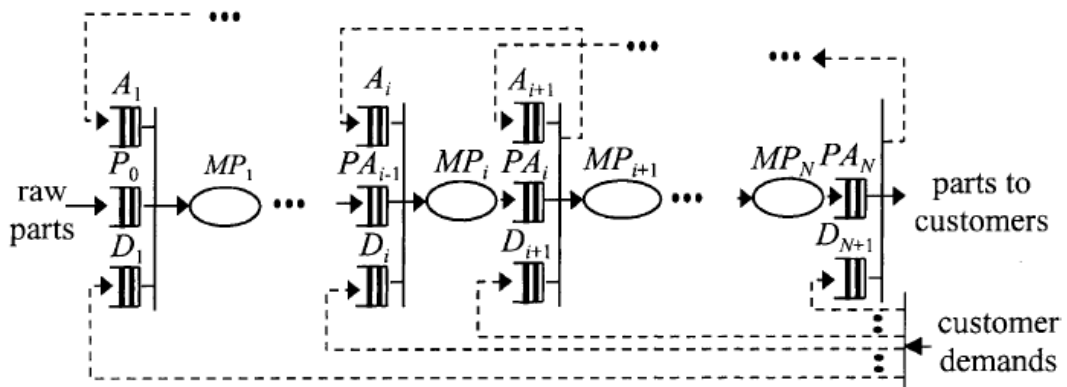


以看板控制机器及其输出缓冲区:



### 5. 扩展看板控制系统(extended kanban control system, EKCS)

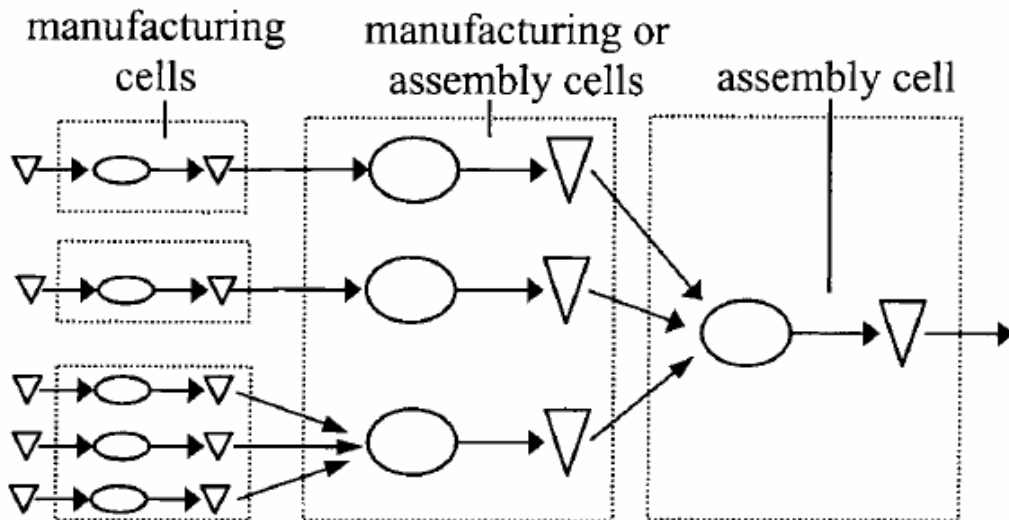
扩展看板控制系统结合了基本库存控制系统和看板控制系统的优点,既能保证每个机器上在制品的基本库存水平,又是由看板来授权上游机器向下游机器释放零件,从而拉动生产进行的。



上图中,  $PA_i$  是机器的输出缓冲区, 其中包含  $S_i$  个零件, 这是基本库存量; 每个零件上附着有一个看板。  $A_i$  中包含的是  $K_i - S_i$  个看板, 因此看板总数是  $K_i$  个。看板数量决定了在制品数量的上界。

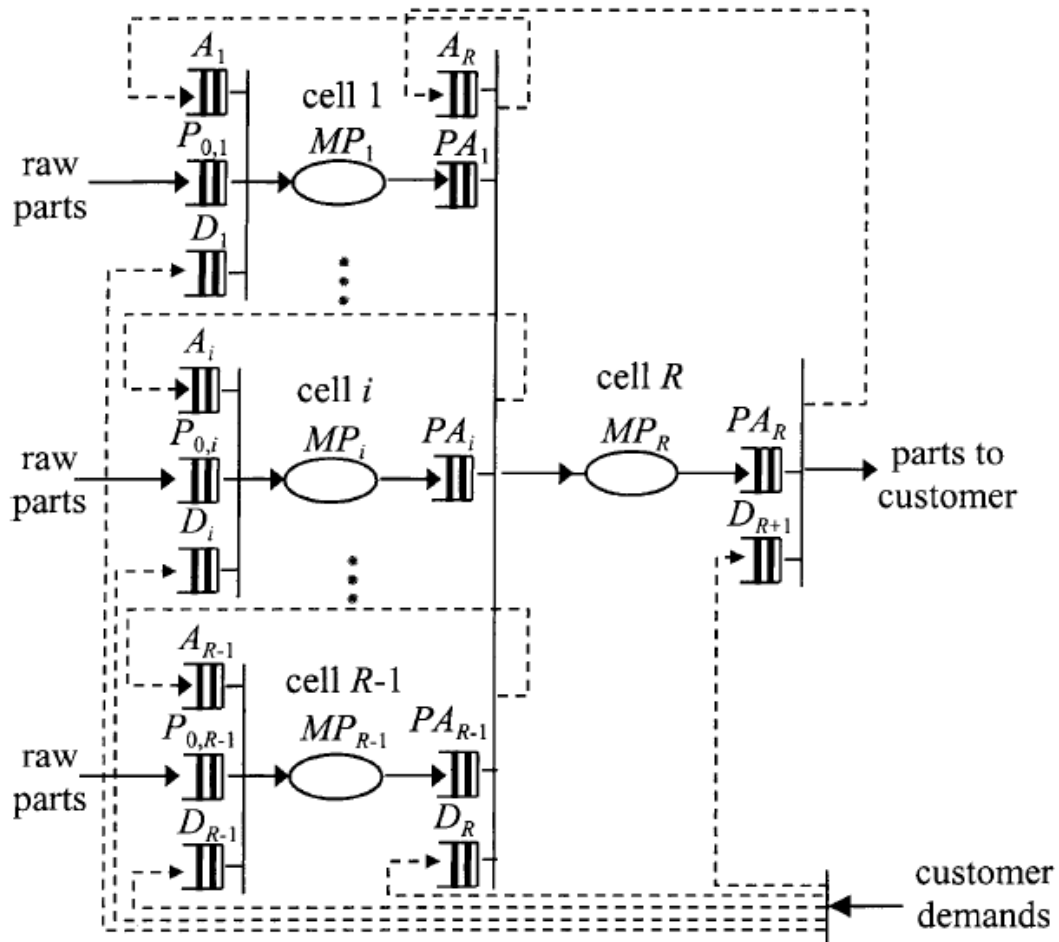
### 6. 并行扩展看板控制系统(simultaneous extended kanban control system, SEKCS)

对于如下图所示的机加工与装配生产系统



可以采用如下形式的扩展看板控制系统





由于装配车间有多个并行的上游生产单位，因此采取并行扩展看板控制系统。

以上各图引自：

1. Liberopoulos, G. and Dallery, Y., 2000. A unified framework for pull control mechanisms in multi-stage manufacturing systems. *Annals of Operations Research* vol.93, pp.325–355.
2. Chaouiya, C., Liberopoulos, G. and Dallery, Y., 2000. The extended kanban control system for production coordination of assembly manufacturing systems. *IIE Transactions*, vol.32, pp.999-1012.

## 2. 基于 Petri 网模型的控制

系统的监控(supervisory control)和避免死锁(dead-lock free)的控制。(见第 5 - 7 章)

## 3. 基于排队网络模型的控制

系统的稳定性分析；生产率控制；工件进入系统(admission control)的控制。(见第 8 - 10 章)

## 第 3 章

### 基于仿真的制造系统分析与控制(1)

制造系统仿真软件在工程上已经得到越来越广泛的应用。常用的仿真软件包括: ProModel, Witness, eM-Plant, QUEST, Simul8, Flexsim 等。MathWorks 公司在其 2006 年版的 Matlab/Simulink 中增加了离散事件系统仿真模块 SimEvents, 这一模块的出现为在 Simulink 环境下对制造系统进行仿真分析提供了良好的工具。本章将首先介绍一个基于 eM-Plant 的装配线仿真分析工程实例, 然后介绍基于 SimEvents 的制造系统分析与控制方法。

#### §3.1 工程实例: 装载机总装生产线仿真分析

某装载机总装生产线包括 1 条总装线, 2 条分装线(前、后车架分装线, 驾驶室分装线)和 7 个分装工序(分配阀分装, 加力泵、储气罐分装, 机油冷却器分装, 油门拉杆分装, 滤清器分装, 发动机罩分装, 轮胎总成分装)。总装线共由 19 个工位组成, 其中有 16 个工位分布在总装车间内, 另外三个工位在完成工作装置装配, 在另一个车间完成。总装线上每个工位要完成若干道工序, 16 个工位共完成工序 68 道。车间共有操作工人 73 人, 其中 38 人在总装线上工作。另外, 总装线上共有大小行车 11 台, 气枪、扳手若干。

装配作业时间测定(work measurement, work study): 以第一和第七两个工位为例。  
一工位:

	上线 2	装下平衡块 3	装后平衡块 4	安装动臂缸 5	安装动臂缸钢管 6	注润滑油脂 7
平均	1:01.6	5:31.3	4:33	5:00	5:41.3	0:55.2
工人分配	白天翔 刘林	白天翔 刘林	白天翔 刘林	张强	徐刚	徐刚
运动中可否装配	F	F	F	F	F	T
工位范围	1	1	1	1-2	1-2	1-2

七工位:

	后转动轴安装 34	滤清器安装 35	装压力油管 36	操纵机构连接 37	连气管 38	电线连接(马达)39	装油门 40
注;		与 31 合(已在 6 工位体现)					
平均	7:36.4	2:02.5	1:48	0:27	1:30	4:23	1:00
工人分配	李品振	贾文迎	薛永芹	张守文	董现宝	刘静	曹幼光
运动中可否装配	F				T	T	T
工位范围	6-7				6-7	6-7	6-8

装配作业先后顺序关系矩阵:

	上线	装下平衡块	装后平衡块	安装动臂缸	安装动臂缸钢管	注油脂	.....	发动机安装	装油管	水箱安装	后转动轴安装	虑清器安装	装压力油管	操纵机构连接	进气管连接	电线连接	油门管连接	油管安装	管路连接	发动机罩安装	
上线																					
装下平衡块	1																				
装后平衡块		1																			
安装动臂缸	1																				
安装动臂缸钢管				1																	
注油脂				1																	
⋮																					
⋮																					
⋮																					
发动机安装																					
装油管								1													
水箱安装								1													
后转动轴安装								1													
虑清器安装									1												
装压力油管								1													
操纵机构连接								1													
进气管连接								1													
电线连接								1													
油门管连接								1													
油管安装								1													
管路连接								1													
发动机罩安装										1	1	1	1	1	1	1	1	1	1	1	1

EM-Plant 仿真软件环境下总装线仿真模型的设计:

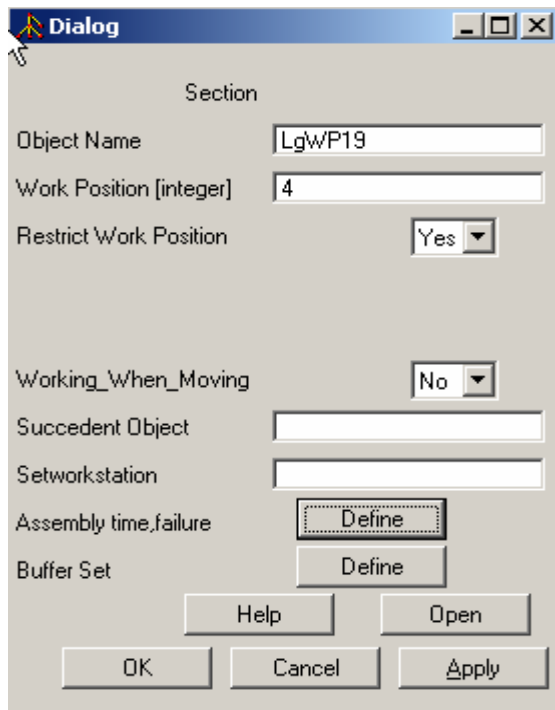
该装载机总装线的特点:

- (1) 总装线为同步线(同走同停), 移动一个工位(约 6.8m)需 2'30", 在每个工位上停留 14'30". 静止与运动交替进行;
- (2) 一个工位中含有多道工序, 每道工序可能由不同的工人来完成; 一个工位上有些工序之间的存在相互约束关系, 使得工位上总的工序的作业存在串行与并行两种情况.

- (3) 工人跨工位作业频繁. 由于工人分配或是由于工人不能按时到位进行操作等原因, 有可能使原来可并行作业的工序实际上以串行方式进行作业.
- (4) 由于行车、空间位置等资源的约束也有可能使原来可并行作业的工序最终以串行方式进行作业.

解决方案:

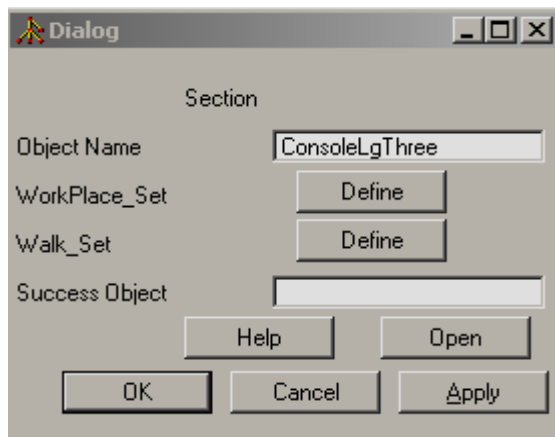
通常生产线的计算机仿真都以工位为对象建模, 然而, 对于这一总装线, 上述特点使得一个工位中工序之间的串行与并行关系很复杂, 并且每个工位的工作持续时间也很难确定. 因此, 难以简单地用工位为对象来建模. 此外, 由于一个工位上存在多个工人操作, 用这种表示方法也很难反映出每个工人的工作量. 因此, 这里采用以工序为对象进行建模的方法.



工序控件:

工序控件是模型中的最基本的单元, 用来模拟实现实际工序装配操作. 工序的主要属性包括:

- Object Name:** 工序名称;
- Work Position:** 分配到哪个工位;
- Restrict Work Position:** 设置该道工序是否一定要在所分配的工位处完成, 默认为是, 如选否, 则出现两个输入框, 用来输入完成该道工序的工位起始处.
- Working\_When\_Moving:** 工件在移动时, 是否可继续操作;
- Succedent Object:** 后续操作内容;
- Setworkstation:** 设置工作站;
- Assembly time, failure:** 该道工序的作业时间及故障时间定义;
- Buffer Set:** 缓冲区容量定义.



装配线走停控件:

用来设定装配线的节拍, 作业时间与走动时间, 并在节拍的作业时间到达时, 判断是否所有工位上的作业内容均已完成, 根据判断结果来推进工件走向下一工位或是停线, 等待直至所有工位上的作业内容均告完成. 其主要属性包括:

- Object Name:** 控件名称;
- Workplace\_Set:** 设置每个生产节拍时间内用于装配作业的时间;

**Walk\_Set:** 设置每个生产节拍时间内装配线走动时间;

**Success Object:** 后续操作内容.

其它控件:

仿真时钟: 用来模拟现实世界中时间的变化;

工件生成序列: 确定工件以什么方式产生;

工件生成器: 根据工件生成序列生成工件;

成品站: 放置装配完的产品, 并对产量进行分析, 显示;

缓冲区: 临时存放待装工件的地方;

装配器: 把部件装配到工件上;

衍生器: 由一个源工件复印生成多个具体同样性质工件的副本;

图表显示器: 以表或图的形式显示分析结果;

工人管理器: 工人的调度管理, 根据请求派出工人, 无需求时招回工人;

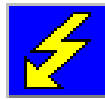
工人: 表示具体的一个工人;

工作站: 提供一个用来装配的空间;

工件: 待装配的原型(如装载机的车架).



仿真时钟



工件生成序列



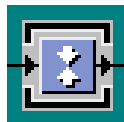
工件生成器



成品站



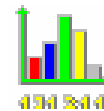
缓冲区



装配器



衍生器



图表显示器



工人管理器



工人



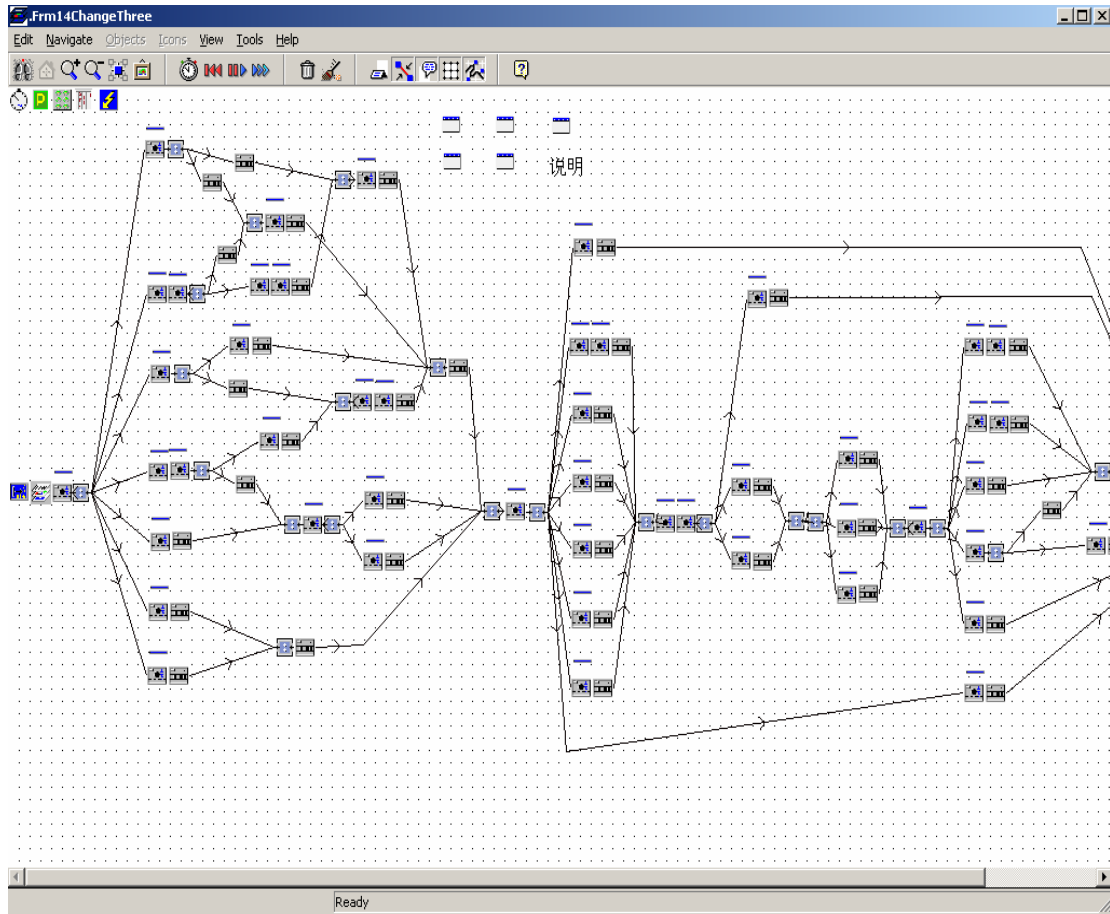
工作站



工件

通过开发与应用以上控件, 以工序为对象, 以工序的先后作业顺序关系及所需的作业资源为约束, 建立装配线的工序网络结构图. 然后, 给每个工序控件的各个属性赋以相应的值(即分配的工位, 作业时间, 分配的工人); 在工序资源约束表中, 给每道工序设定相应的资源(行车、工具、空间位置等)约束; 在工件生成序列控件中, 设置工件的产生模式(单品种或是混合品种); 在装配线走停控件中, 设定装配线节拍(工作时间与走动时间). 完成所有的操作后, 装配线的建模工作就完成了.

最终模型如下图所示.



现行总装线运行状况分析:

在完成系统的建模后,运行仿真模型(生产节拍:17分钟,其中走动 2:30 秒,工作 14:30),得到如下结果(见下表),从所得结果分析可知;

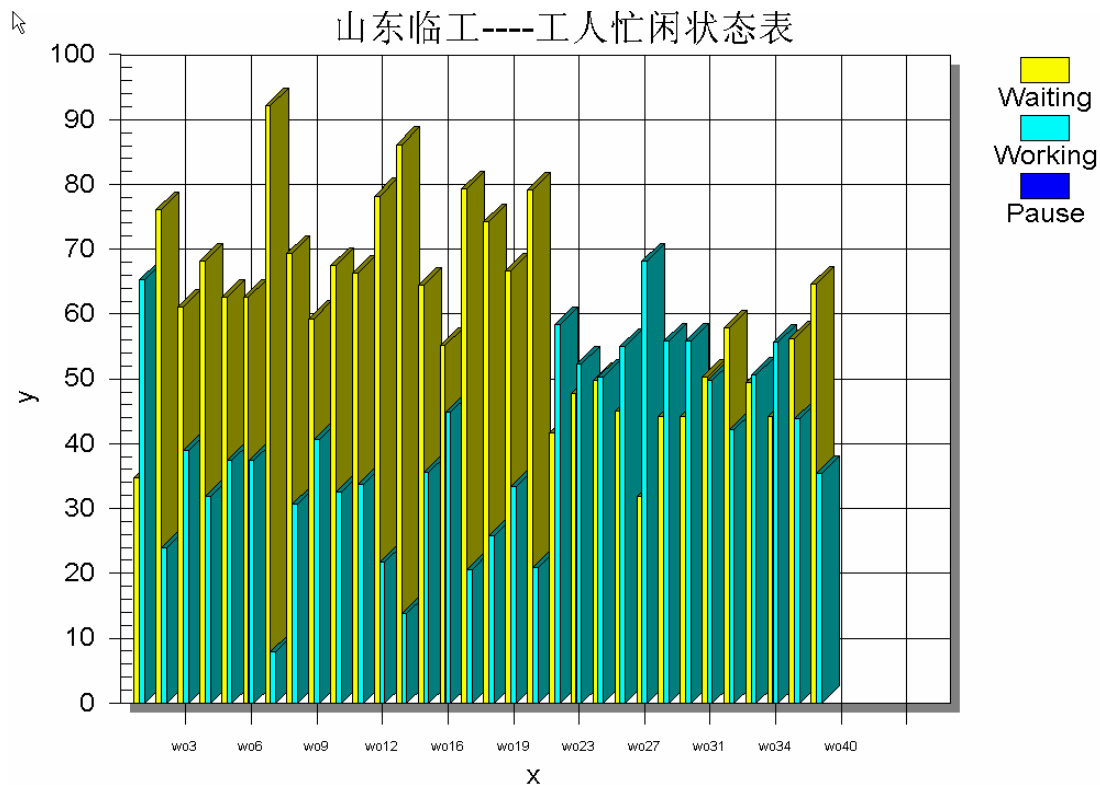
- (1) 工位负荷分布不均;
- (2) 生产节拍为 17 分钟时,总装线上各工位并没有处于满负荷状态下运行.

每工位装配时间(生产节拍:17分钟):

工位	耗时	工位	耗时
1	11:41.1000	9	8:40.0000
2	6:21.0000	10	9:20.0000
3	2:36.8000	11	11:35.0000
4	8:57.0000	12	9:29.5000
5	8:07.3000	13	8:53.0000
6	6:54.6000	14	9:28.0000
7	7:36.4000	15	6:00.0000
8	9:55.0000	16	

注:第 16 工位的工作内容为加油加水,开动装载机下线,并无实质的装配工作,并且时间较短,因此在仿真分析中不予考虑。

工人负荷分析:



- (1) 总装线上工人的负荷极其不均, 最大的达到节拍时间的 70% (14 分钟), 这已接近工节拍所允许的最大作业时间. 而最小的不足 10% (不到 2 分钟).
- (2) 总装线上有一半工人的工作负荷小于 50%, 也就是说, 其休息的时间比工作的时间更长, 劳动强度明显不足.

限制日产量提高(即生产节拍加快)的原因分析:

- (1) 工序分配不合理: 对于在作业上有前后约束关系的工序, 没有尽量分配到不同的工位上.
- (2) 工序划分不合理: 某些单道工序的作业时间都很长, 某些工序还可以进行进一步的拆分.
- (3) 资源的分布不合理: 例如行车的分配.
- (4) 工位的空间布局不合理的情况: 未考虑到空间位置的约束.

多种改进方案仿真实验分析: 节拍压缩到 14 分钟时(走动 2'30", 工作 11'30")  
 现有装配线上, 1 工位与 11 工位的作业时间均超过 11'30", 形成瓶颈. 因此要压缩这两个工位的作业时间. 为此可采取如下措施: 一是把有先后约束关系的工序分配到不同的工位上; 二是拆分工序, 重新分配工人与工位; 三是减少工序在总装线上作业时间.

方案一:

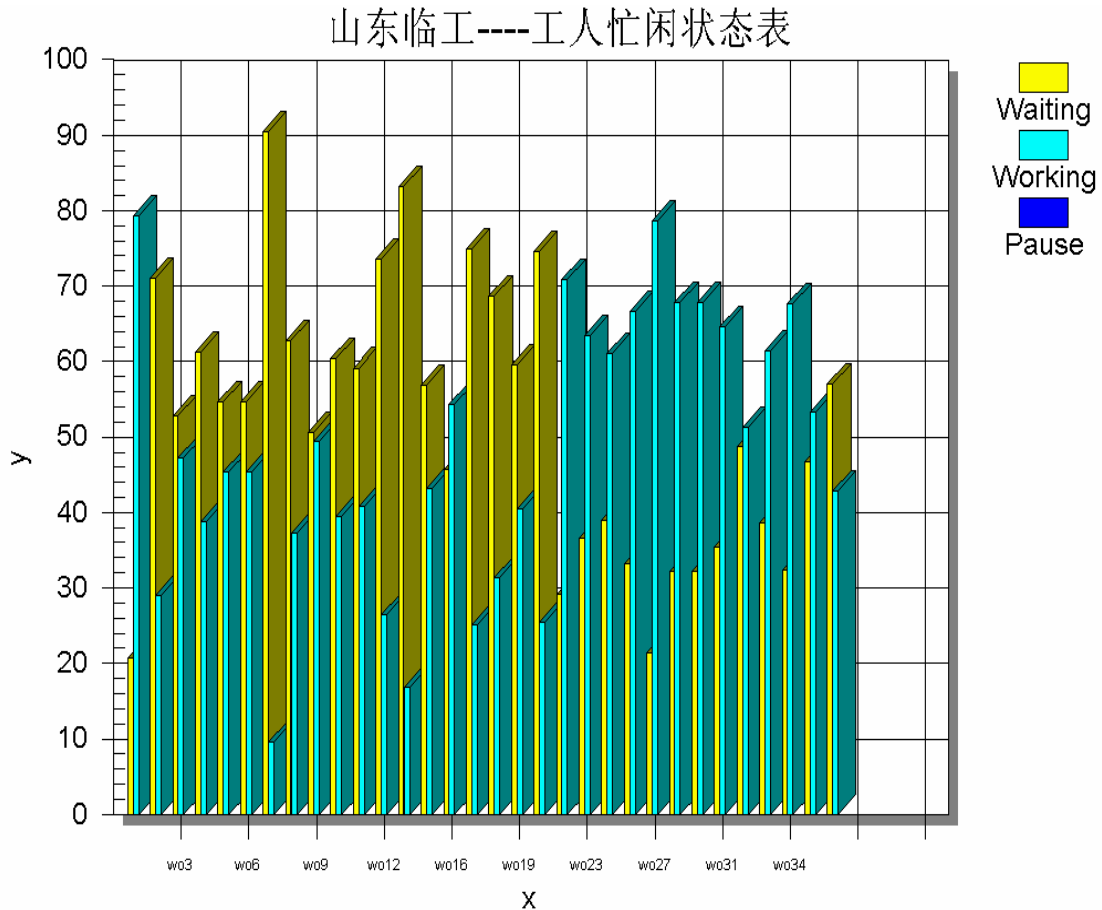
把 1 工位上的安装动臂缸钢管与注润滑油脂调整到 2 工位上;

把 11 工位上的操纵软轴固定调整到 12 工位, 并将此任务分配给原仪表盘安装的工人.

每工位装配时间:

工位	耗时	说明	工位	耗时	说明
1	11:05.9000	工位作业时间缩短	11	11:00.0000	工位作业时间缩短
2	7:51.0000		12	9:29.5000	

工人负荷:



仿真结果:

仿真运行一个月(工作日为 22 天, 每天实际工作 7 小时)折合时间为 154 小时, 另外去除在程序运行时的布线时间(16 工位, 节拍为 14 分钟的布线时间)3 小时 44 分钟. 故一个月的实际工作时间在仿真运行程序中的反映就是:  $154(H)+3(H)44(M) = 157(H)+44(M)$ . 由于仿真软件内的仿真时钟是以每天 24 小时计, 折合为 6 天 13 小时 44 分钟.

方案 1 的仿真结果

实际耗时 (H)	仿真时钟时间 (24H/D)	总产量 (台)	单台耗时	下线数量 (台/每小时)	台/7H	去除布线时间	限制工位
154	6:13:44:00.0000	660	3:44:00.0000	4.28	30.01	是	是

第一种调整方案中, 调整内容只牵涉到 1、2、11 与 12 工位, 调整工序只有 3 道, 只有一个工人分配到新的任务, 因此实施较易. 但总装线上原来工位不平衡, 工人负荷不均等情况并没有得到多大的改善.



## 方案二

把 1 工位上的安装动臂缸钢管与注润滑油脂调整到 3 工位上；把 11 工位上的暖风机、工具箱安装进行改进，即把工具箱先进行组装，然后，再拿到总装线上来装配。其余不变。

经过调整之后，1 与 11 工位处的工位负荷均降到节拍允许的作业时间(11:30)之内，而由于 1 工位处部分工序调整到 3 工位，使得 3 工位处的关键路径产生了变化，该工位负荷有较长的增加；另外 11 工位处的工位负荷降低较大，使得 8 工位成为第二瓶颈。

第二种调整方案中，调整内容只牵涉到 1、3 与 11 工位，调整工序只有 3 道，总装线上无工人分配到新的任务，因此实施起来比第一种方案更为容易。经过这样调整之后，工位相对比较平衡（因为 3 工位的负荷从原有的不到 3 分钟到现在的 7 分多钟，大致达到平均水平）；工人的负荷不平衡状况虽有一定的改善，但仍然比较严重。

## 方案三

在尽量平衡每个工人的作业时间和尽量减少工人跨工位作业的指导原则下，做如下调整：

把 1 工位上的安装动臂缸钢管与注润滑油脂调整到 3 工位上，操作者不变；把 11 工位上的暖风机、工具箱安装进行拆分，暖风机的安装仍然由原来操作者来完成。把工具箱的安装放在 13 工位，由原装座椅的工人来完成。

将下部机构车上安装移到第 1 工位，由安装动臂缸的工人操作(增加其劳动强度)；

将转向缸进油胶管安装改由安装动臂缸钢管的工人操作(减少跨工位作业)；

把部装线上的前桥油管与前桥油管线下组装，安排到 3 工位处(便于部件运输与管理，也增加 3 工位上的工作强度)，由原来的操作者完成；

将单稳阀车上安装调整到 4 工位处，由原下部机构组装的工人来操作。同时把 5 工位上的机油冷却器安装调整到 4 工位，也由此人来做。另外，把连铰接头/两软轴车上安装也改由此人来做(因为两软轴车上安装的工人在总装线上就做这一道工序，这样就可以解放出此人，留待备用)；

把动臂缸胶管连接、转向缸油管连接改由多路阀总成车上安装的工人来完成(操作连贯、方便)；

将机油冷却器安装和油水分离器安装改由装后桥油管连接及放气嘴的工人来完成；(增加其工作强度，原先装后桥油管连接及放气嘴的工人只做一道工序)；

5 工位上剩下工内容，由储气罐安装与后桥油管连接及放气嘴的工人一块协同完成(提高工人的利用率，工作比较紧凑)；

将滤清器安装，装压力油管，操纵机构连接，装油门均改由装连气管的工人来完成；

将发动机罩安装的螺栓紧固由原来燃油箱安装的工人改由来发动机罩安装的工人完成；

将装烟筒空滤器改由装油管连接(单稳/油箱.多路阀/油箱,转向器/回油管)的工人来完成，并安排在 9 工位完成；

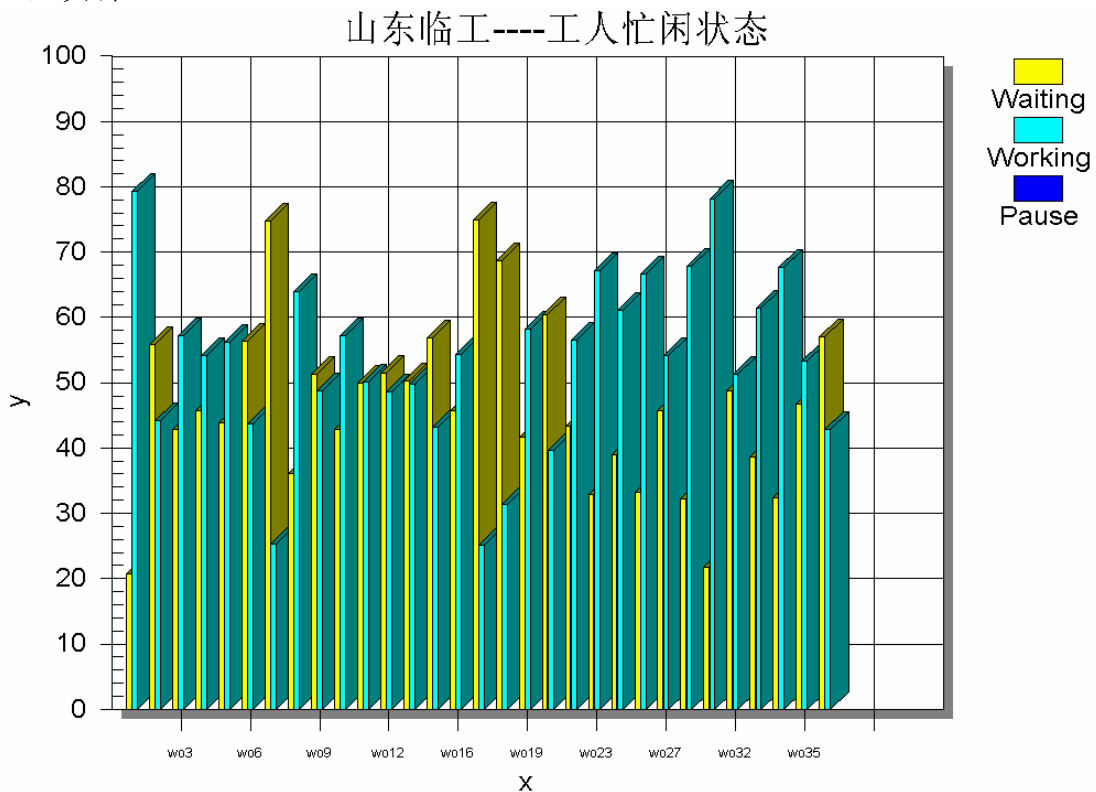
将连手刹车、油门拉杆连接、左右支架安装与仪表盘线下组装启动继电器改由原来操作两软轴车上安装的工人来完成；

其余不变。

每工位装配时间:

工位	耗时	说明	工位	耗时	说明
1	11:05.9000	工位作业时间减少	9	8:40.0000	
2	7:51.0000		10	9:20.0000	
3	8:00.0000	工位作业时间增加	11	7:35.0000	工位作业时间减少
4	8:57.0000		12	9:29.4000	
5	7:38.0000	工位作业时间减少	13	9:24.0000	工位作业时间增加
6	8:18.6000	工位作业时间减少	14	9:28.0000	
7	7:36.4000		15	6:00.0000	
8	9:03.0000	工位作业时间减少	16		

工人负荷:



第三种调整方案中,调整内容涉及大部分工位,调整工序也比较多.通过调整之后,工位与工人的负荷均相对比较平衡(因为由于工位数量是固定的,分装线的分布是固定的,行车等资源的分布也是固定的,故调整将受到一定的限制,节拍压缩到一定程度之后,将难以再提高).然而,由于工人操作的内容有较大的变更,有可能需要对工人进行培训,且需给工人一定的时间使其达到熟练操作,因此实施起来需要一定的投入和时间.

通过对前面三种方案的比较分析,可见第三种方案是比较优的方案,这个方案虽然需要调整的内容要多一些,但是经此调整后,每个工位上的负荷较均衡,大部分工人在总装线上的工作时间也较平衡.

生产节拍压缩到 12 分钟 30 秒 (走动 2:30,工作 10:00):

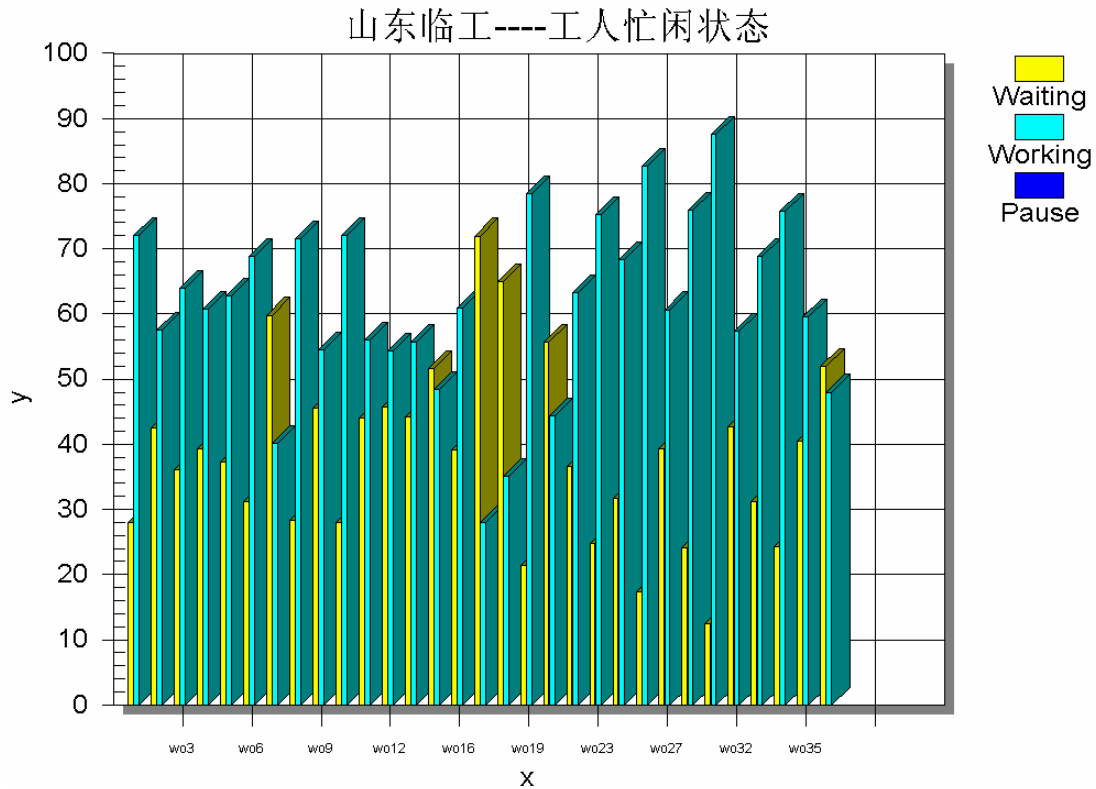
在方案三的基础上,只需在工位 1 上增加 1 个工人,帮助操作上线、装下平

衡块和后平衡块，不需对其它内容进行任何变动，就可把节拍压缩到 12 分 30 秒（走动 2:30,工作 10:00），则日产量可高达到 33 台。

每工位装配时间:

由于在工位 1 上增加了工人，所以装配时间缩短，其余工位上的时间不变。工位 1 上的装配时间降至 9 分钟。

工人负荷:



节拍为 12:30 时工人的忙闲状态 (包括部分工人在线下组装时间)

仿真结果:

实际耗时 (H)	仿真时钟时间 (24H/D)	总产量 (台)	单台耗时	下线数量 (台/每小时)	台/7H	去除布线时间	限制工位
154	6:13:12:00.0000	739	3:20:00.0000	4.79	33.58	是	是

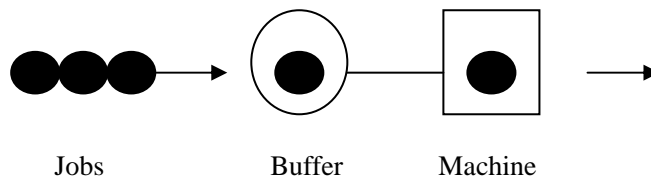
仿真运行一个月(工作日为 22 天, 每天实际工作 7 小时)折合时间为 154 小时, 另外去除在程序运行时的布线时间(16 工位, 节拍为 12 分钟的布线时间) 3 小时 12 分钟。故一个月实际工作时间在仿真运行程序中为 154(H)+3(H)12(M)=157(H)+12(M), 以每天 24 小时计, 折合为 6 天 13 小时 12 分钟。

## §3.2 基于仿真的无故障单机器制造系统分析

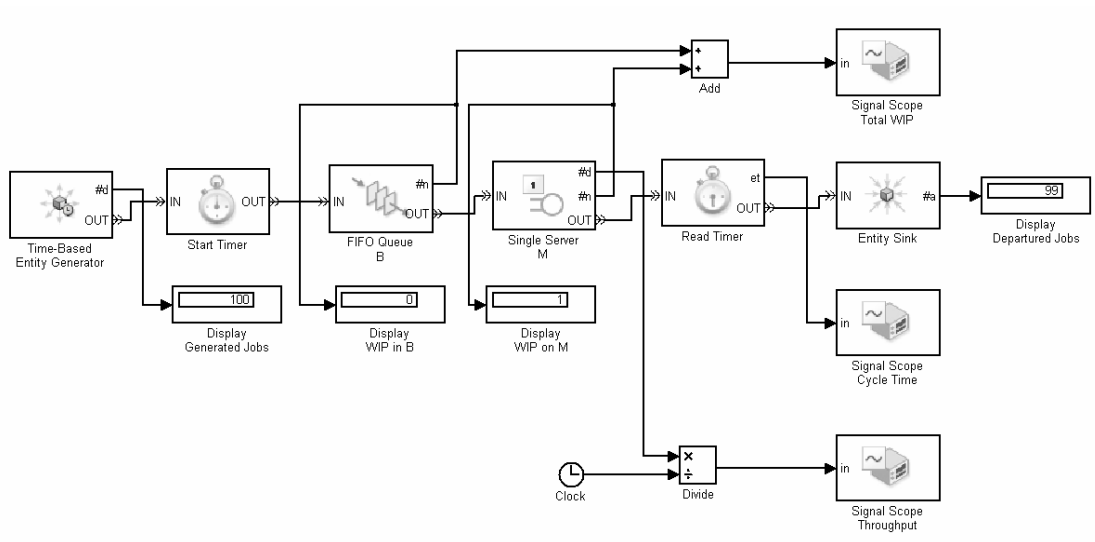
对制造系统进行分析的目的是为了获得其参数与性能之间的关系. 这里, 参数包括: 机器加工一个零件所需时间、平均正常工作时间(MTTF)、平均故障修复时间(MTTR)、零件到达时间间隔、缓冲区容量等. 性能包括: 制造系统的生产率、在制品数量、生产周期、设备利用率等. 控制的目的是通过改变某些容易改变的参数(例如: 机器的生产率)来改善系统的性能. 本节讨论用利用仿真实现上述目的的方法.

### 1. 确定的无故障单机器制造系统

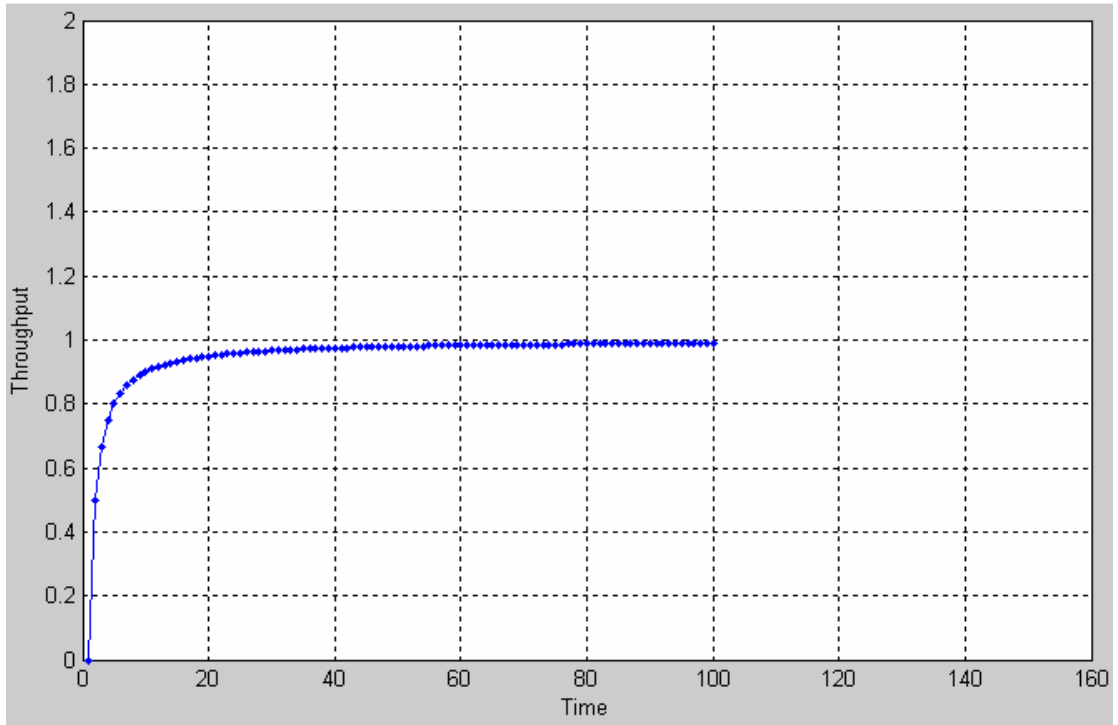
设有一单机器制造系统: 机器生产 1 个零件需要 1 个单位时间; 机器不发生故障. 机器前的缓冲区容量为无穷大; 零件到达的时间间隔为 1 个单位时间.



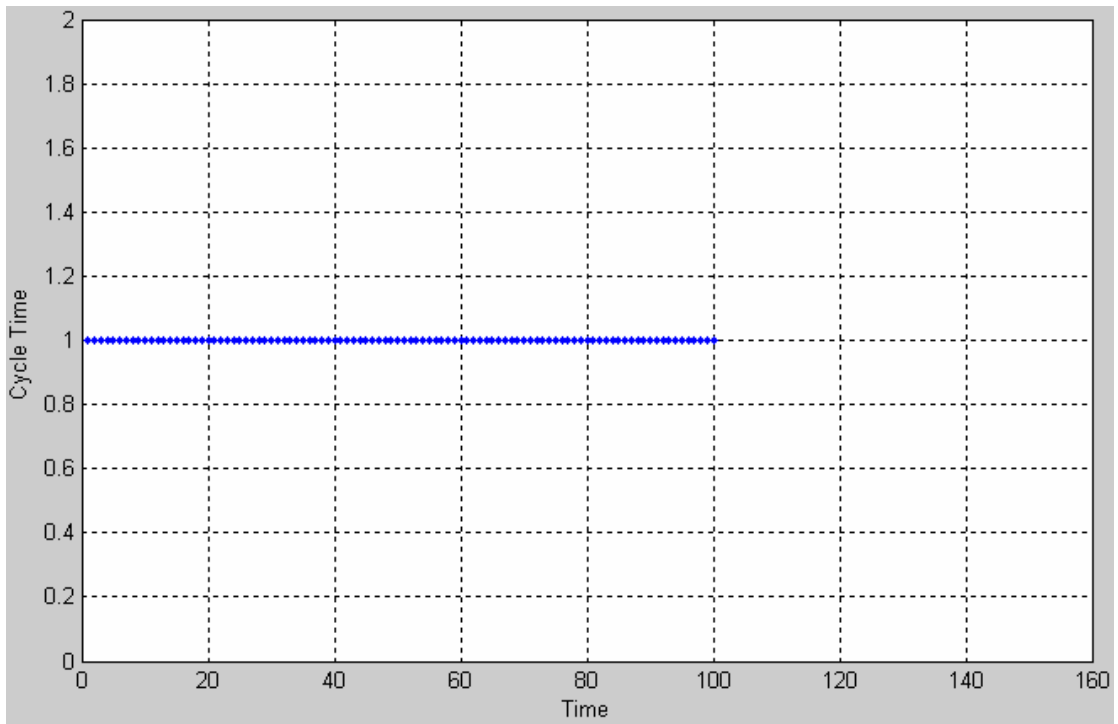
在 Matlab/SimEvents 环境下建立上述制造系统的模型如下:



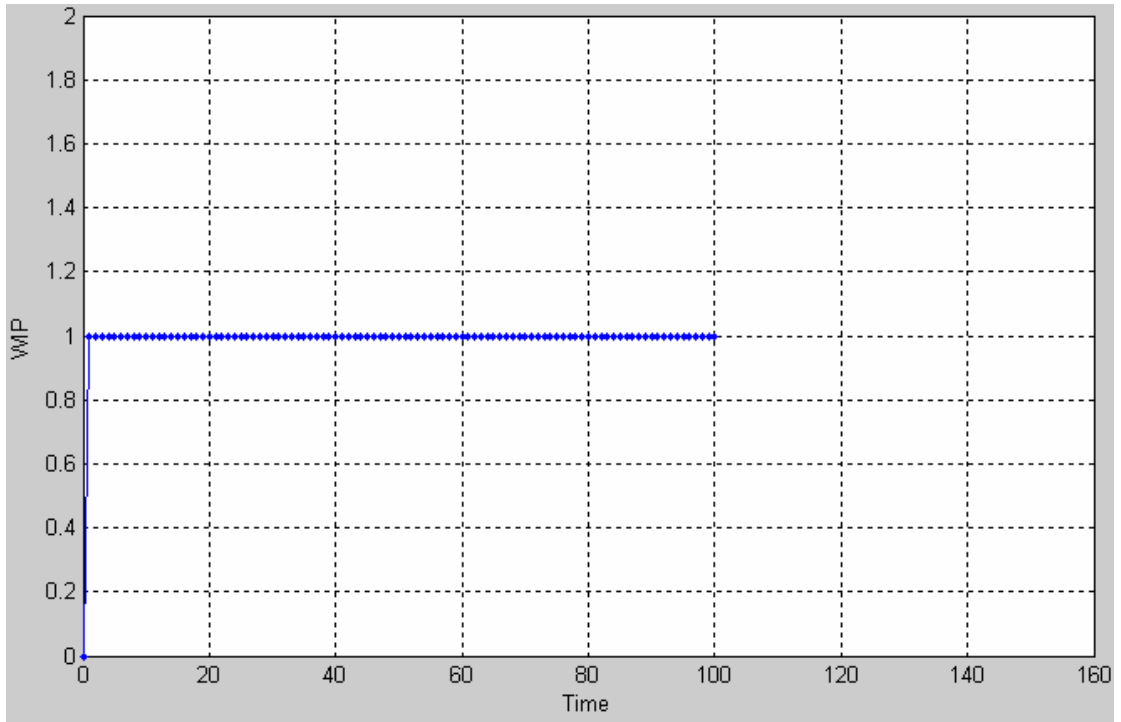
仿真时钟从 0 运行到 100, 获得三个主要参数(系统生产率、生产周期和总的在制品数量)随时间变化的情况. 如下图所示.



系统生产率(稳定时接近 1)

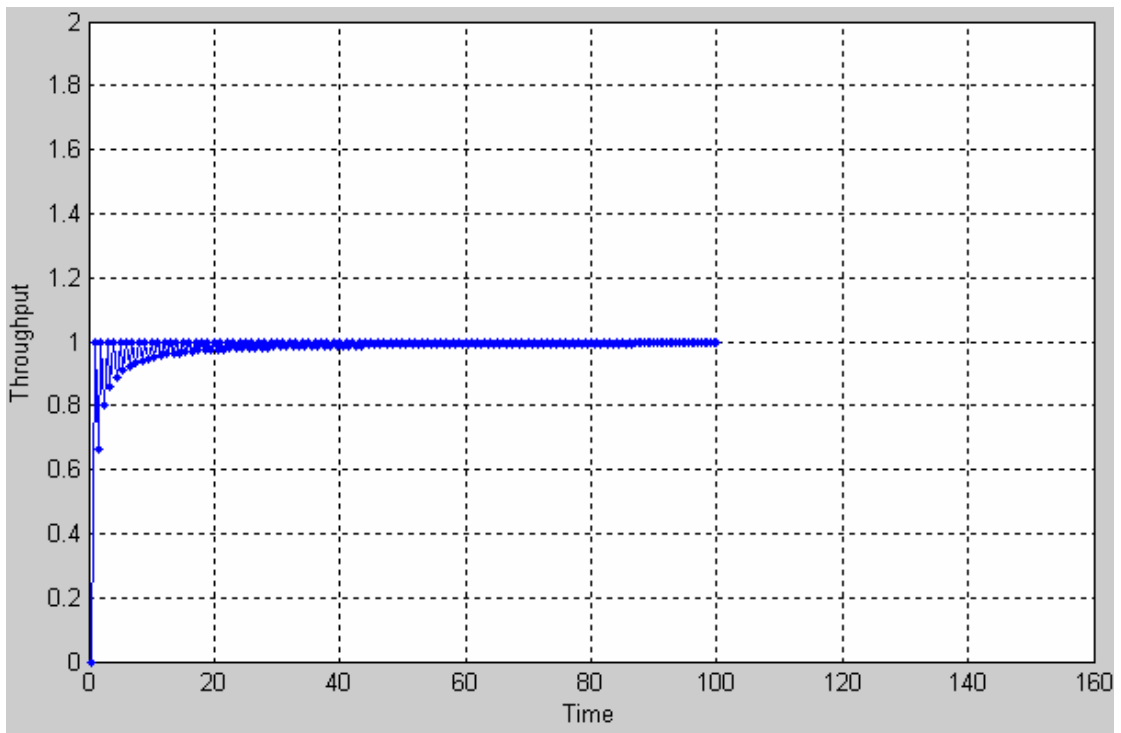


生产周期(=1)

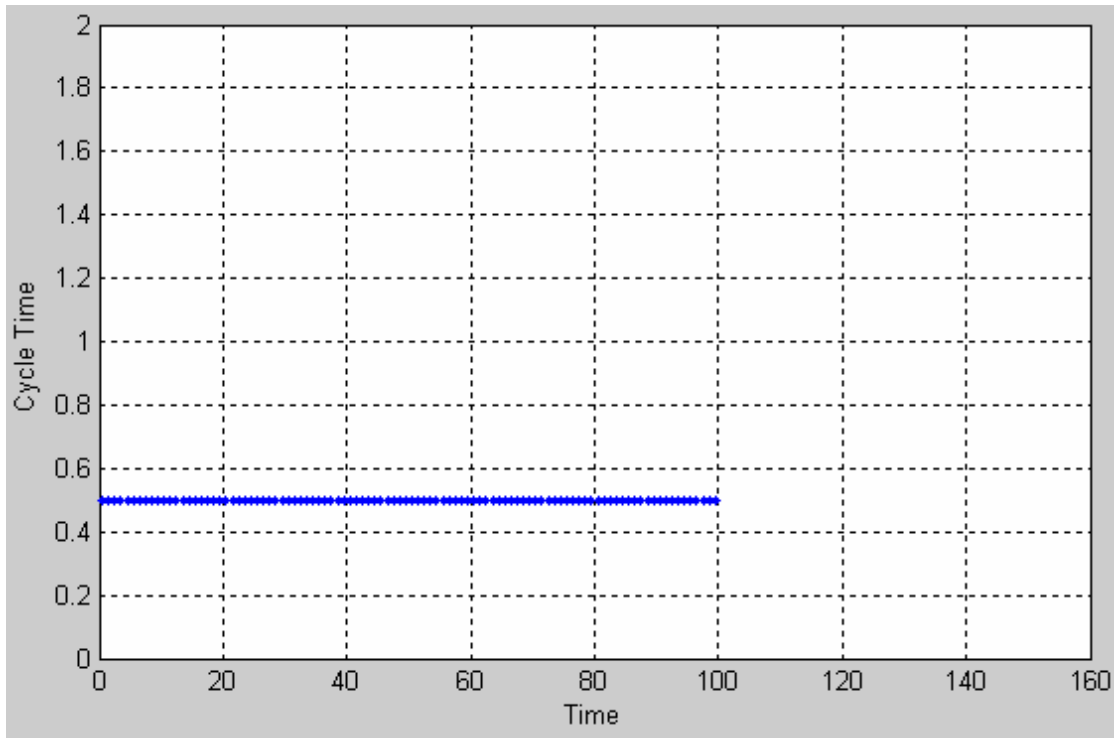


总在制品数量(=1)

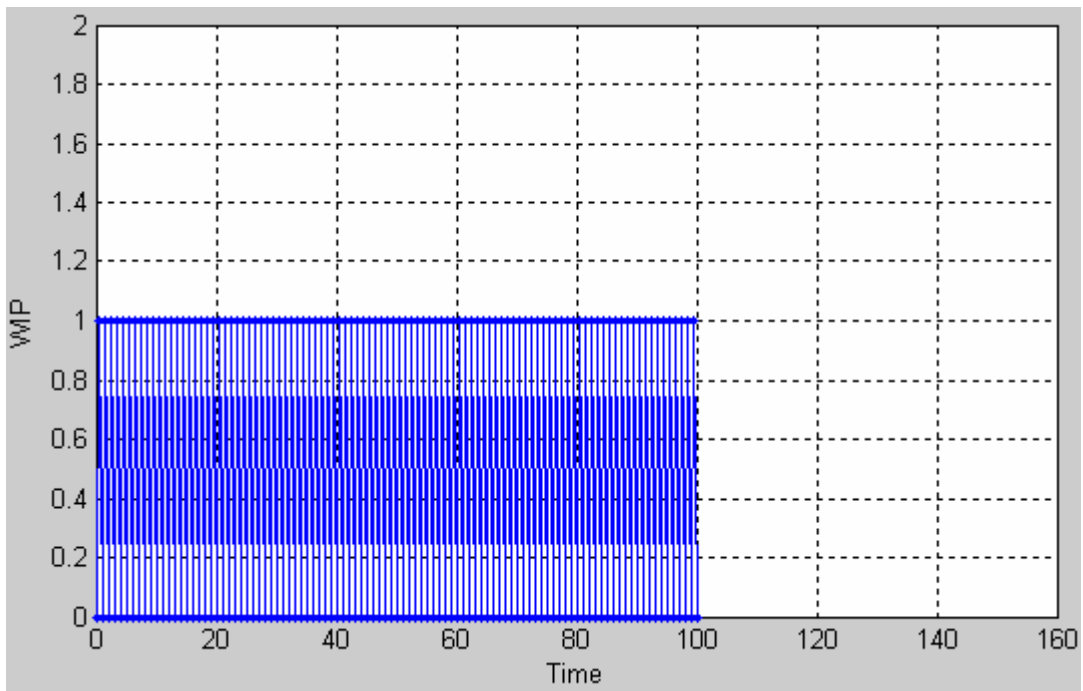
改变系统参数：把机器生产一个零件的时间改成 0.5，再进行仿真，则上述三个性能随时间变化的曲线如下：



系统生产率

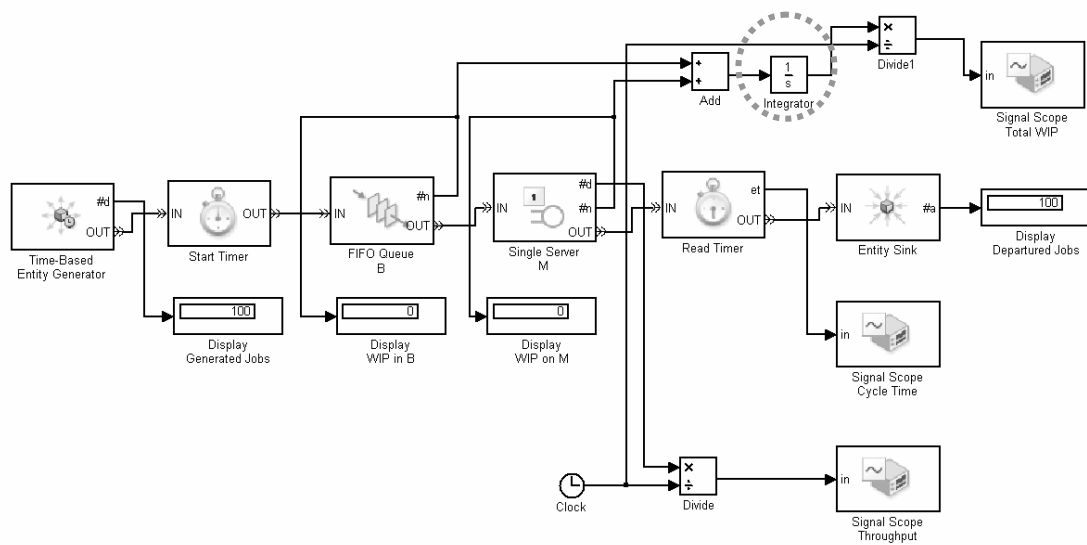


生产周期

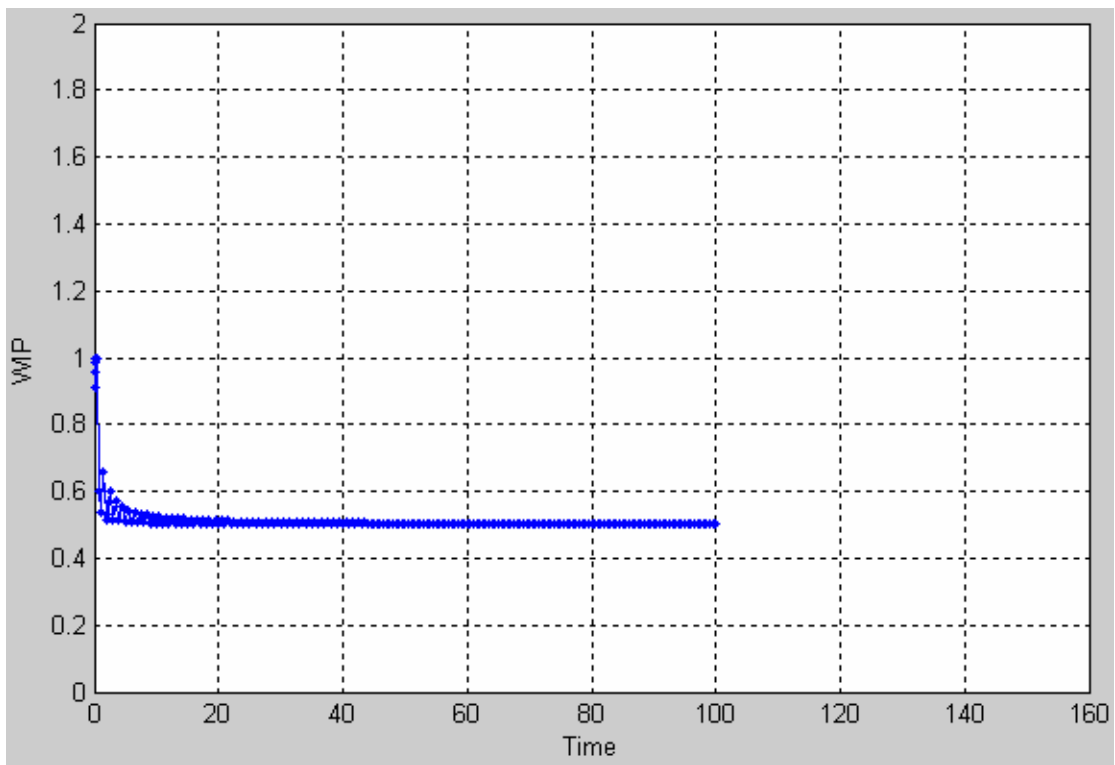


总在制品数量

参数改变后，生产周期缩短为 0.5；在生产稳定后，系统生产率仍然接近 1，这是因为它受零件到达率的影响；而在制品数量在 0 与 1 之间来回波动，在时间意义上的平均在制品水平为 0.5，这一点可通过在原来的模型上加一个积分环节来实现(如下图所示)。



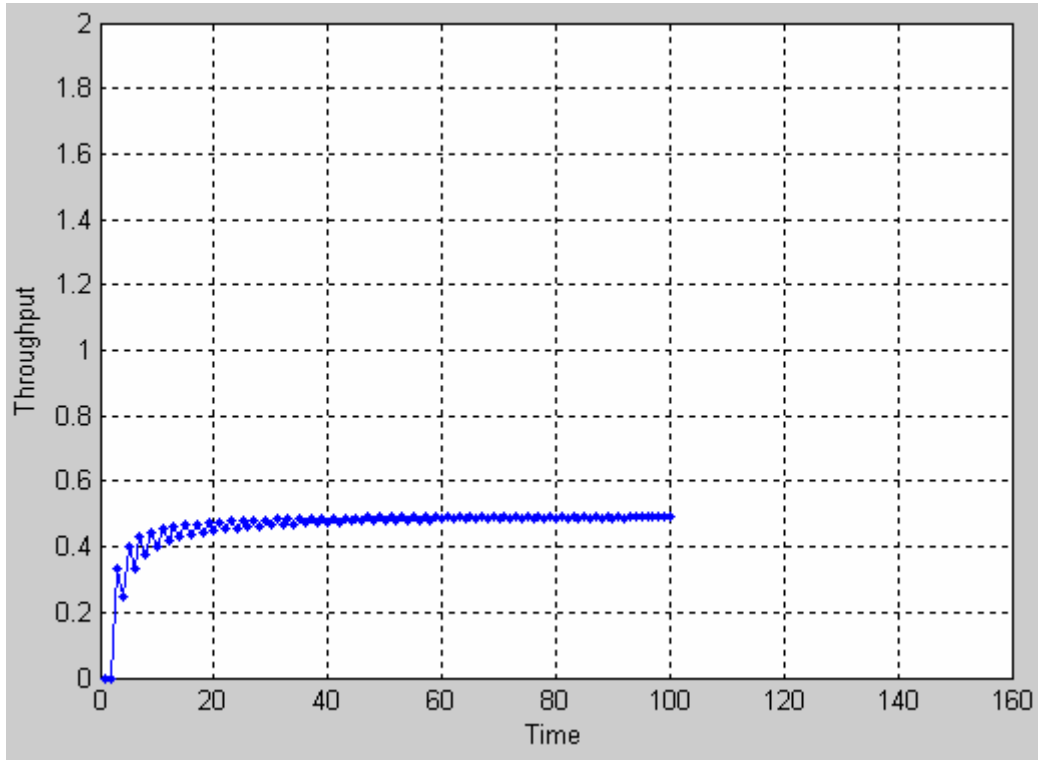
加上积分环节的系统模型



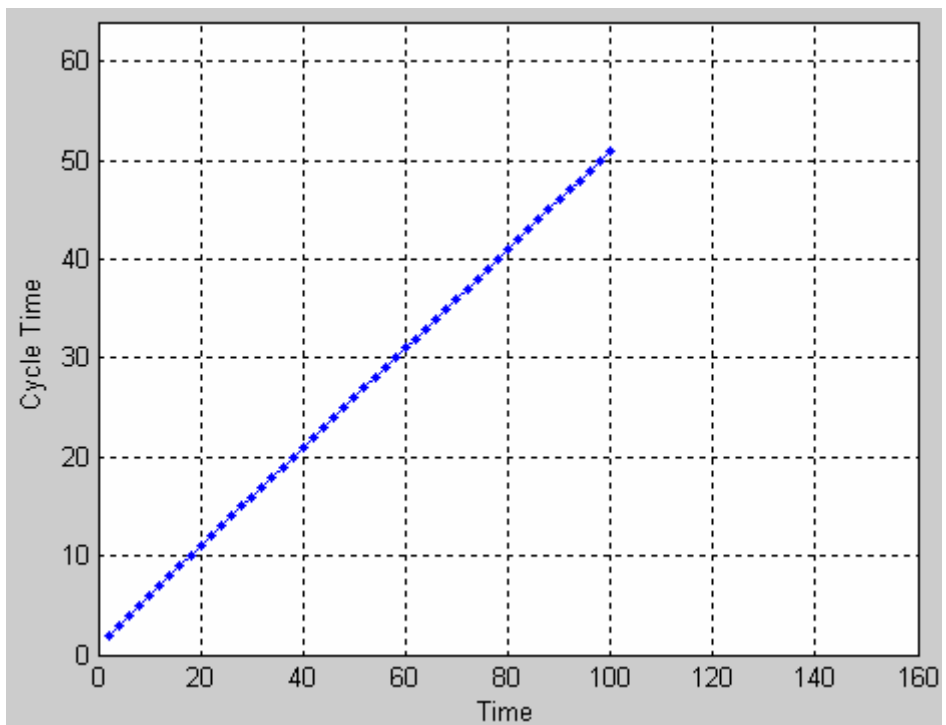
平均在制品水平

继续改变系统参数：把机器生产一个零件的时间改成 2，再进行仿真，则上述三个性能随时间变化的曲线如下：

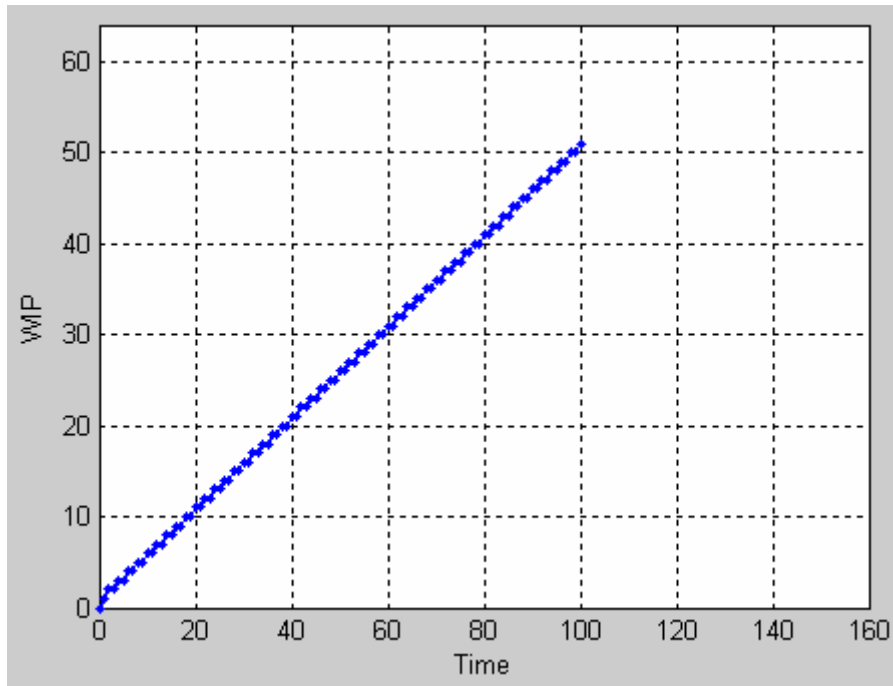




系统生产率

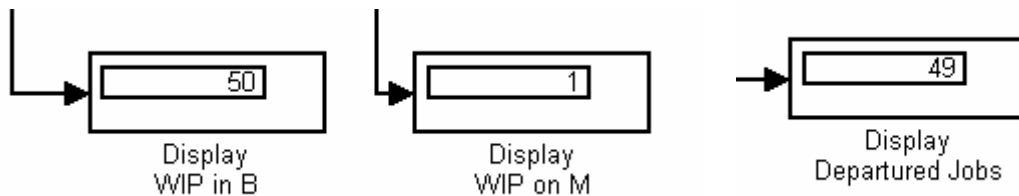


生产周期



总的在制品数量

在这种情况下，系统生产率下降到 0.5；仿真结束时，在制品数量为 51(其中 50 个在缓冲区中，1 个在机器上)；加工完成的零件数为 49，见下图。



生产周期图描述的是已经完成的 49 个零件的生产周期。

由上述三组图可见，前两种情况下，系统是稳定的；第三种情况下，系统是不稳定的，因为在制品数量不断上升。在稳定的情况下，系统生产率(throughput, TP)、生产周期(cycle time, CT)和在制品数量(WIP)之间存在这样的关系：

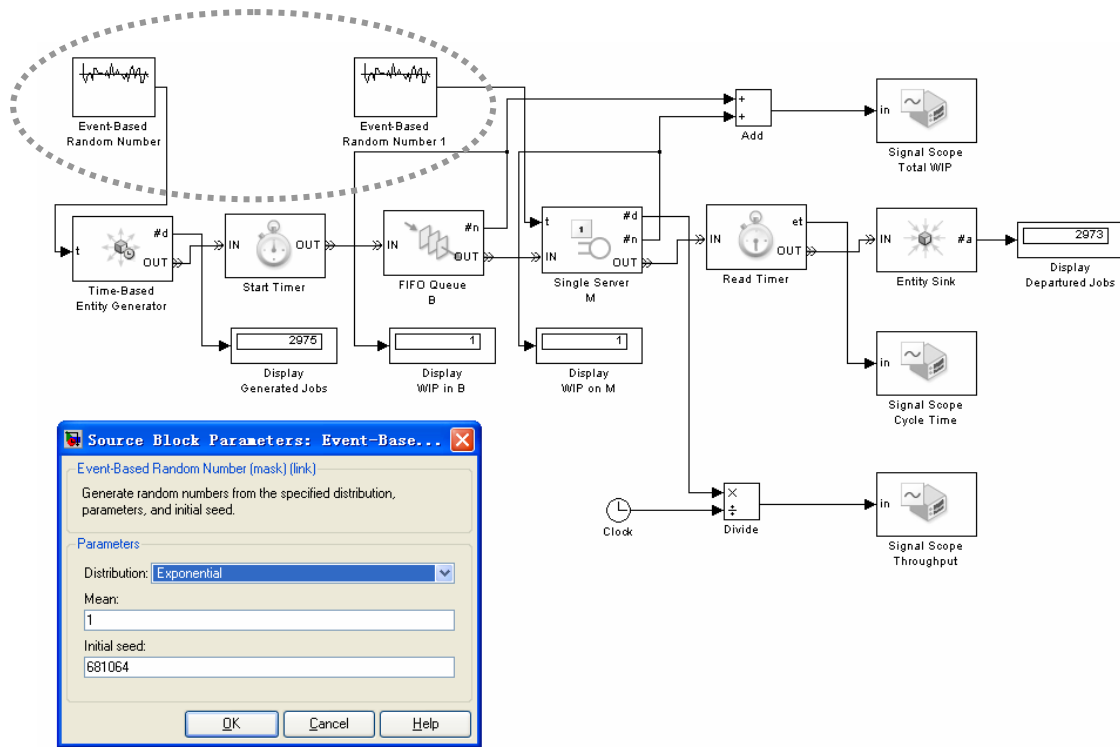
$$WIP=CT \times TP$$

这一关系叫做 **Little's Law**。

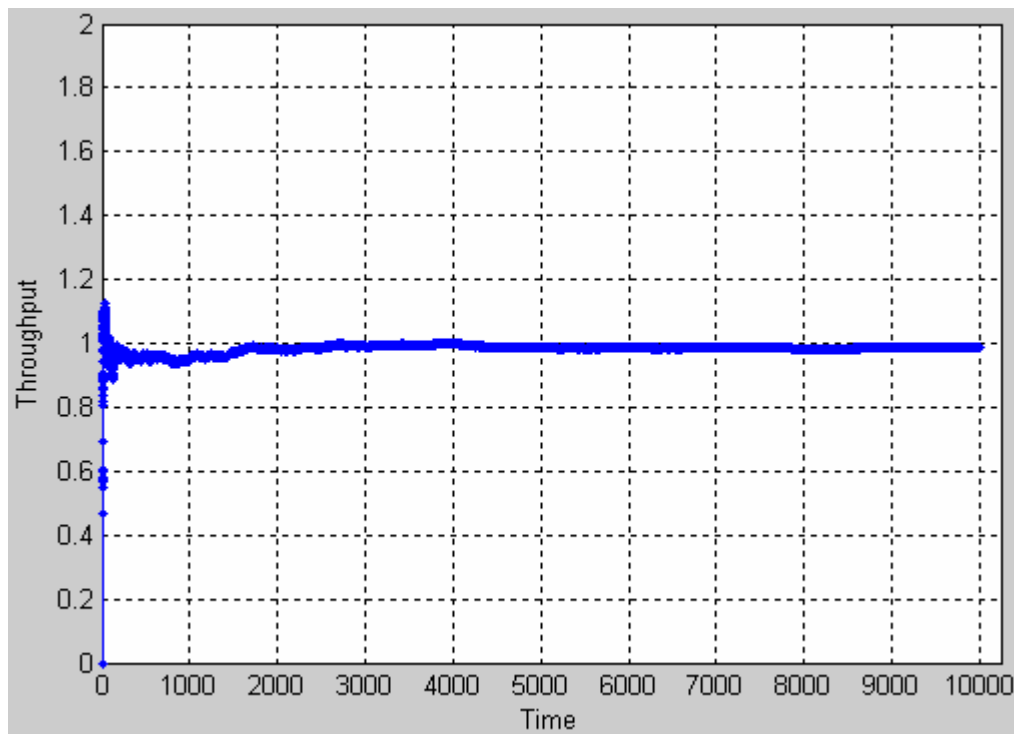
## 2. 随机的无故障单机器制造系统

单机器制造系统的随机因素主要体现在以下几个方面：零件到达时间间隔的随机性，零件加工时间的随机性，机器故障的随机性。一般地，假定零件到达时间间隔、零件加工时间、正常工作时间和故障维修时间服从指数分布，当然，参数各不相同。继续考虑上节的例子，假定零件到达时间间隔服从参数为 1 的指数分

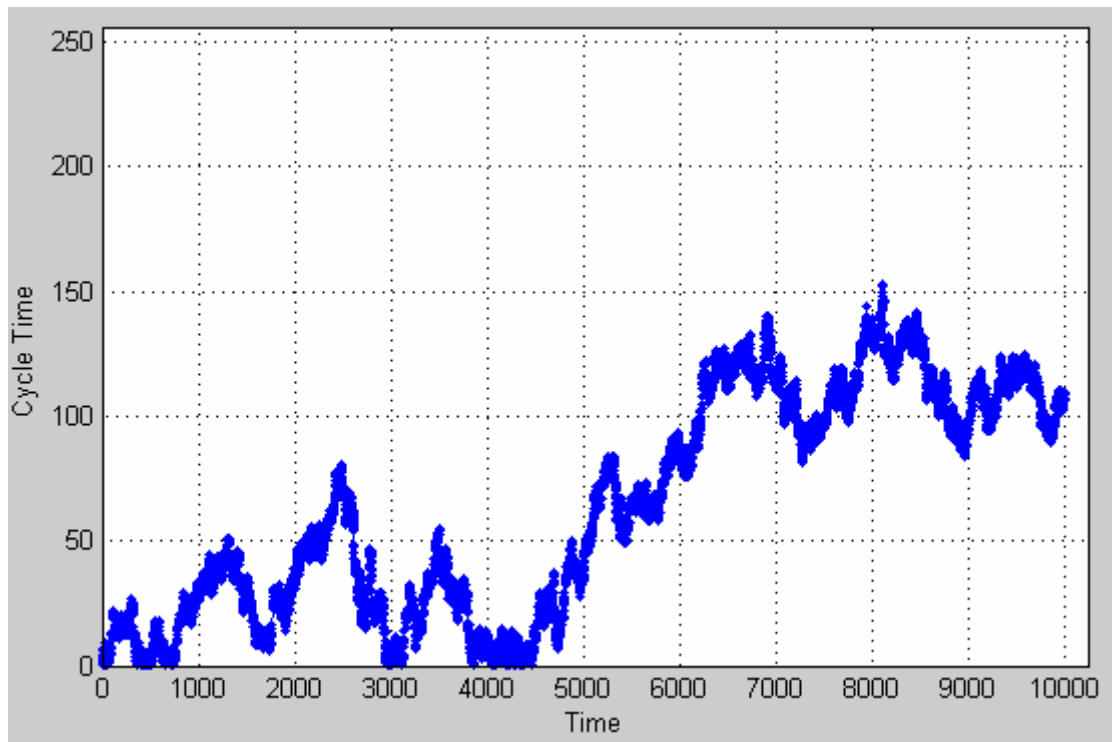
布, 零件加工时间也服从参数为 1 的指数分布, 这里暂不考虑机器故障. 在 SimEvent 环境下, 这样一个系统的模型可建立如下:



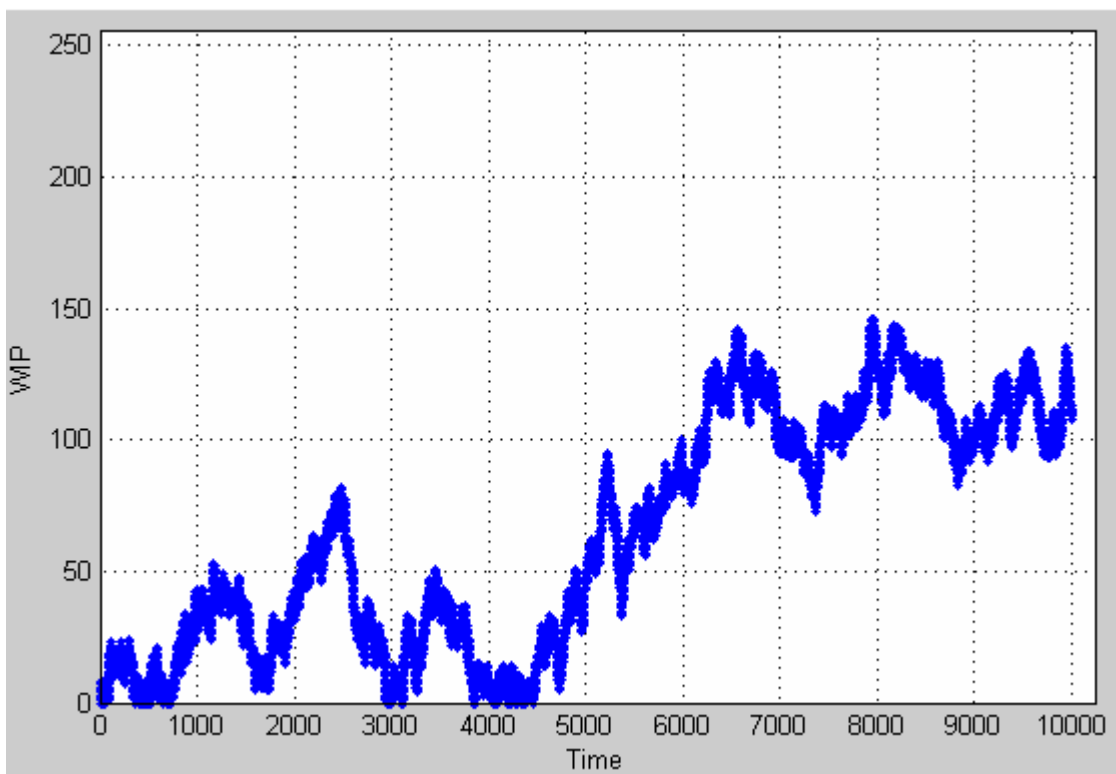
与上节的模型相比, 多了两个随机数发生器, 根据不同的初始种子数产生指数分布的随机数. 经过仿真, 可得系统的生产率、生产周期和在制品数量随时间的变化曲线如下:



系统生产率



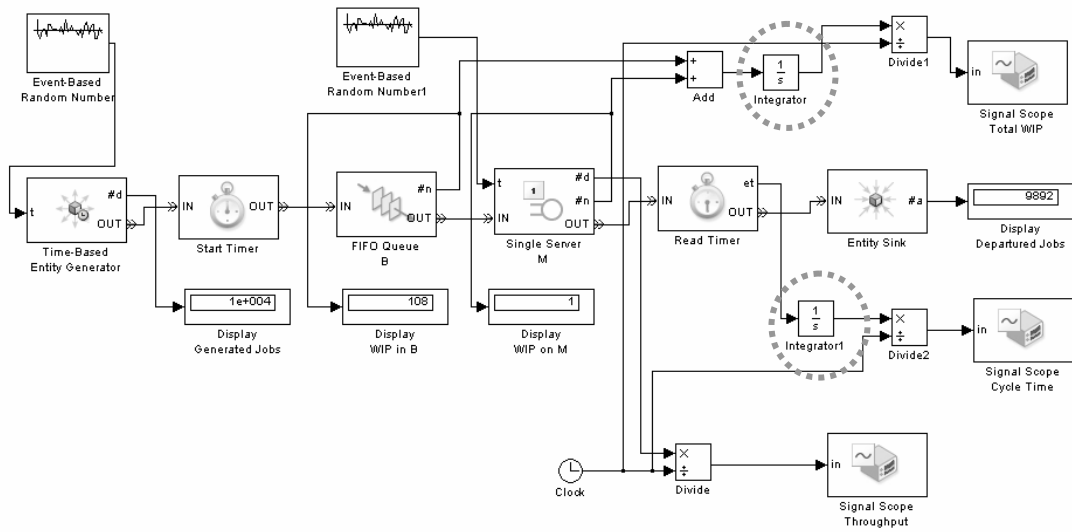
生产周期



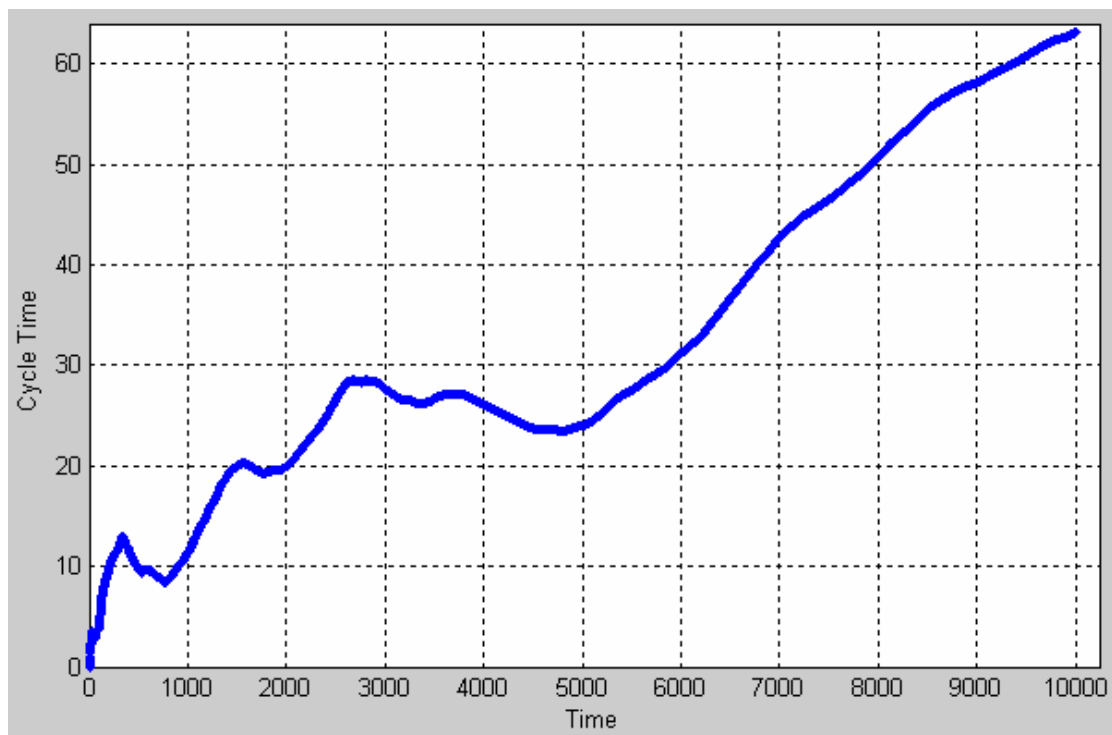
在制品总数

由上述三图可见，系统生产率逐渐趋近于 1，而生产周期和在制品总量却有逐渐增加的趋势，但两者的变化曲线形状基本一致，这也体现了 Little's Law.

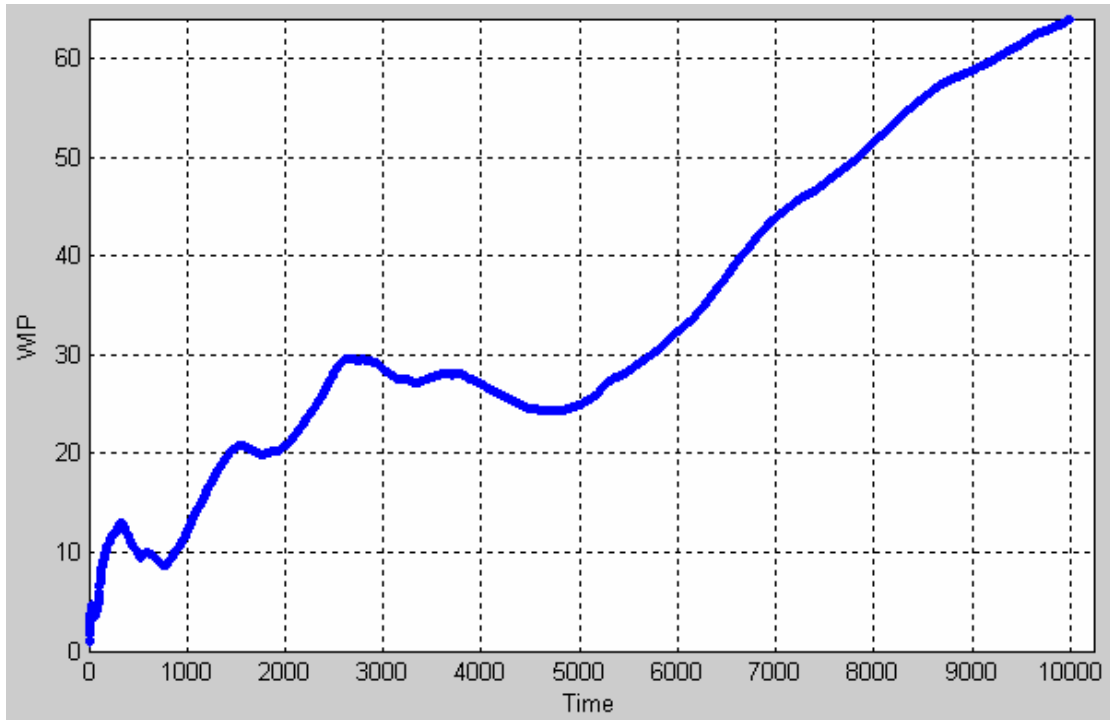
为考察平均在制品和平均生产周期，可以在模型中加上积分环节，如下图所示。



平均生产周期和平均在制品数量分别如下二图所示，二者的形状也是相似的。



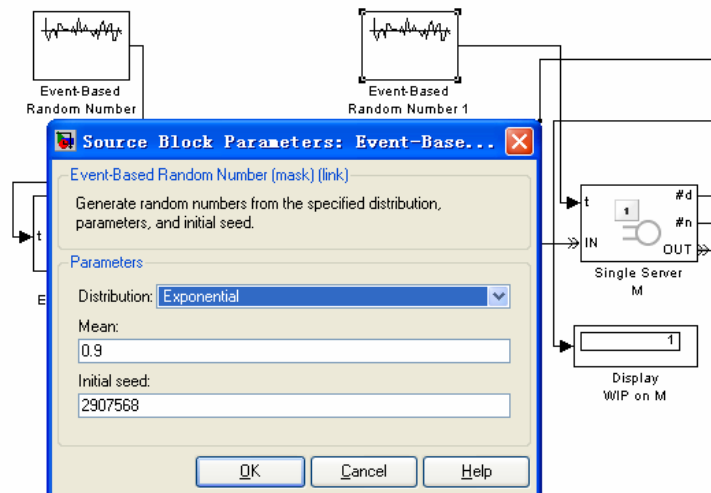
平均生产周期



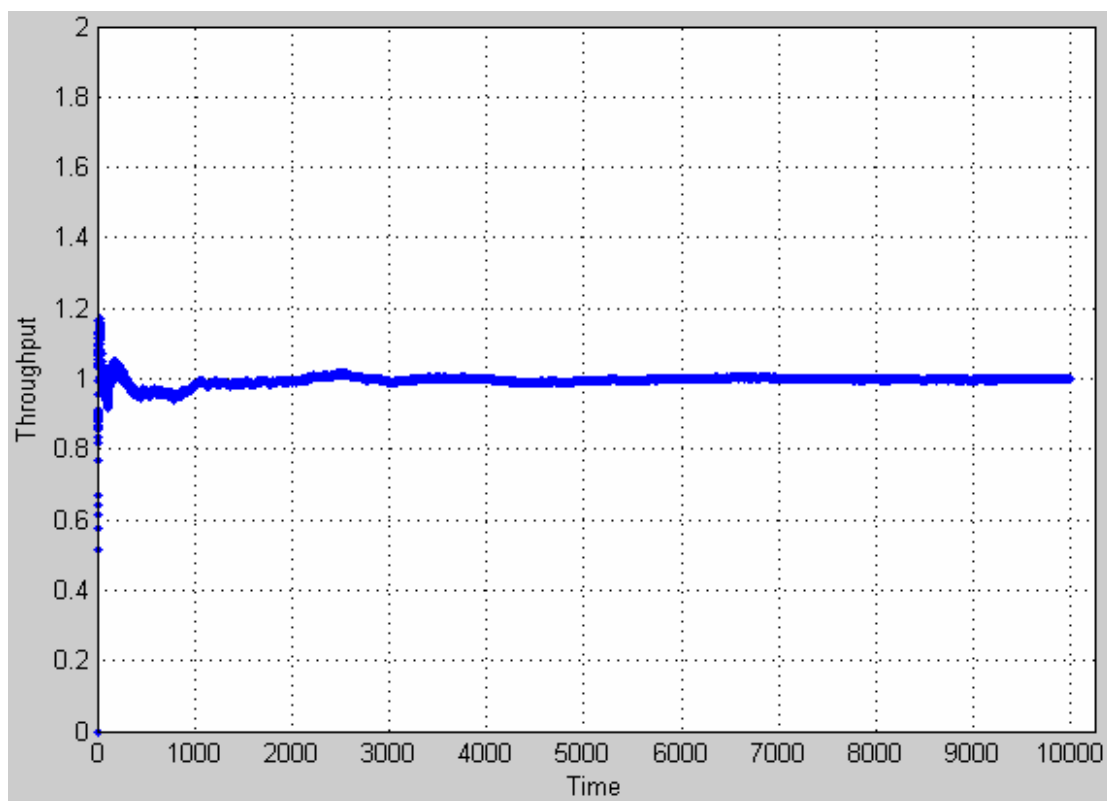
平均在制品数量

从上面的例子可以观察到，在零件到达时间间隔和机器加工一个零件的时间都是服从指数分布的情况下，尽管它们的参数(或者说均值)是相同的，但系统的在制品总量(或者说生产周期)仍然会趋于无穷大，即系统仍然会不稳定。而确定性系统在零件到达时间间隔和机器加工一个零件的时间相同的情况下却是稳定的，这在上节的仿真例子中已经得到验证。这是为什么呢？是因为在随机因素的作用下，零件的加工时间很可能比到达时间间隔长，因此，在制品数量大于 1 的情况经常发生，而在制品数量最小为 1。

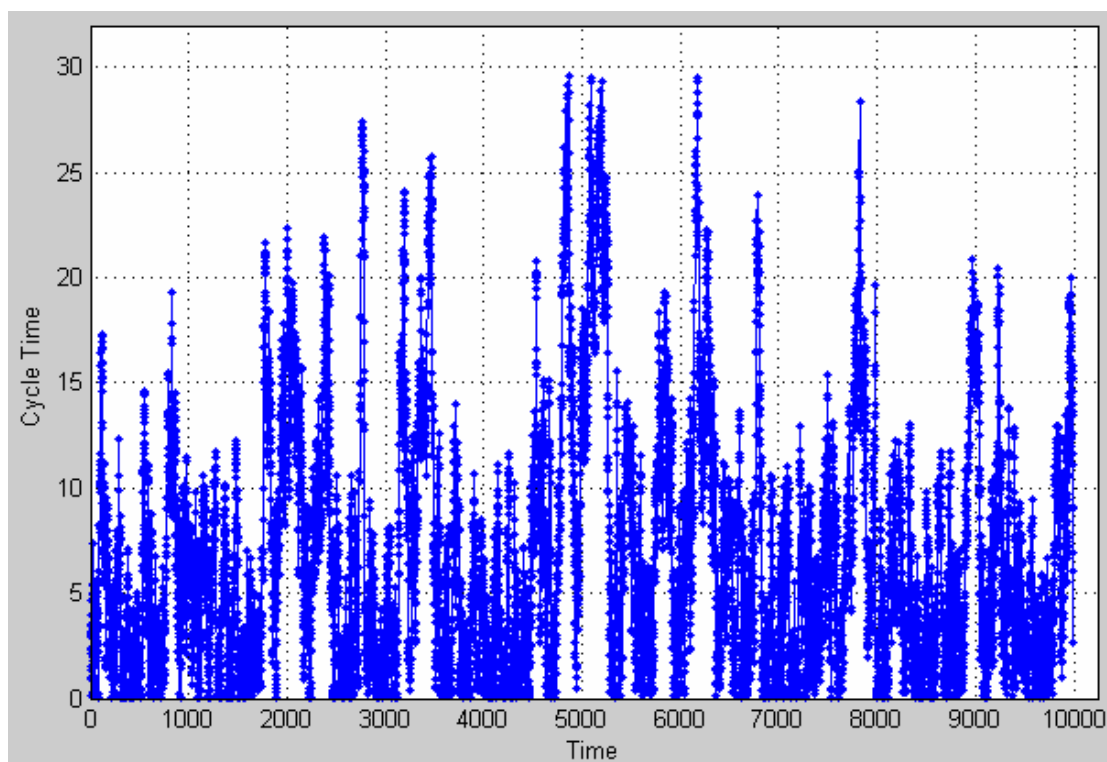
那么，如果把机器生产一个零件的时间缩短(即把它服从的指数分布的均值缩短)，或者说提高机器的生产率，会怎么样呢？在上面的模型中，把机器生产一个零件的时间的均值改为 0.9，如下图所示。



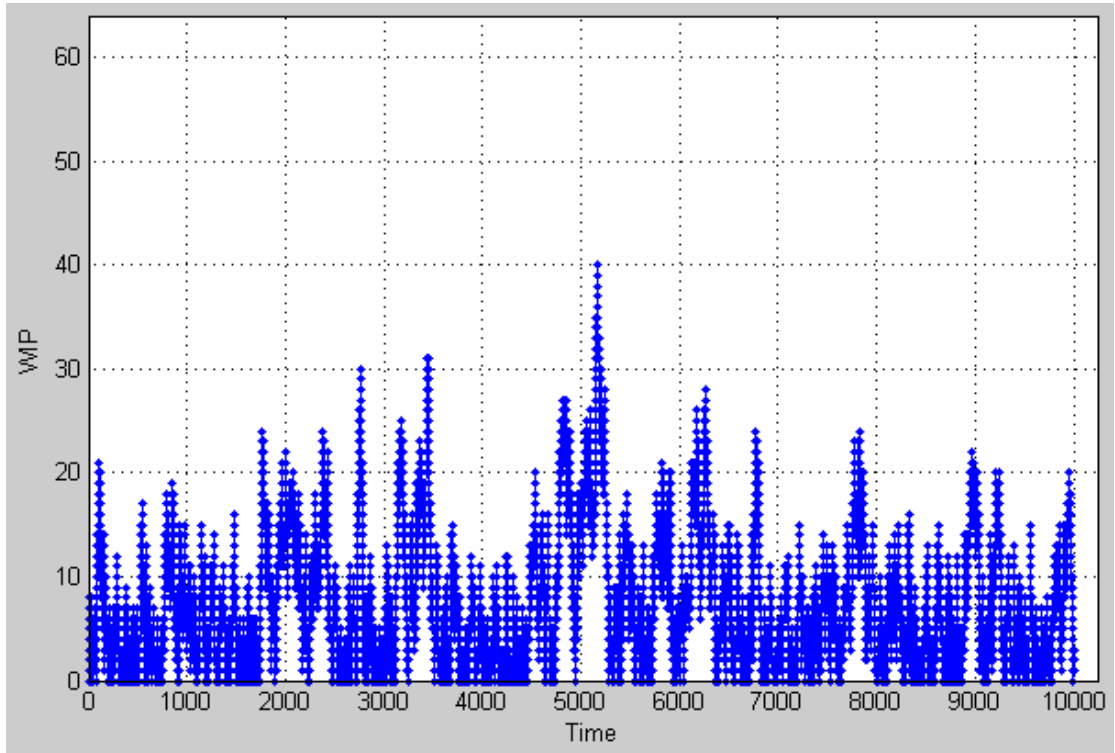
经过仿真, 可得系统生产率、生产周期和在制品数量如下:



系统生产率

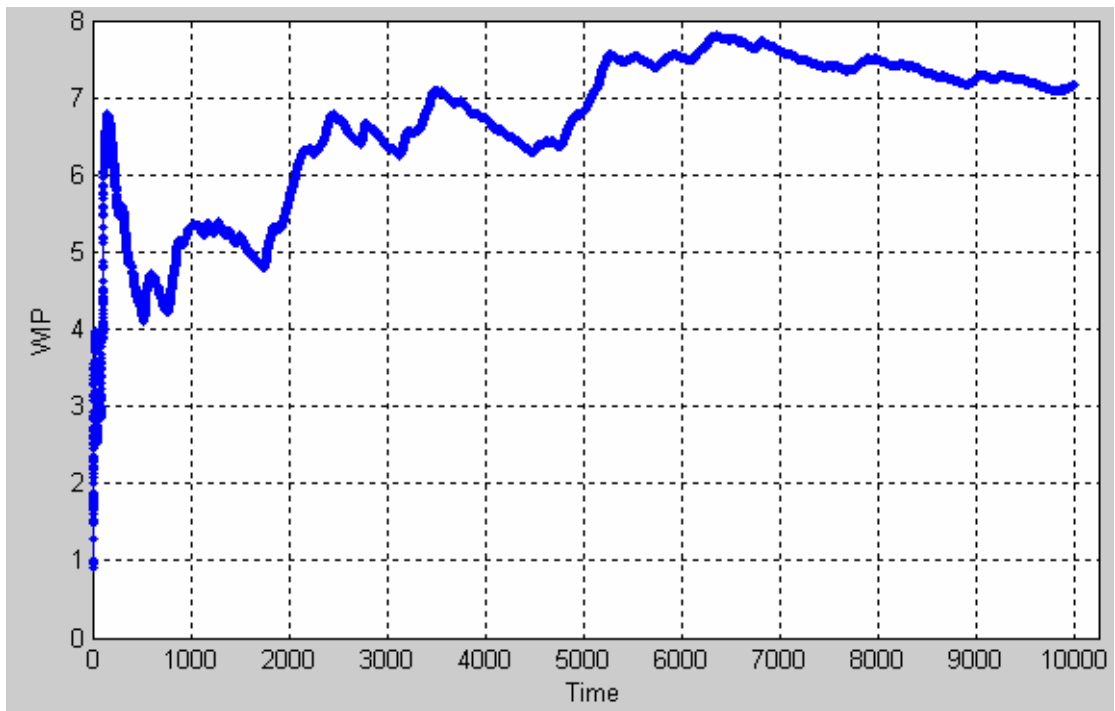


生产周期



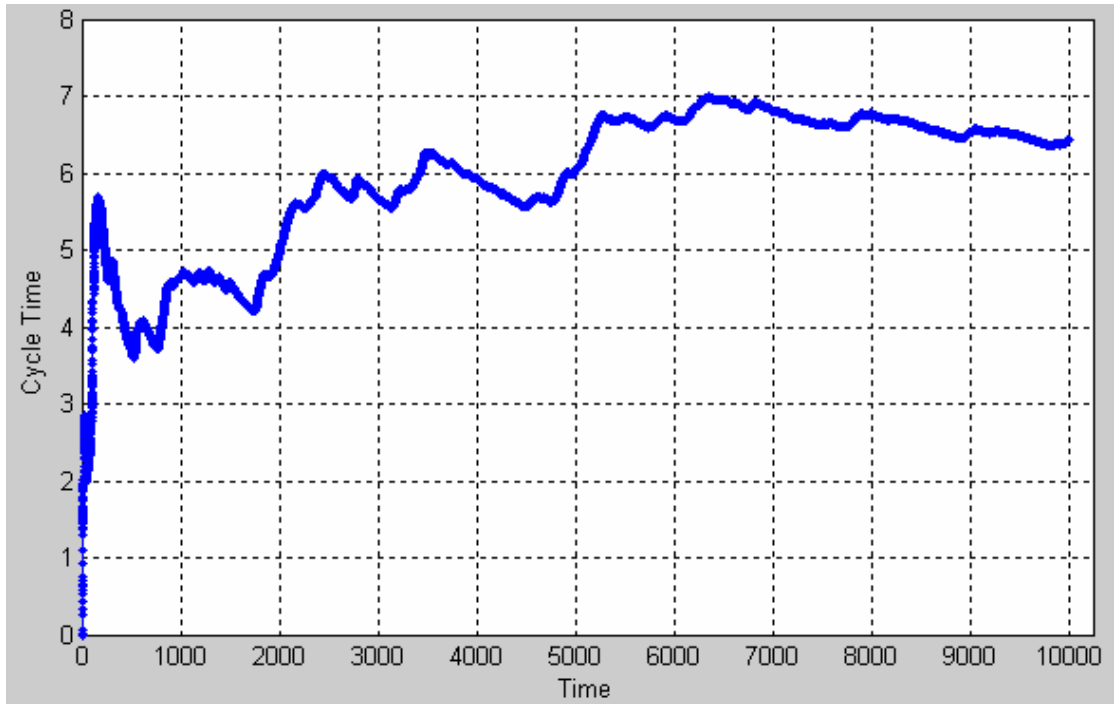
在制品数量

可见，系统生产率仍然趋近于 1，这是因为它受到零件到达率的限制，尽管机器生产率提高了。而在制品数量和生产周期不再趋于无穷大，系统稳定了。通过加上积分环节，通过仿真可得在制品与生产周期的平均值曲线如下：



在制品平均值

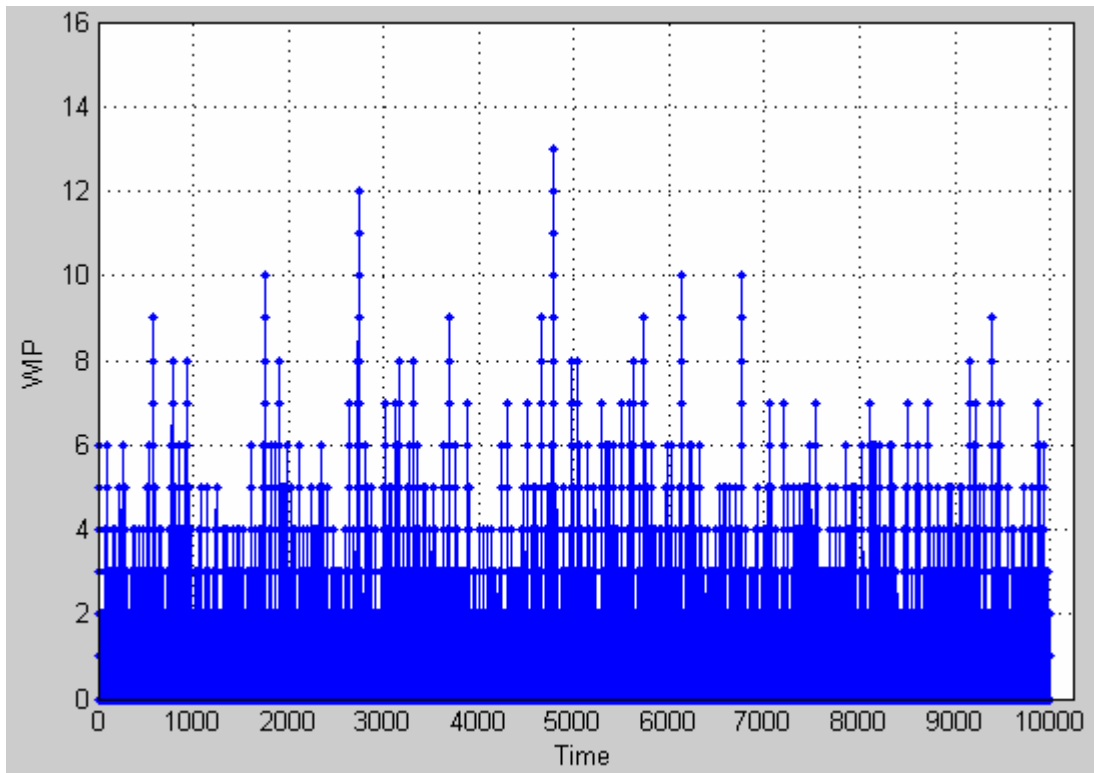




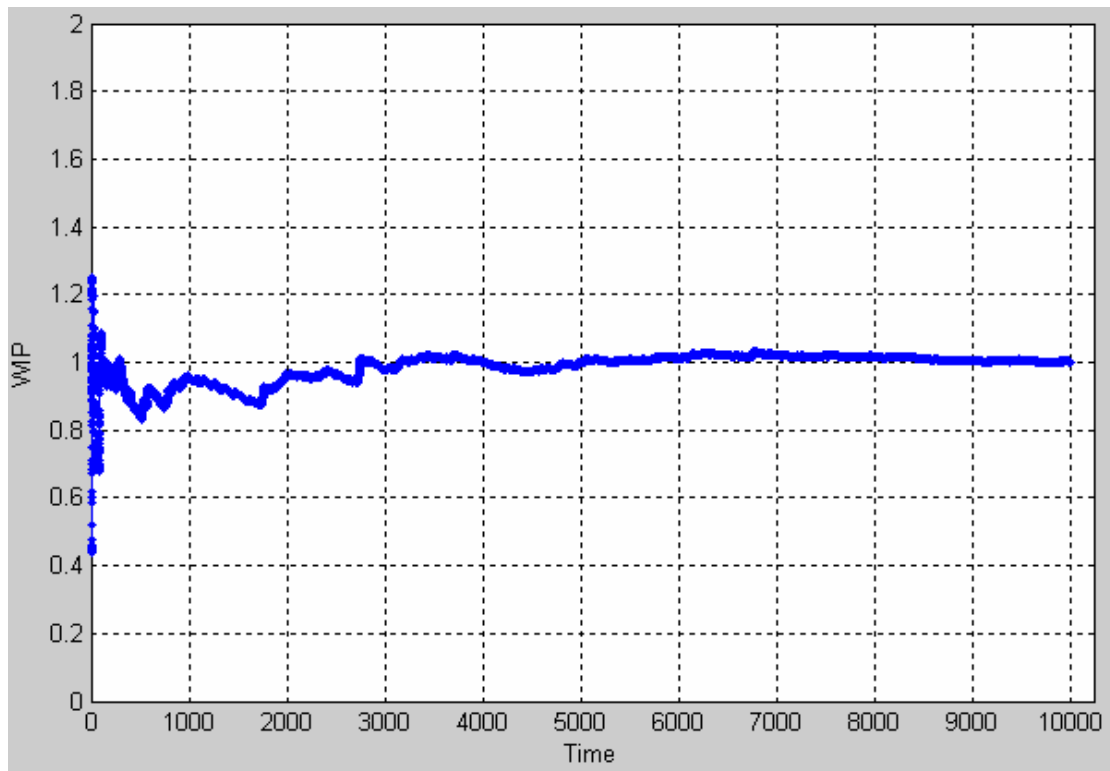
生产周期平均值

可见，在制品平均值稳定在 7.1 左右，生产周期平均值稳定在 6.5 左右，似乎不符合 Little's Law. 这是由于系统生产周期是从零件进入生产系统算起的，不受到达时间间隔限制.

进一步缩短机器加工零件的时间均值至 0.5，经过仿真，可得在制品曲线如下：



在制品平均值曲线如下:



可见,系统是稳定的.与加工时间均值为 0.9 的情况相比,在制品数量的平均值进一步减少,趋近于 1,波动也减弱了.

从上面的例子,可以猜想这样一个结论:对于随机的单机器制造系统,只要加工时间的均值小于零件到达时间间隔的均值,系统就可以稳定.这个结论是正确的,其理论上证明将在基于排队网络的制造系统分析课程中介绍.

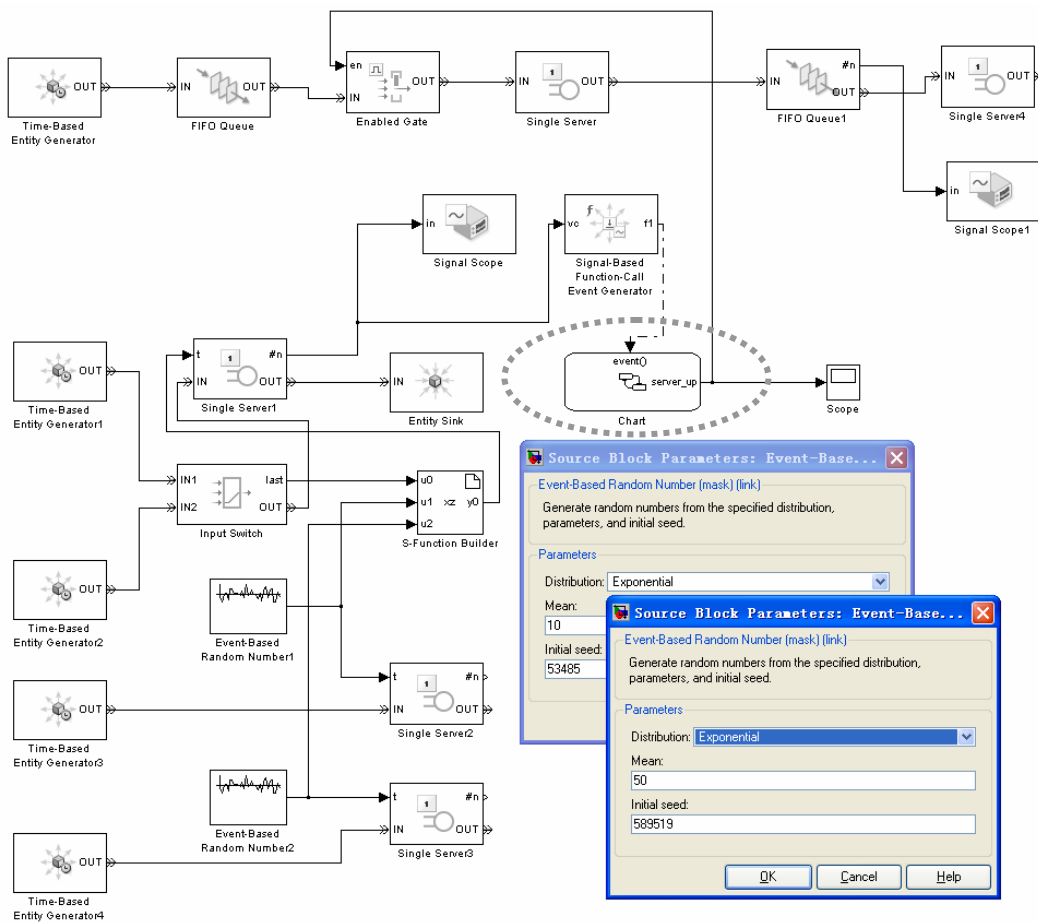
## 第 4 章

# 基于仿真的制造系统分析与控制(2)

本章将继续上一章的内容,介绍有故障的单机器制造系统、多机器串行生产线、以及可重入生产系统的分析与控制.

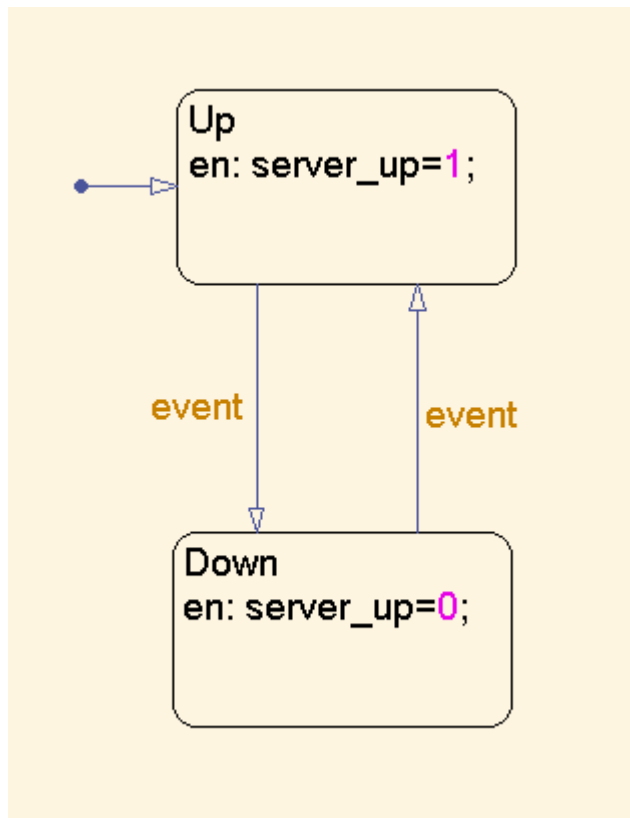
### §4.1 基于仿真的有故障单机器制造系统分析与控制

设有一单机器制造系统,机器正常工作时间服从指数分布,均值为 50;机器的故障时间也服从指数分布,均值为 10;机器加工一个零件所需时间为 1,加工完成的零件存储在缓冲区中;每 1 个单位时间,缓冲区中的 1 个零件被取走,以满足需求.在 SimEvents 环境中,可建立这样如下图所示的系统模型.

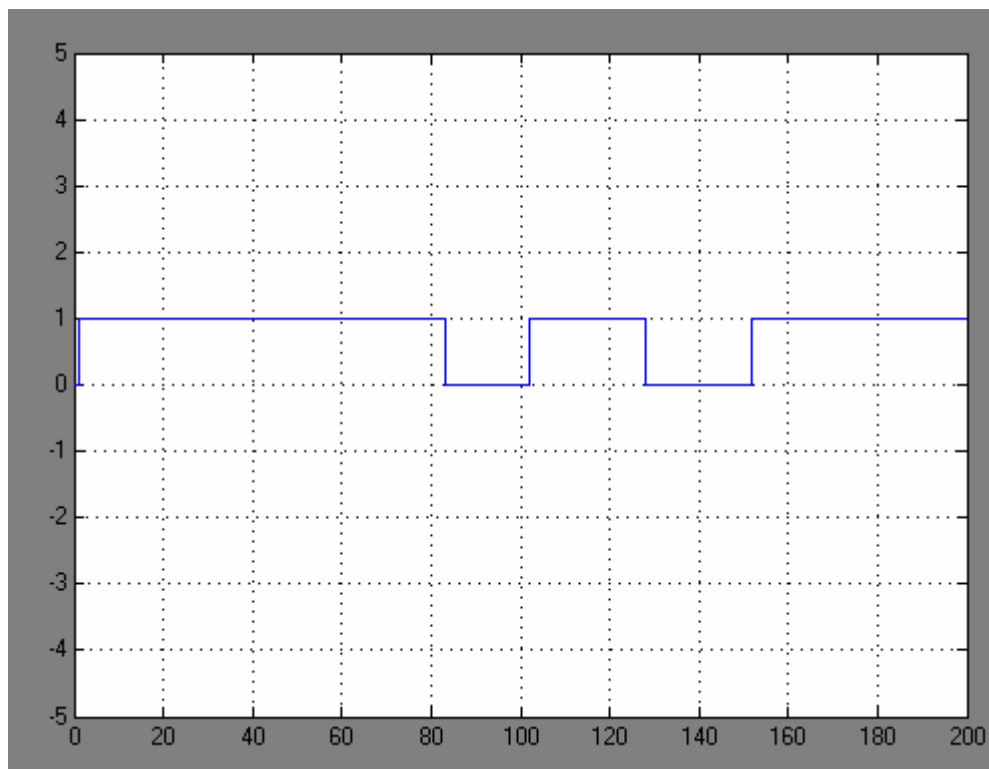


其中包含一个用 Stateflow 建立的机器状态变化模型 (见上图中虚线椭圆标出的

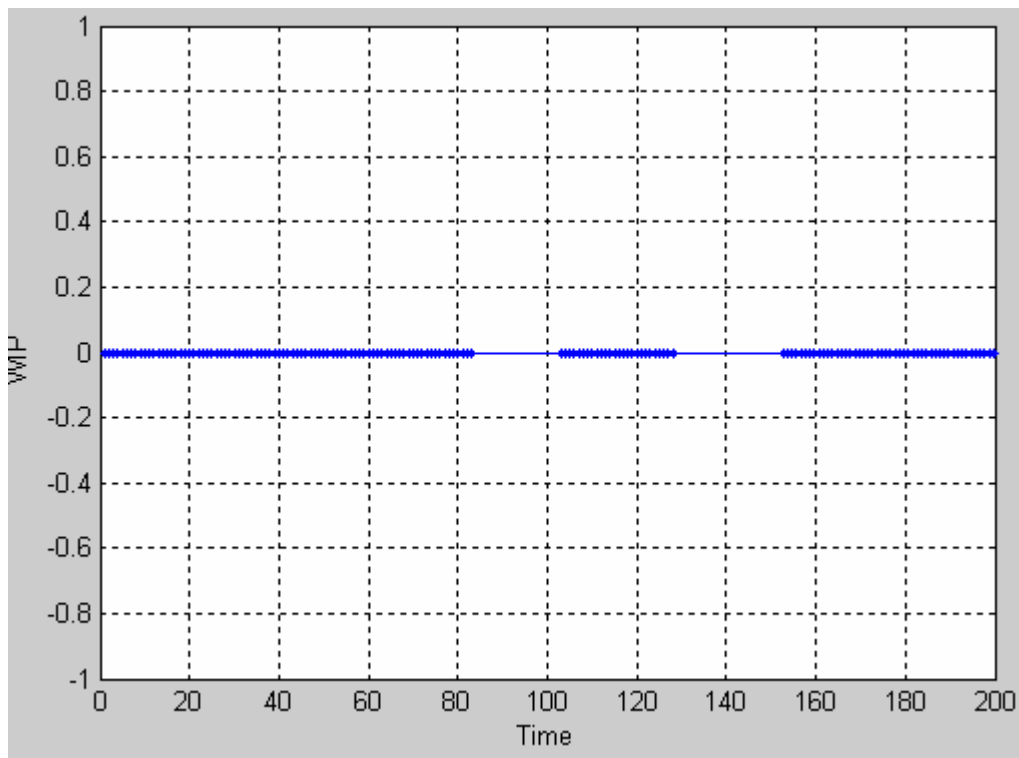
部分), 其内部结构如下图所示.



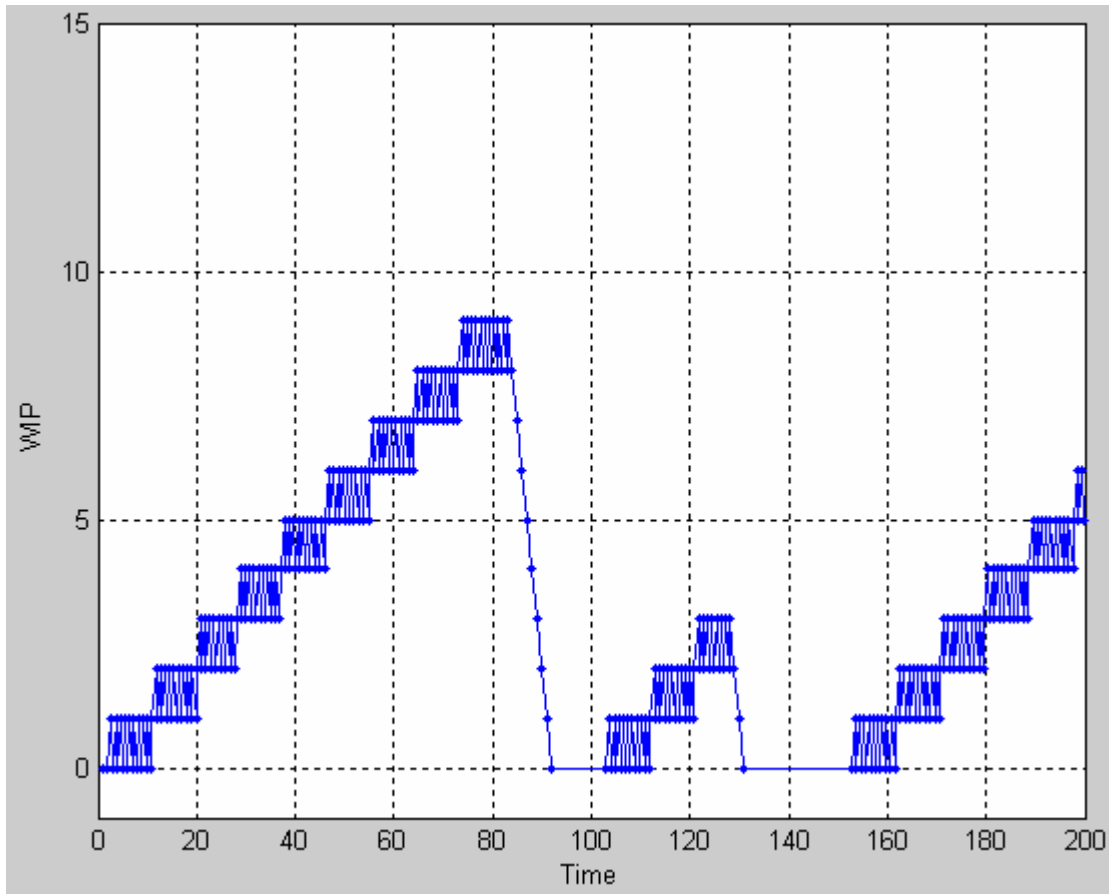
通过仿真, 机器状态随时间变化的情况为:



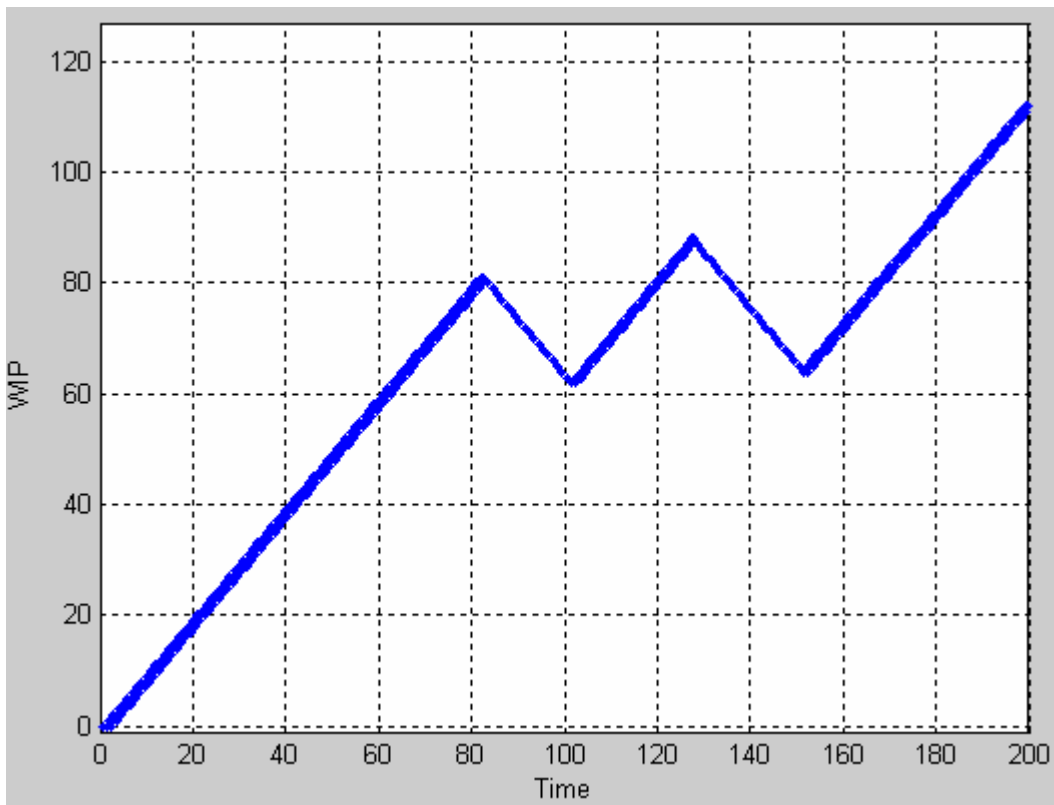
其中,1 表示机器正常工作,0 表示机器发生故障.  
缓冲区中在制品数量随时间变化的情况如下图所示.



如果把机器的生产一个零件的时间缩短到 0.9, 则缓冲区中在制品数量随时间变化的曲线为:



进一步把机器生产一个零件的时间缩短到 0.5，则缓冲区中在制品数量随时间变化的曲线为：

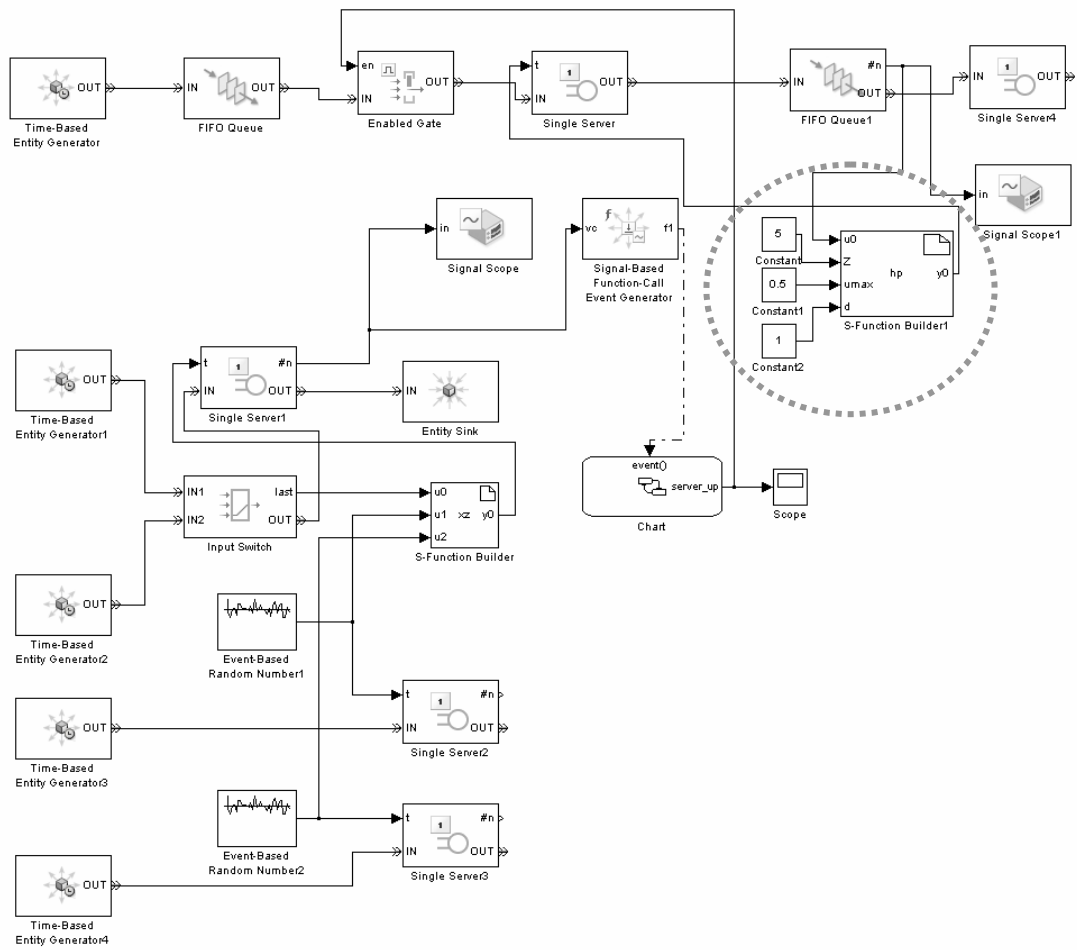


可见在制品数量呈现无限上升的趋势. 而曲线中下降的部分是由于机器的故障造成的.

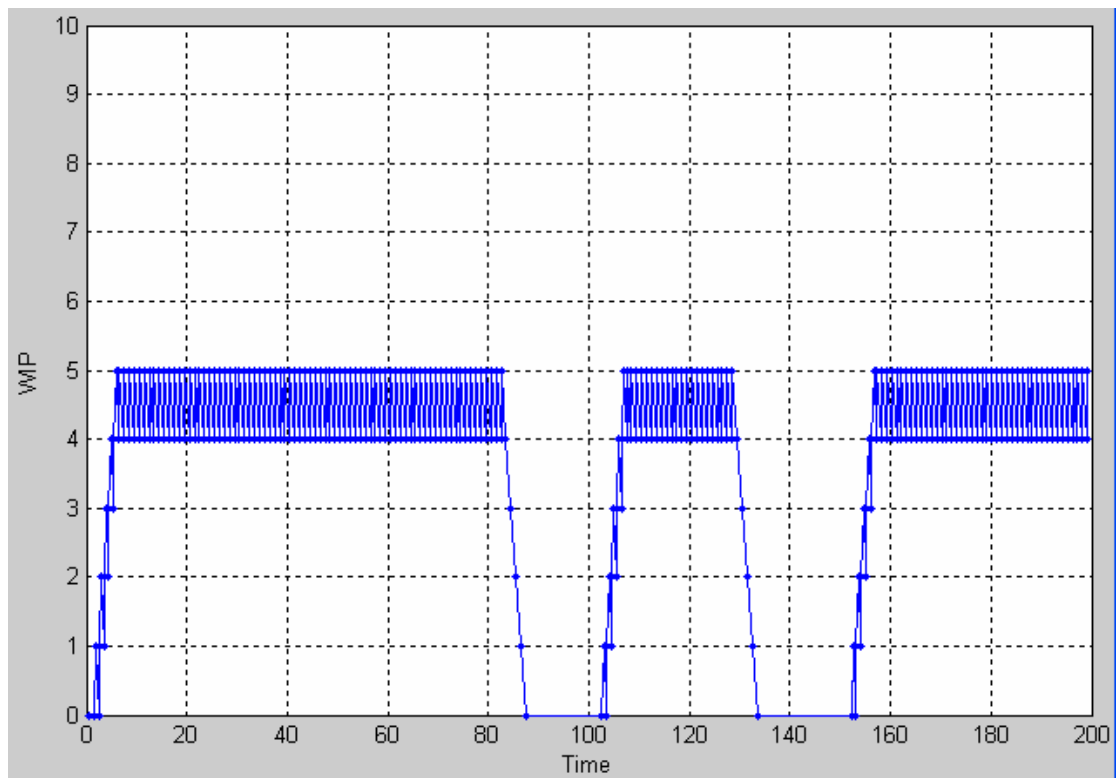
为了保证系统超产和欠产的成本尽量小, 需要对机器的生产率进行控制. 由于机器在正常状态下生产 1 个零件所需的时间是 0.5 个单位时间(即生产率为 2), 而零件的需求率为每单位时间 1 个, 很明显, 生产率大于需求率. 考虑到机器的正常与故障时间均值, 可估算出机器的有效生产率为  $2*50/(50+5)=1.82$ , 仍然大于生产率. 因此需求是可以满足的. 正常工作时, 机器的生产率可在  $[0, 2]$  区间内变动, 如果生产率一直较大, 则在制品数量会不断增加; 如果生产率一直较小, 例如, 刚好等于需求率, 则在制品为 0, 但这时一旦机器发生故障, 由于没有在制品作为缓冲, 则可能出现欠产情况. 因此, 还需要保持一定水平的在制品. 所以, 单机器制造系统通常采取如下的 **Heging Point** 控制策略.

$$u = \begin{cases} 0, & \text{if } x > Z \text{ and } \alpha = 1 \\ d, & \text{if } x = Z \text{ and } \alpha = 1 \\ \mu, & \text{if } x < Z \text{ and } \alpha = 1 \\ 0, & \text{if } \alpha = 0 \end{cases}$$

这里,  $\mu$  为机器在正常状态下的最大生产率;  $d$  为需求率为;  $\alpha = 1$  表示机器是正常的;  $\alpha = 0$  表;  $x$  为系统内的在制品状态;  $u$  为  $t$  机器的实际生产率. 加上这样的控制之后, 有故障单机器制造系统在 **SimEvent** 中的模型如下图所示. 图中虚线椭圆标出的部分为控制器, 它是用 **Simulink** 中的 **S** 函数实现的.

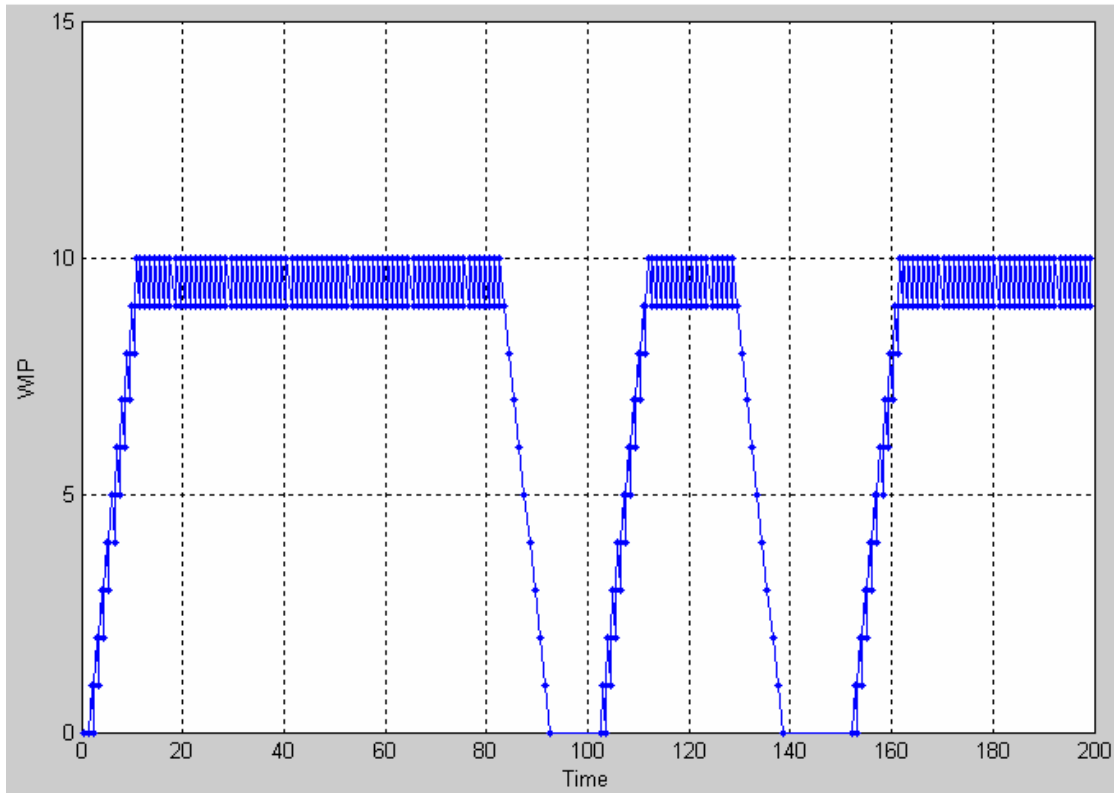


当取  $\mu = 2$ ,  $d = 1$  以及  $Z = 5$  时, 系统在制品数量随时间的变化如下:

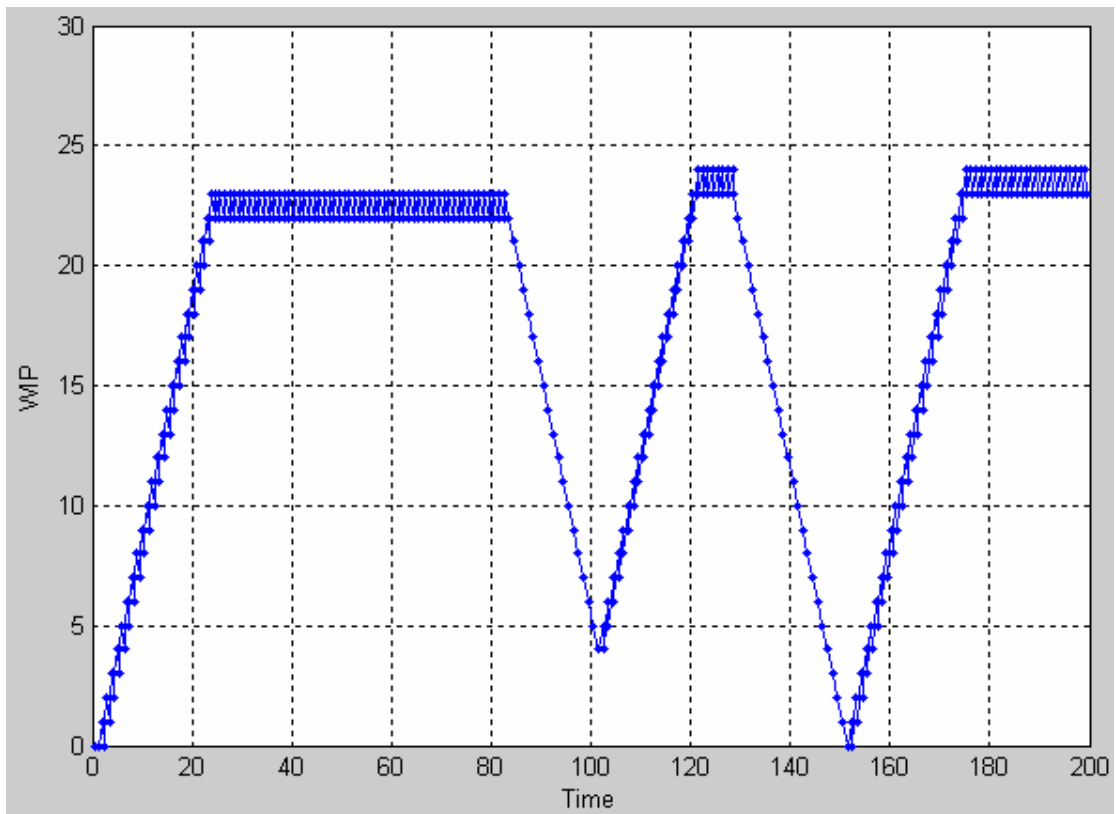




由上图可见, 还有在制品为 0 的情况, 即欠产. 若增大 Hedging Point 的值, 取  $Z=10$ , 可见欠产时间有所缩短.

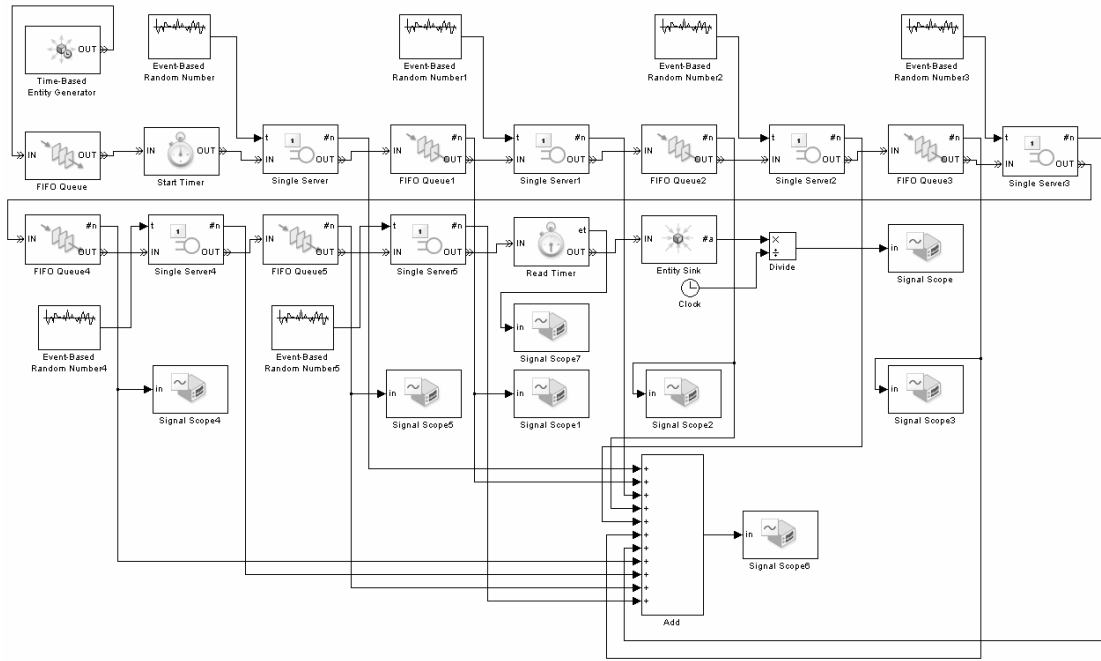


继续增大 Hedging Point 的值, 取  $Z=23$ , 则欠产可基本消除. 其代价是, 在制品数量多, 超产成本增加.

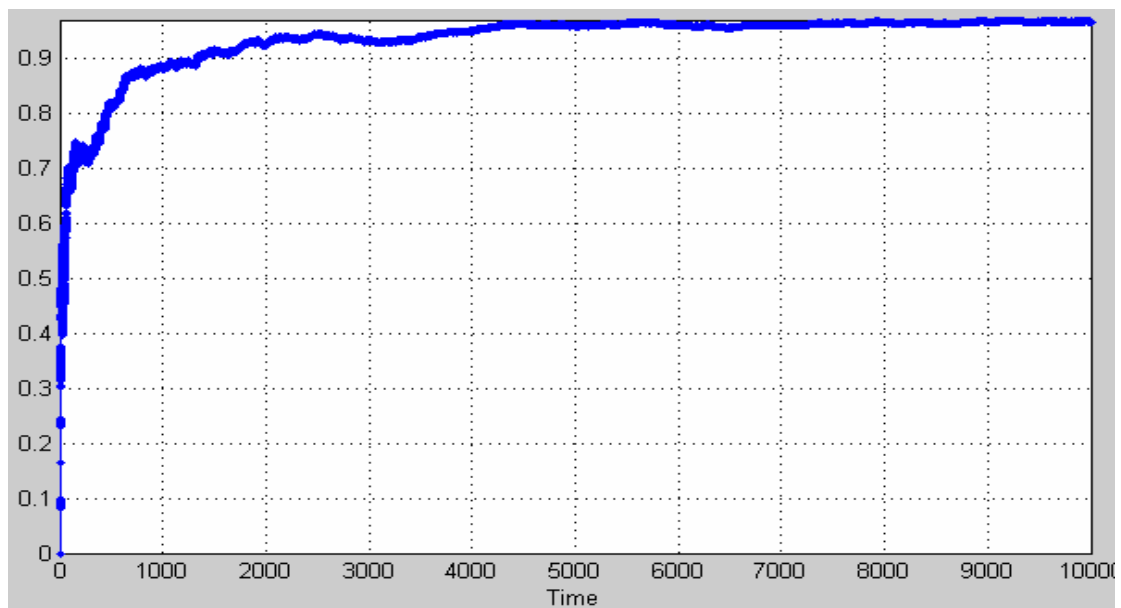


## §4.2 基于仿真的多机器串行生产线分析

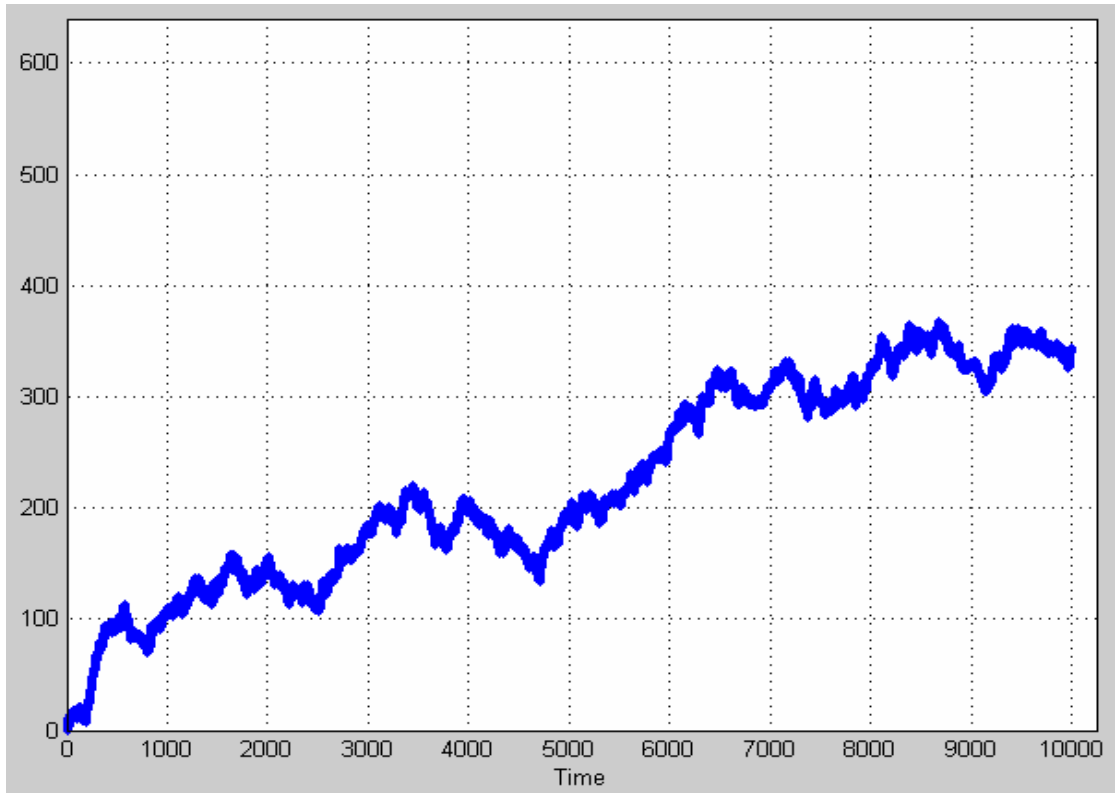
设有一多机器串行生产线，一种零件从生产线的起始端进入生产线，依次经过每台机器，完成若干道工序后离开生产线。在 SimEvents 环境下，建立该生产线的模型如下。



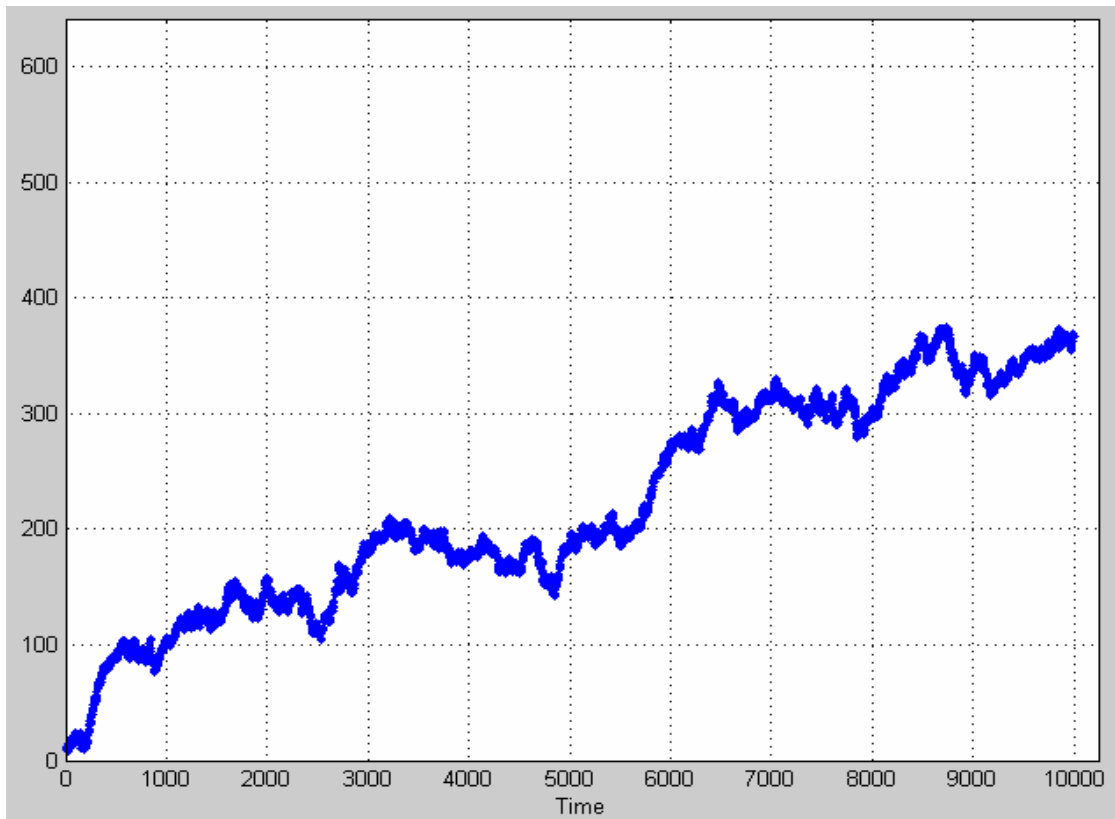
上图所示为一条由 6 台机器和 5 个缓冲区构成的串行生产线，零件在每台机器上的加工时间是独立同分布的随机变量，服从参数为 1 的指数分布。零件进入生产线的间隔为 0.1。缓冲区容量为无穷大。仿真 10000 个时间单位后，其生产率、生产周期和在制品数量随时间变化的情况如下图所示。



系统生产率随时间变化的情况。



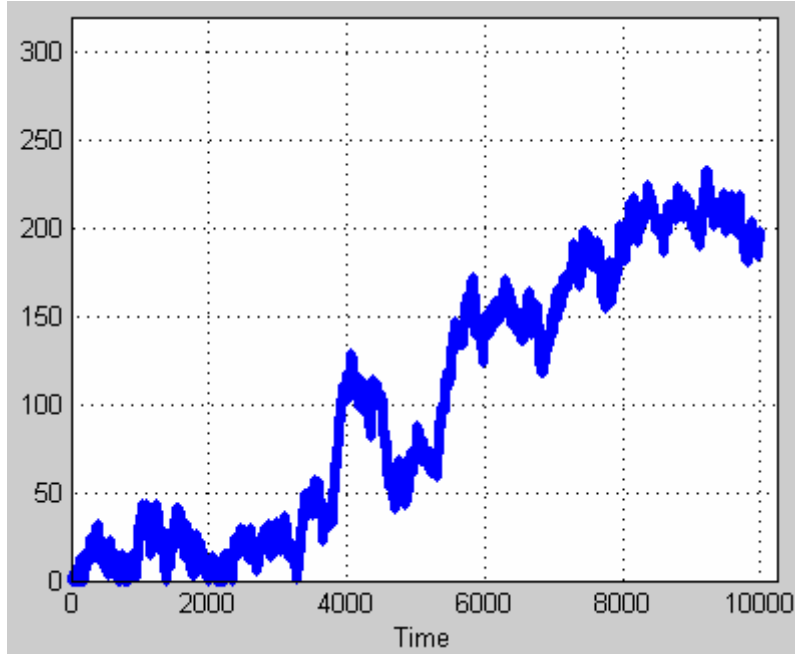
生产线中在制品总量随时间变化的情况.



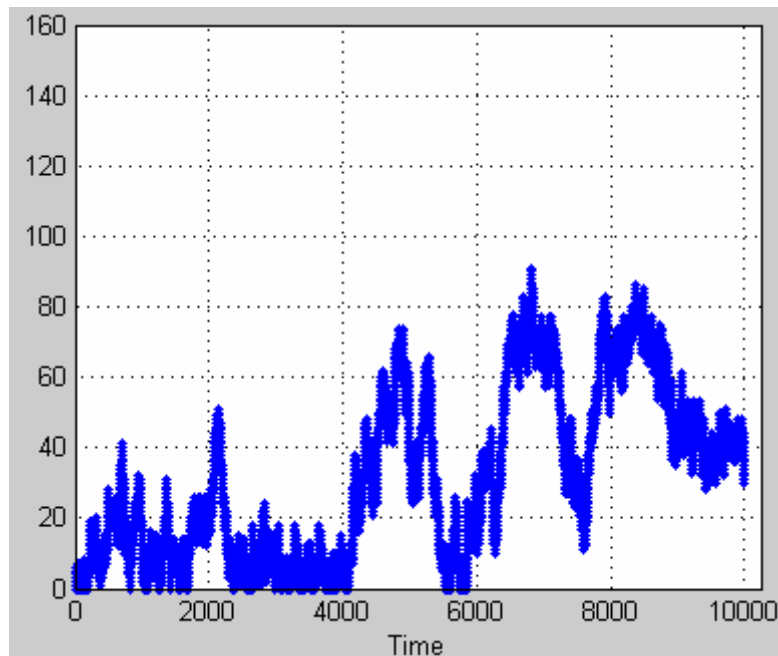
生产周期随时间变化的情况

由上两图可见, 在制品总量不断增加, 系统不稳定. 在制品总量和生产周期的变化曲线形状基本一致, 而生产率趋近于 1. 这是符合 Little's Law 的.

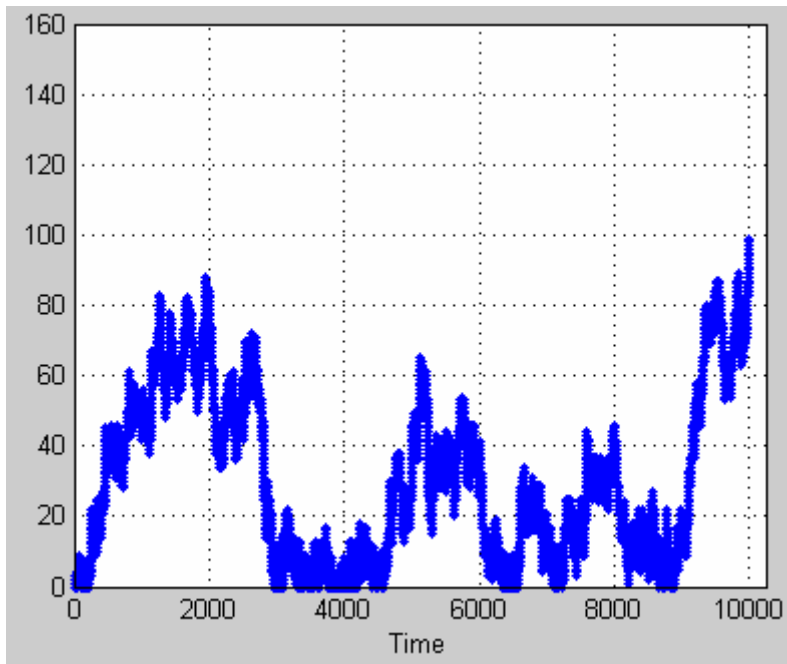
继续观察每个缓冲区中在制品数量的变化情况:



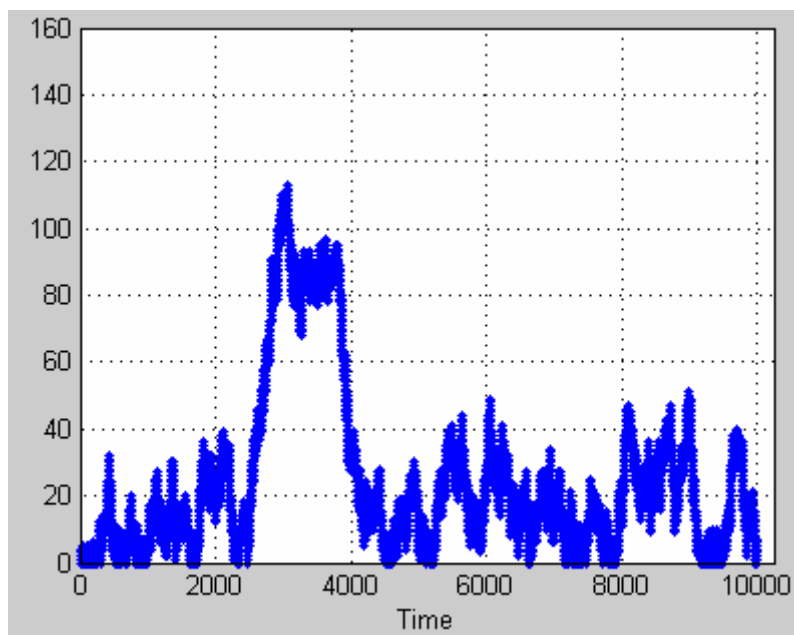
缓冲区 B1(模型中的 Queue1)中在制品数量的变化情况



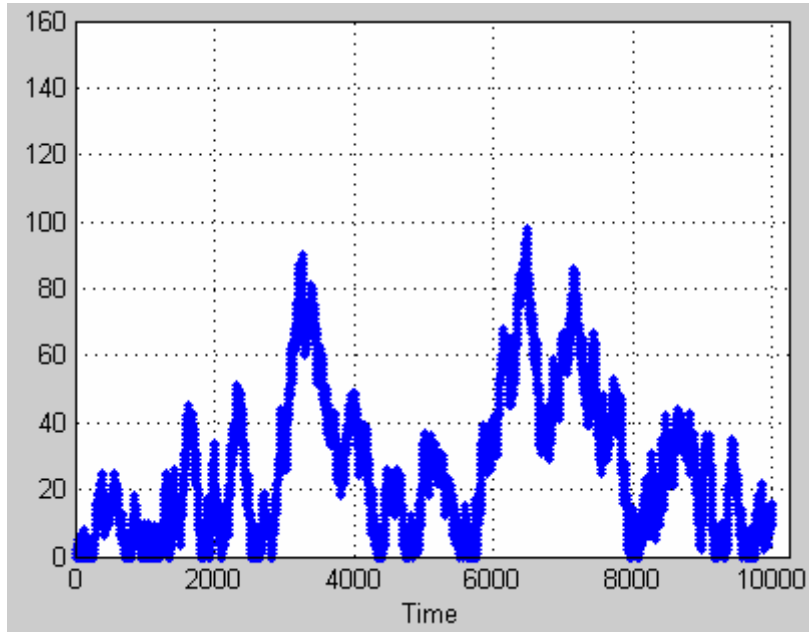
缓冲区 B2(模型中的 Queue2)中在制品数量的变化情况



缓冲区 B3(模型中的 Queue3)中在制品数量的变化情况

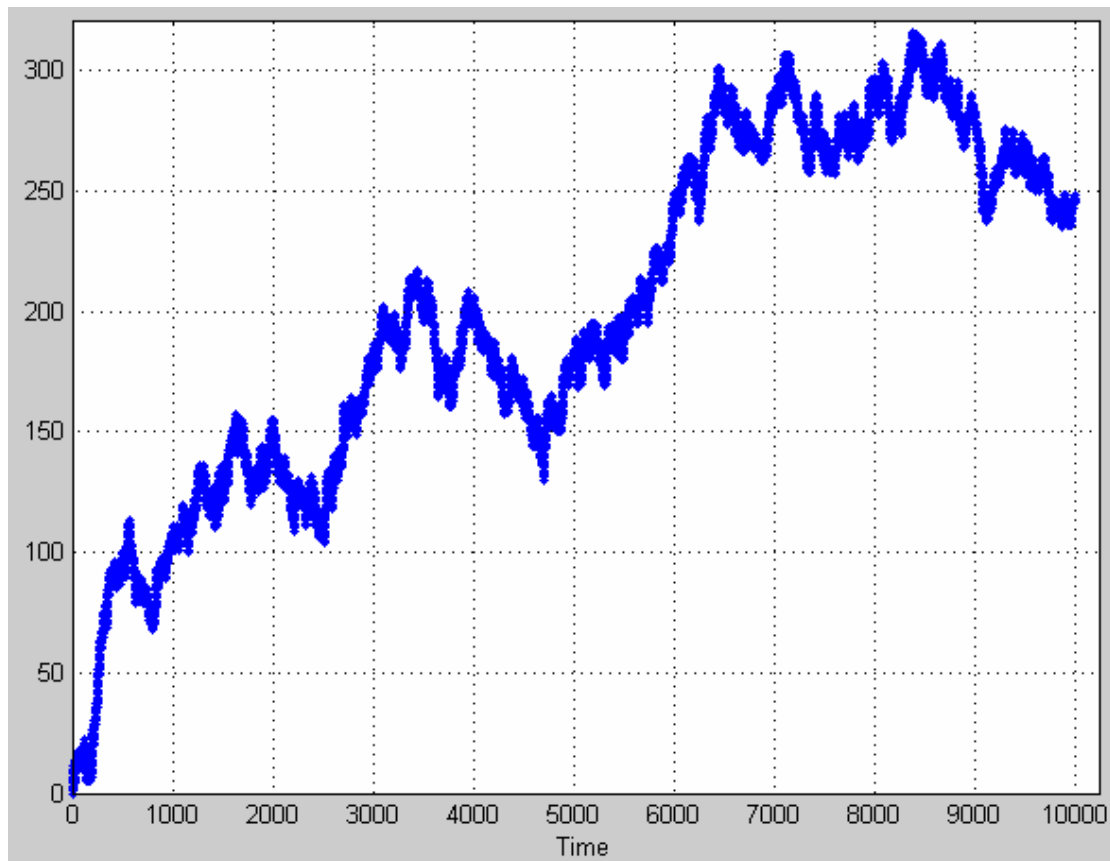


缓冲区 B4(模型中的 Queue4)中在制品数量的变化情况



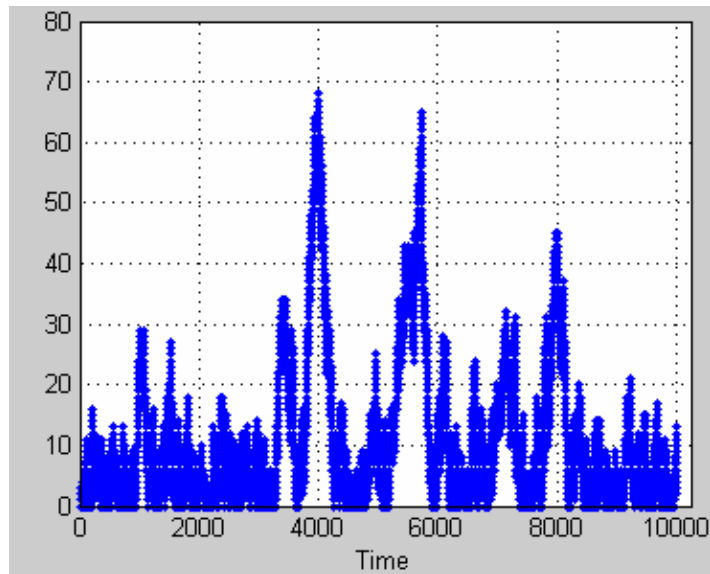
缓冲区 B5(模型中的 Queue5)中在制品数量的变化情况

从上面5张图可以发现，只有缓冲区B1中的在制品数量呈不断上升趋势，而其他缓冲区中在制品数量没有呈现这一趋势。似乎整个生产线的的不稳定是由于第一个缓冲区中在制品数量不断上升引起的。是这样的吗？现在把机器Single Server1加工一个零件的时间均值缩短到0.9，再进行仿真。得到在制品总量的变化情况如下：

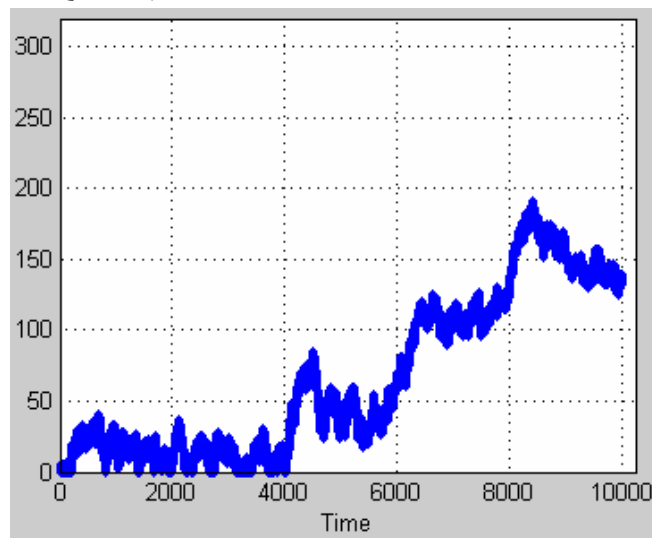


每个缓冲区内制品数量变化如下:

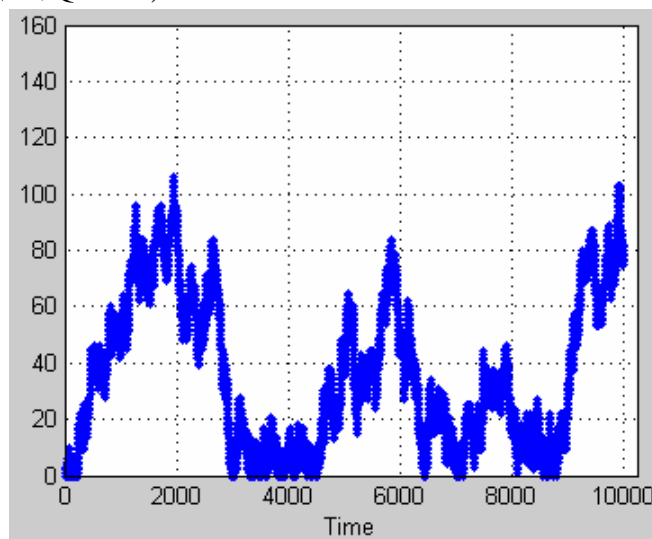
缓冲区B1(模型中的Queue1):



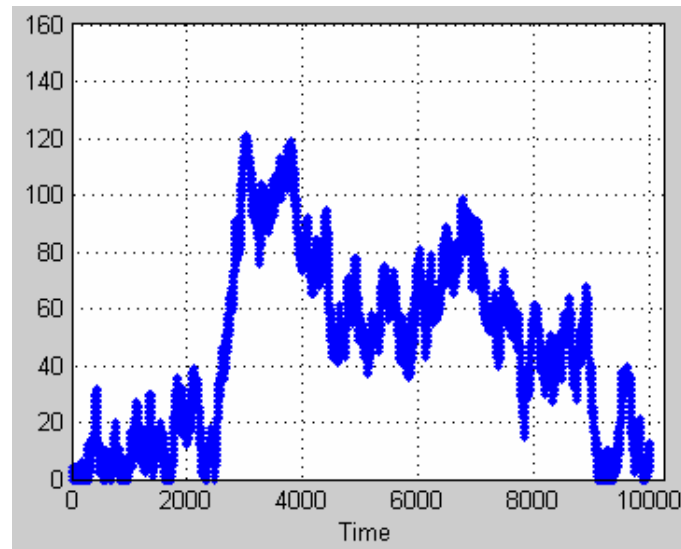
缓冲区B2(模型中的Queue2):



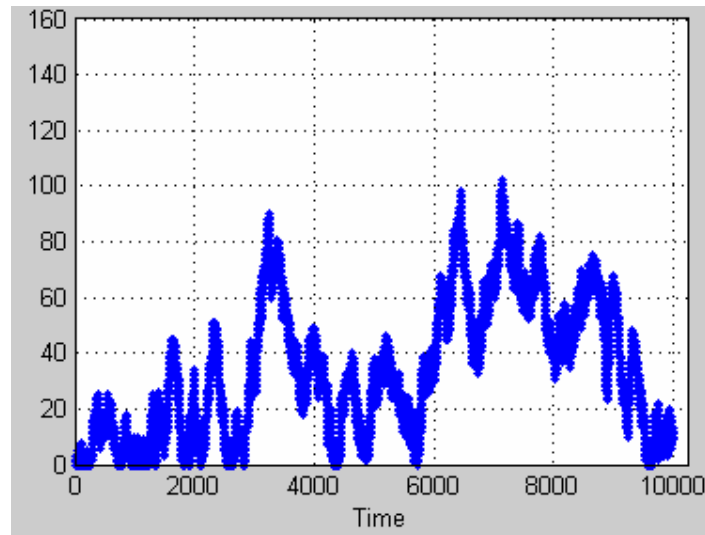
缓冲区B3(模型中的Queue3):



缓冲区 B4(模型中的 Queue4):



缓冲区 B5(模型中的 Queue5):

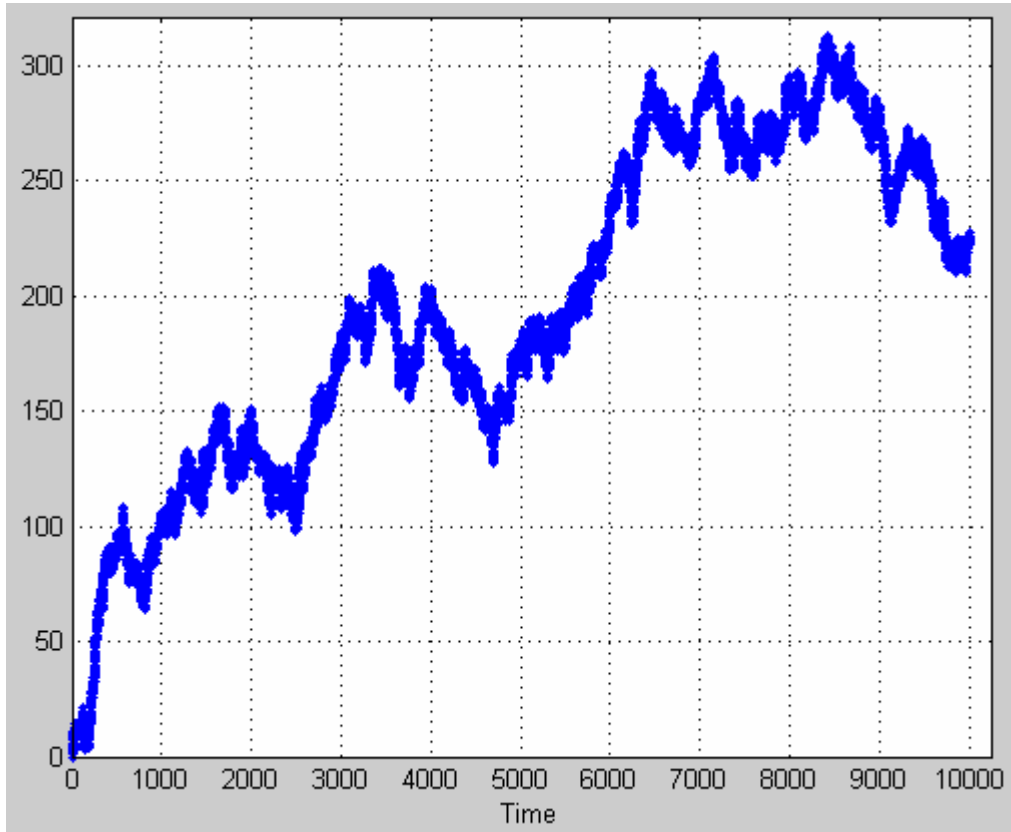


由以上各图可见,总的在制品数量仍然呈上升趋势. 在 5 个缓冲区中,缓冲区 2 中的在制品数量呈上升趋势,它对在制品总量的贡献最大;而其他缓冲区中在制品数量不呈上升趋势.

继续改变参数,把机器 **Single Server2** 加工 1 个零件的时间缩短为 0.8, 进行仿真. 可以猜测: 是否缓冲区 3 中的在制品数量呈上升趋势,而其他缓冲区中在制品数量稳定,整个系统不稳定. 仿真结果如下:

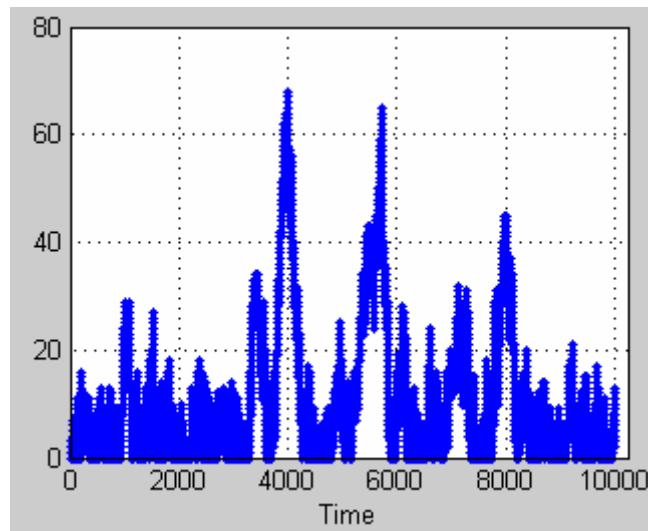
下图为总的在制品数量随时间变化的情况,可见系统仍然是不稳定的.



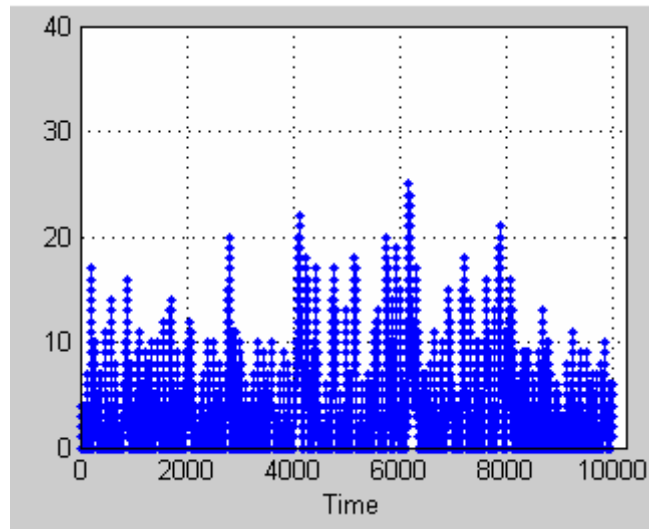


各缓冲区内在制品数量随时间变化的情况如下:

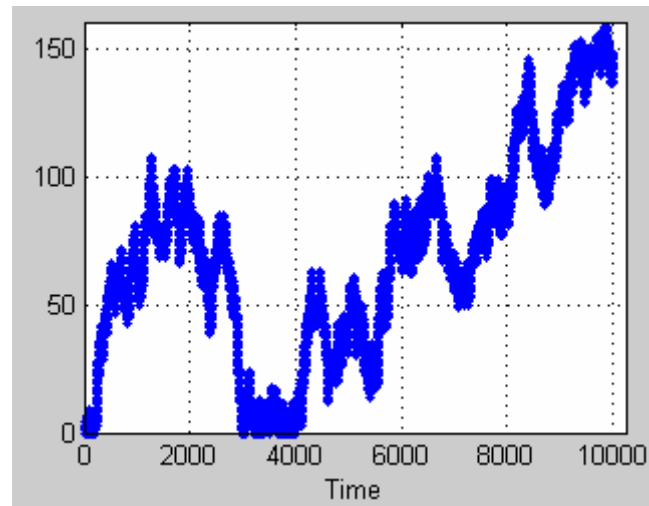
缓冲区 B1:



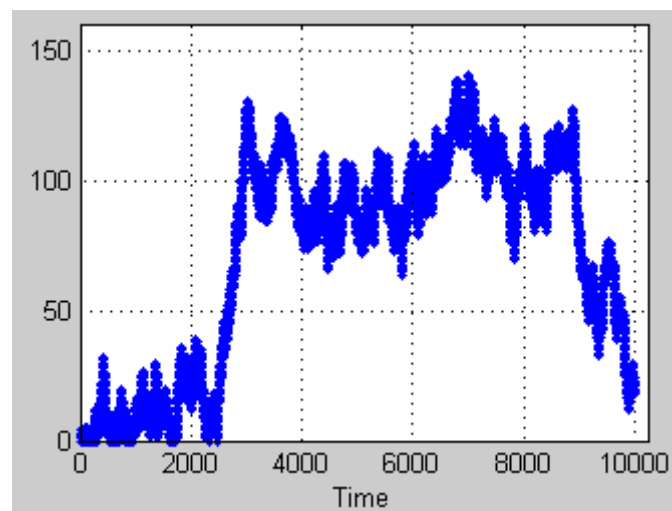
缓冲区 B2:



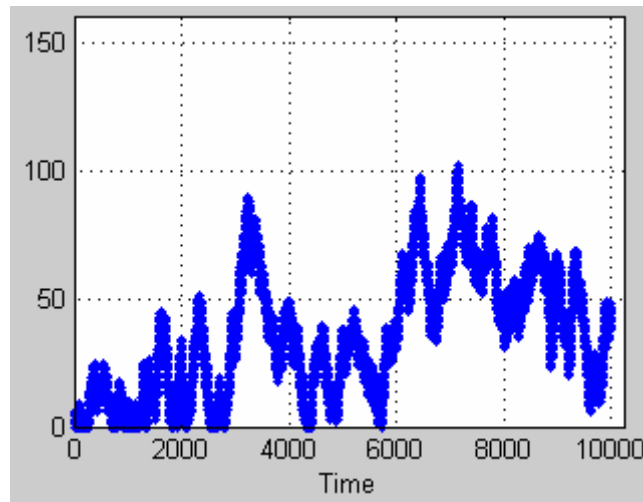
缓冲区 B3:



缓冲区 B4:

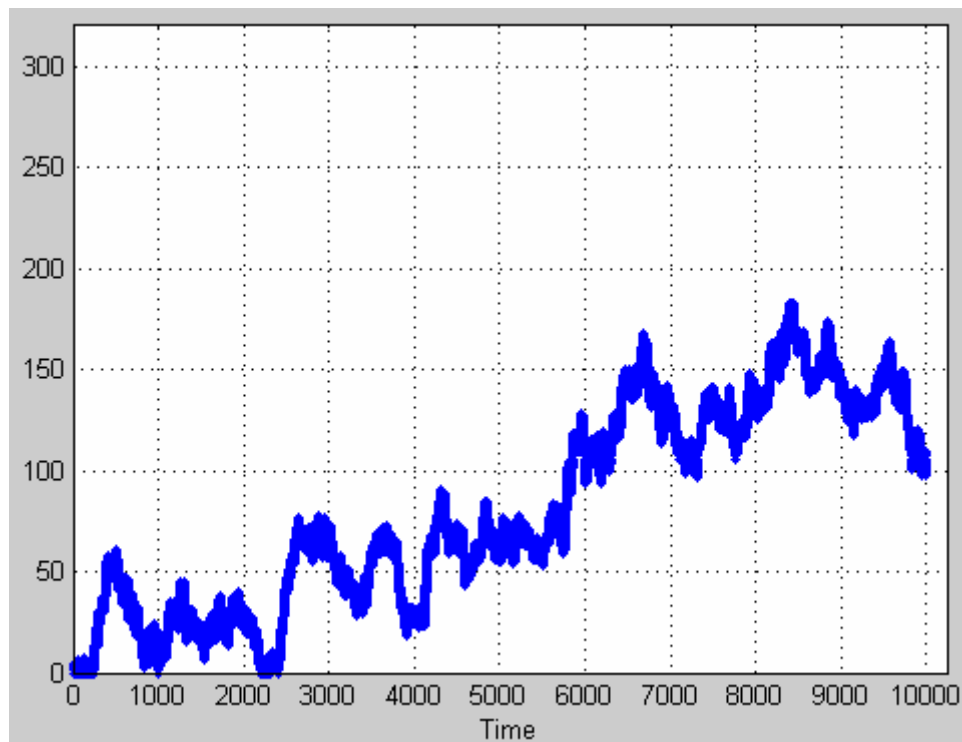


缓冲区 B5:

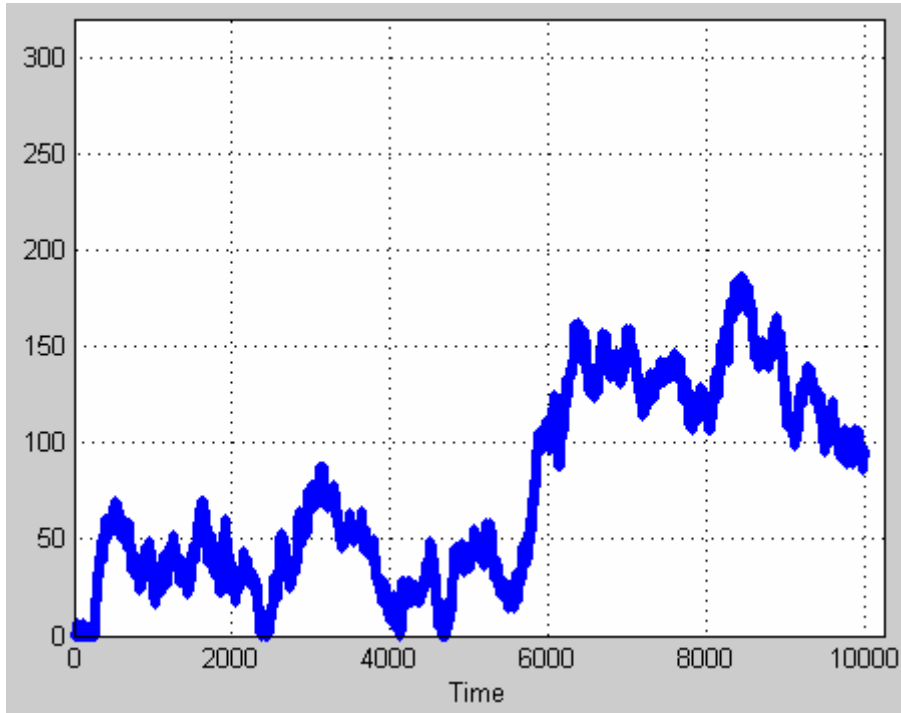


由上图可见, 结果和我们猜测的基本相符. 缓冲区 B3 中的在制品数量呈上升趋势, 对在制品总量贡献最大.

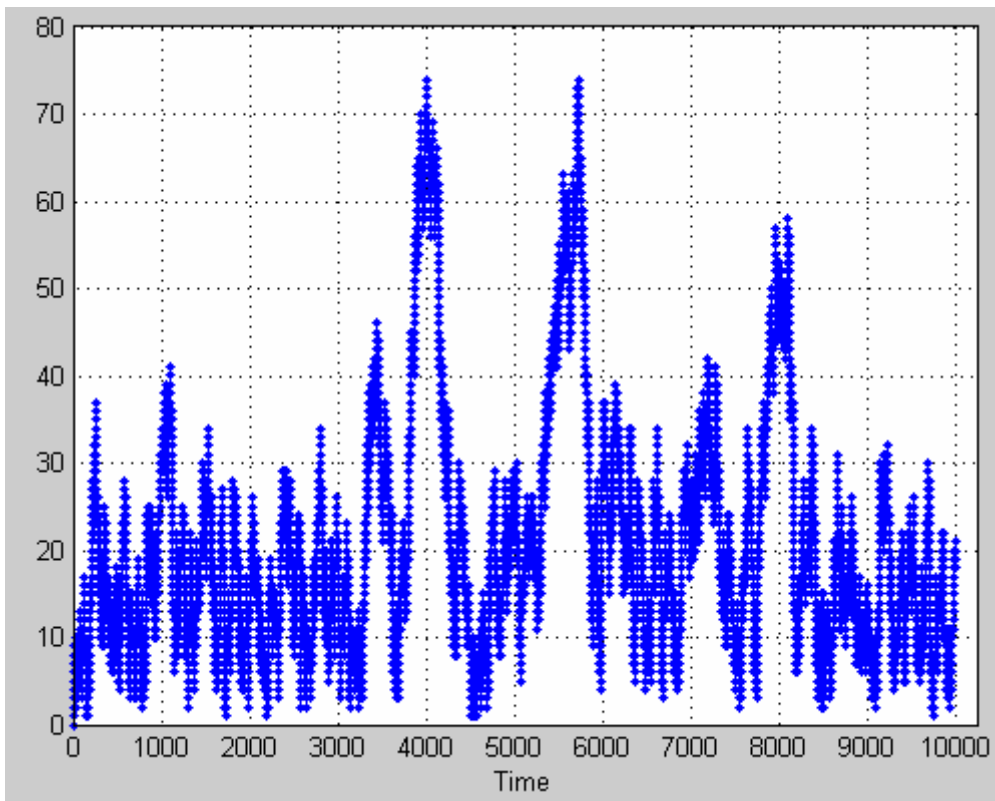
继续依次改变机器加工零件的时间: 把机器 Single Server3 加工 1 个零件的时间缩短为 0.7; 再把机器 Single Server4 加工 1 个零件的时间缩短为 0.6; 最后把机器 Single Server5 加工 1 个零件的时间缩短为 0.5, 分别进行仿真. 由仿真结果可知 (见下图), 对于前两种情况, 系统不稳定, 分别是缓冲区 B4 和 B5 中的在制品数量呈上升趋势. 而对于最后一种情况, 系统稳定.



机器 Single Server3 加工 1 个零件的时间缩短为 0.7 时  
缓冲区 B4 中在制品数量随时间变化的情况



机器 Single Server4 加工 1 个零件的时间缩短为 0.6 时  
缓冲区 B5 中在制品数量随时间变化的情况

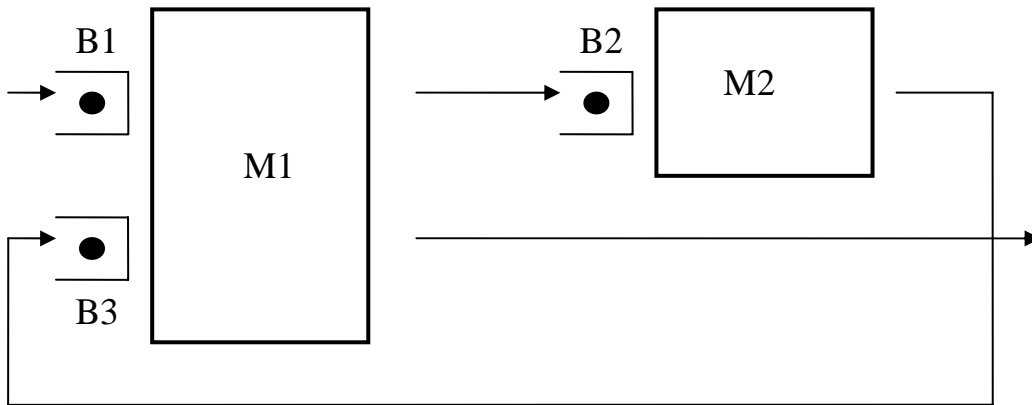


6 台机器加工 1 个零件的时间的均值分别是 1.0, 0.9, 0.8, 0.7, 0.6, 0.5 时  
生产线上在制品总量随时间变化的情况

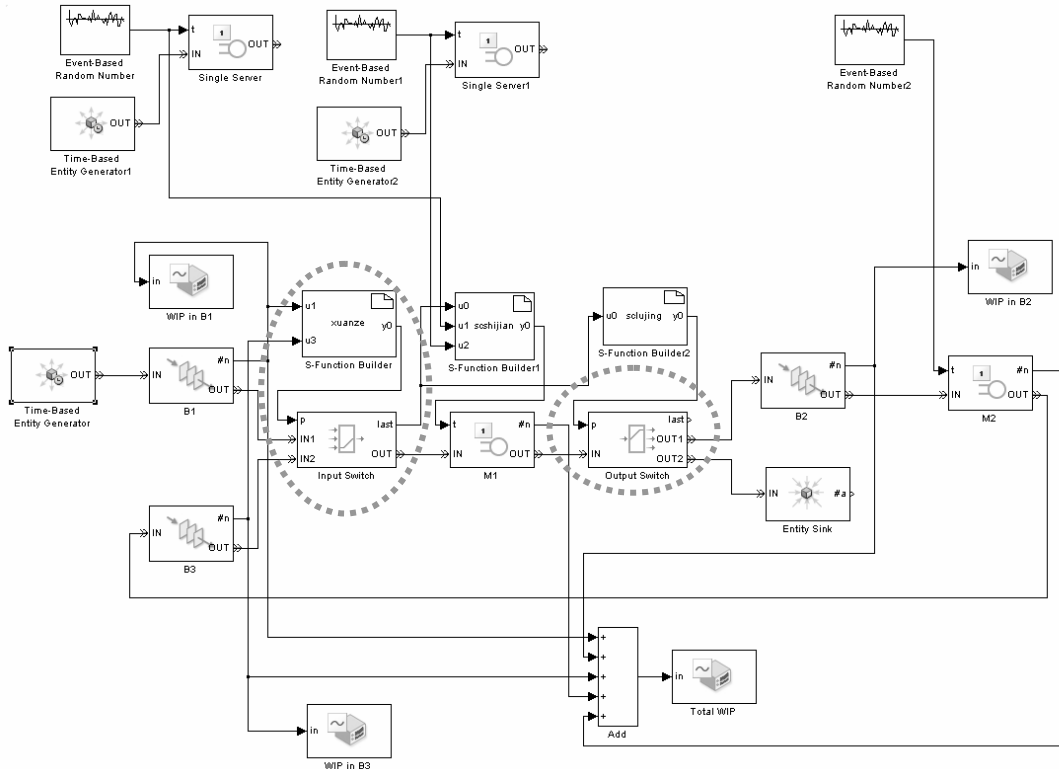
由此可得出这样的结论：对于随机的串行生产线，要使系统稳定，每个机器加工一个零件的时间的均值应依次减小。

### §4.3 基于仿真的可重入制造系统的分析

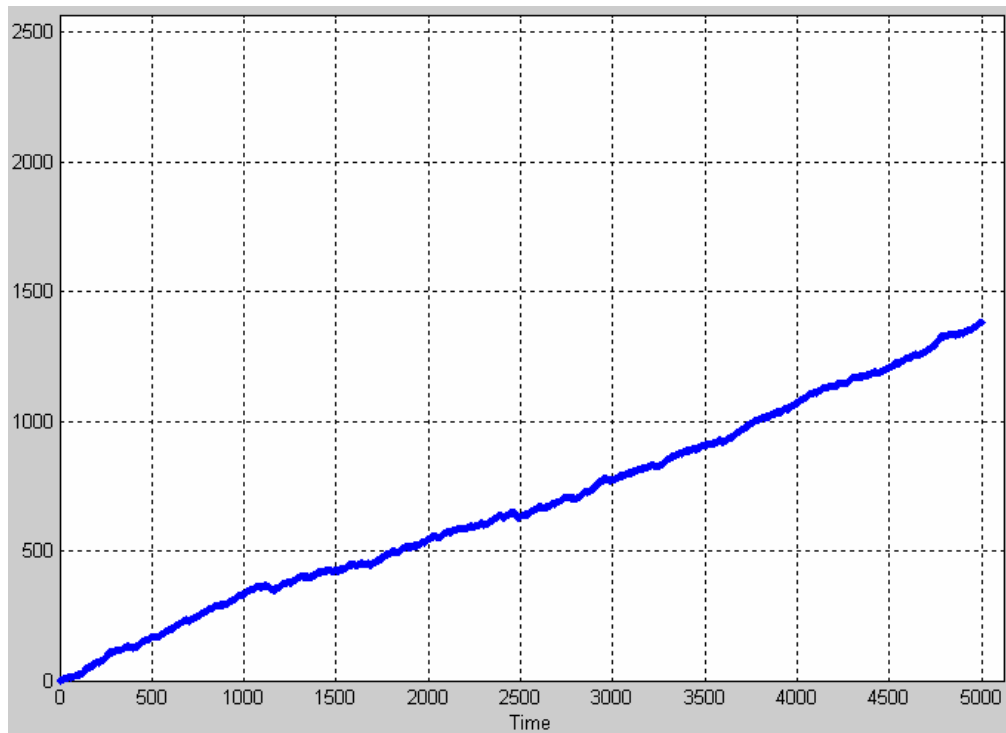
设有一两机器的可重入制造系统，如下图所示。



机器 M1 加工缓冲区 B1 中和 B3 中零件所需时间服从均值为 1 的指数分布；机器 M2 加工缓冲区 B2 中零件所需时间服从均值为 0.5 的指数分布. M1 按照 B1 优先于 B3 的规则来调度零件. 在 SimEvents 环境下建立仿真模型如下：

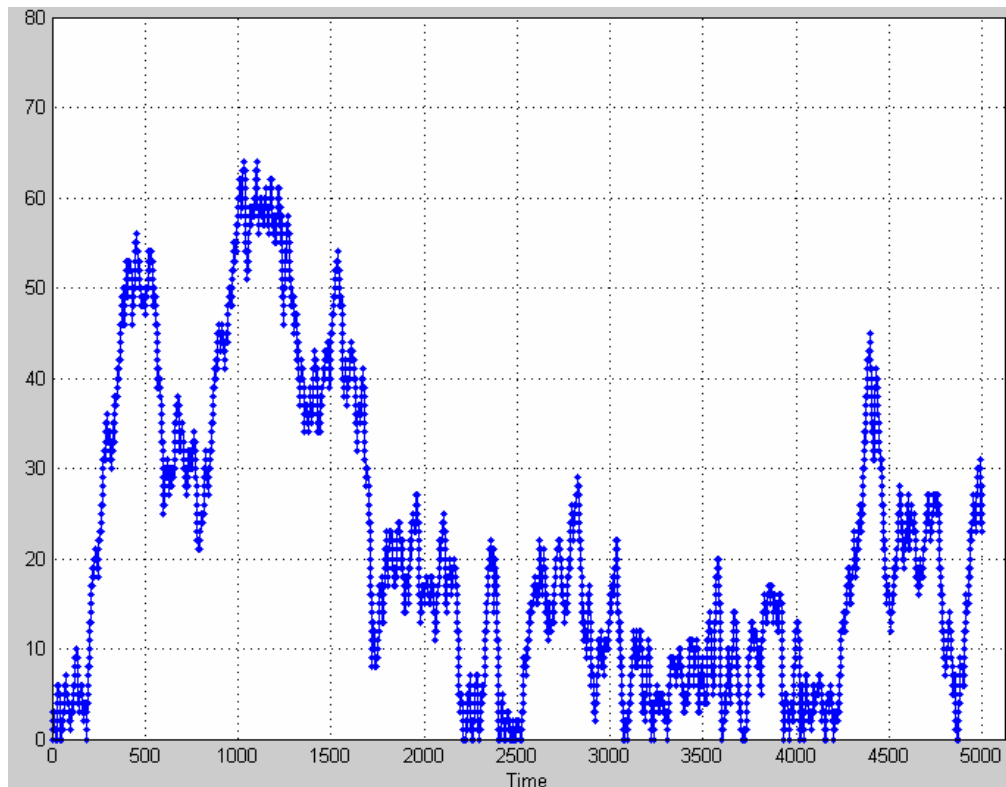


模型中，调度规则由 Input Switch, Output Switch 元件和 S 函数共同实现(见上图中用虚线椭圆标出的部分)。对该系统进行仿真，当零件的到达时间间隔服从参数为 2 的指数分布时，系统内部总的在制品数量随时间变化的情况如下：

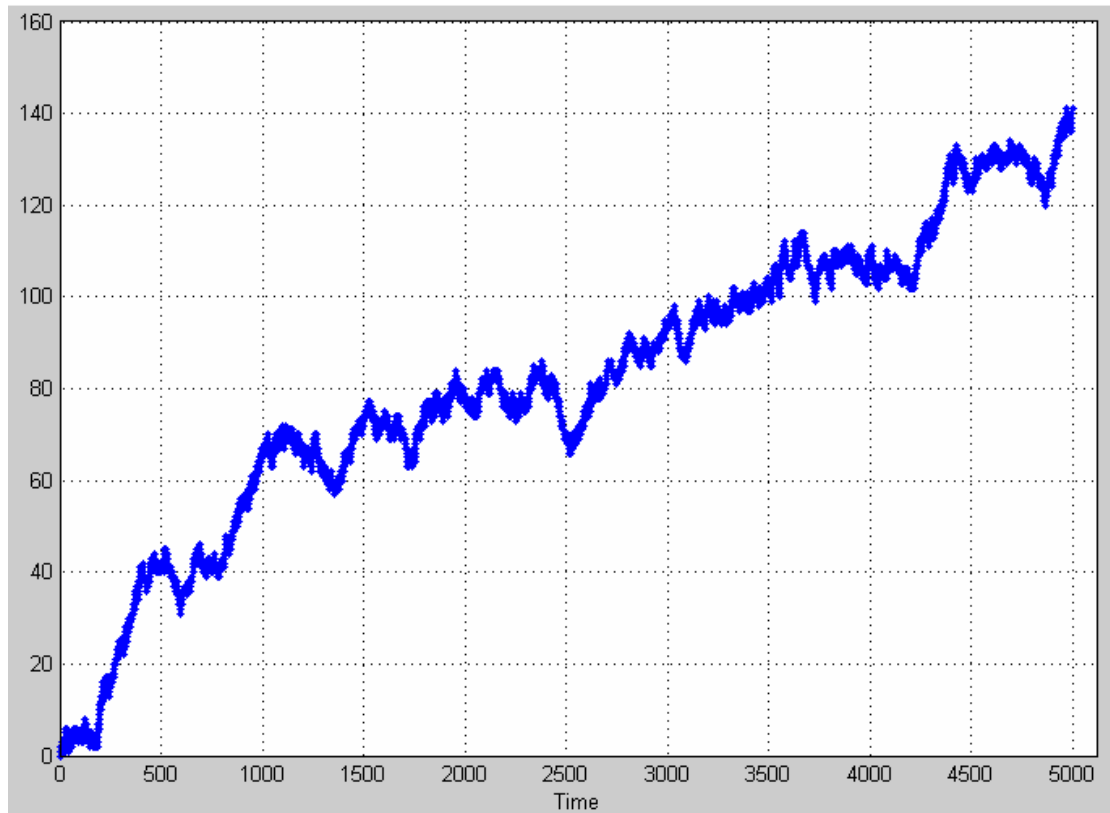


可见系统是不稳定的。

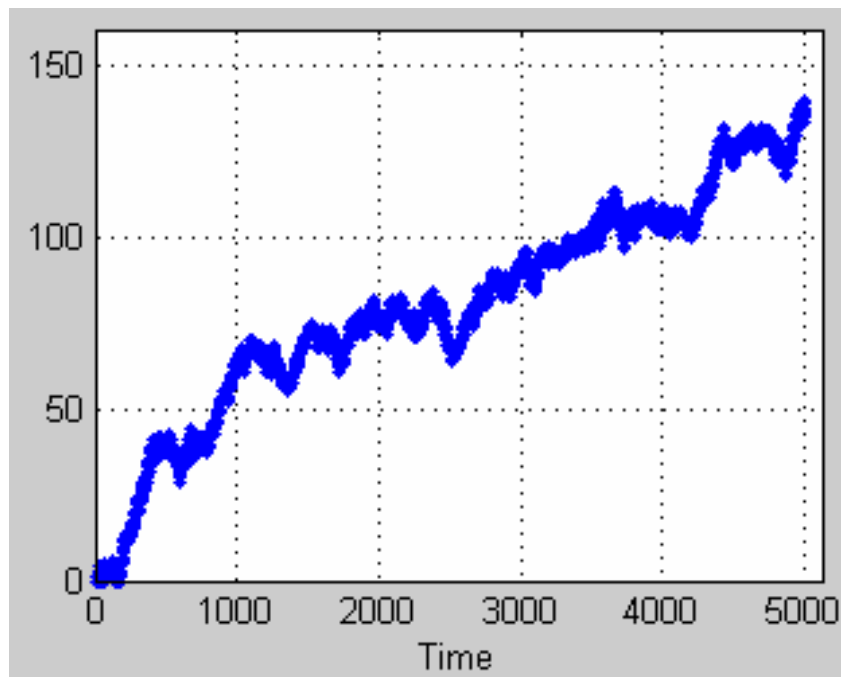
增大零件到达时间间隔：当零件的到达时间间隔服从参数为 3.0 的指数分布时，系统内部总的在制品数量随时间变化的情况如下：



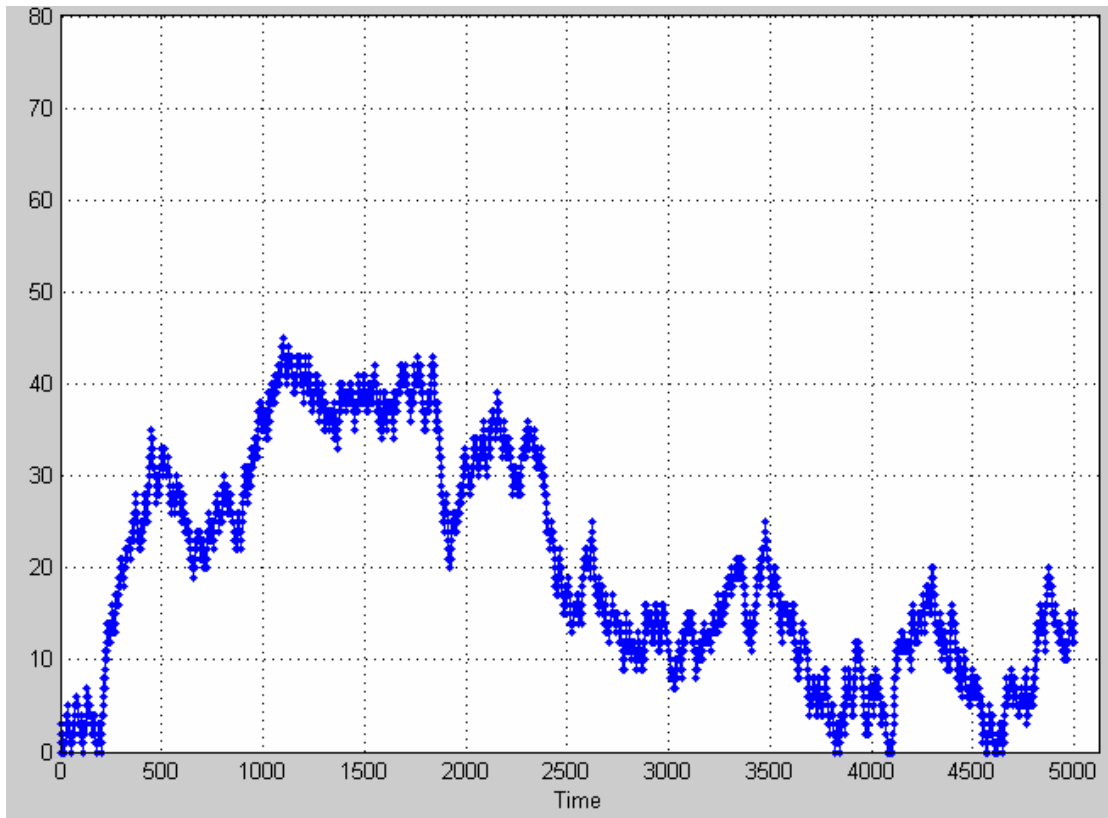
系统稳定. 如果把调度规则改为 B3 优先于 B1, 则系统又变得不稳定, 如下图所示.



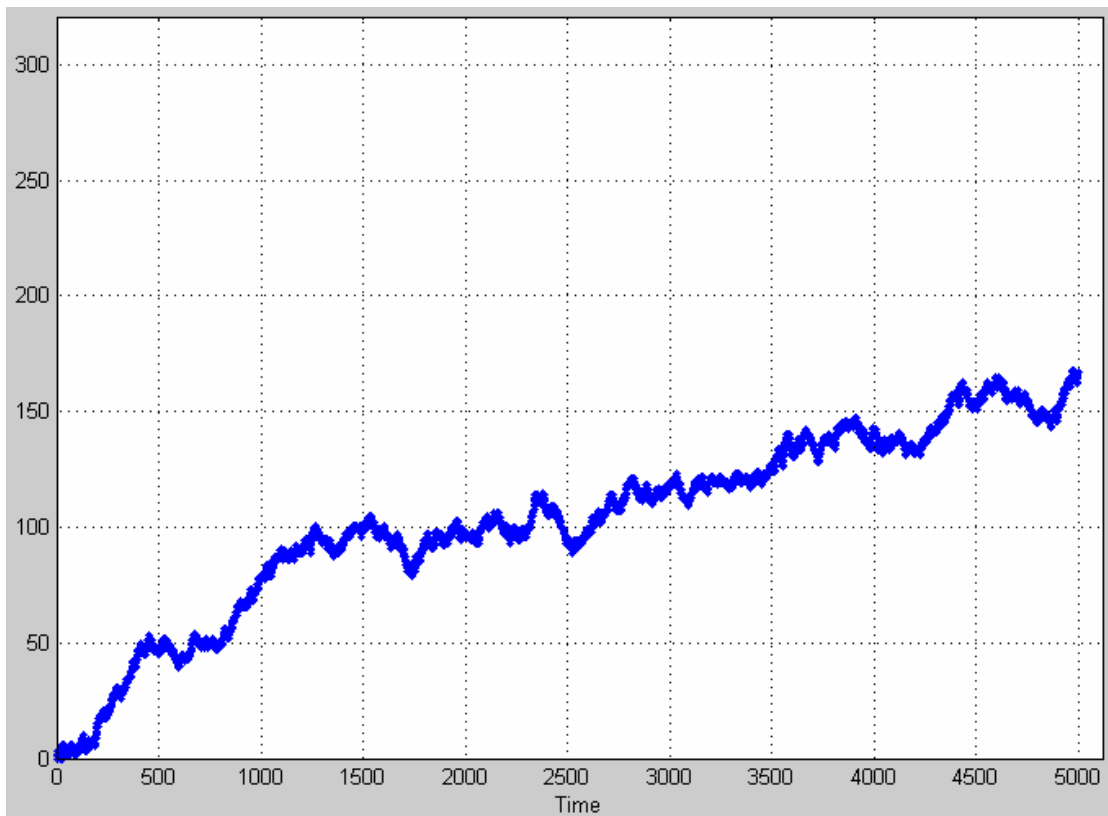
三个缓冲区中, B1 中的在制品数量有不断上升的趋势, 如下图所示.



把零件到达时间间隔的均值增加到 3.3, 则系统又恢复稳定, 如下图所示:

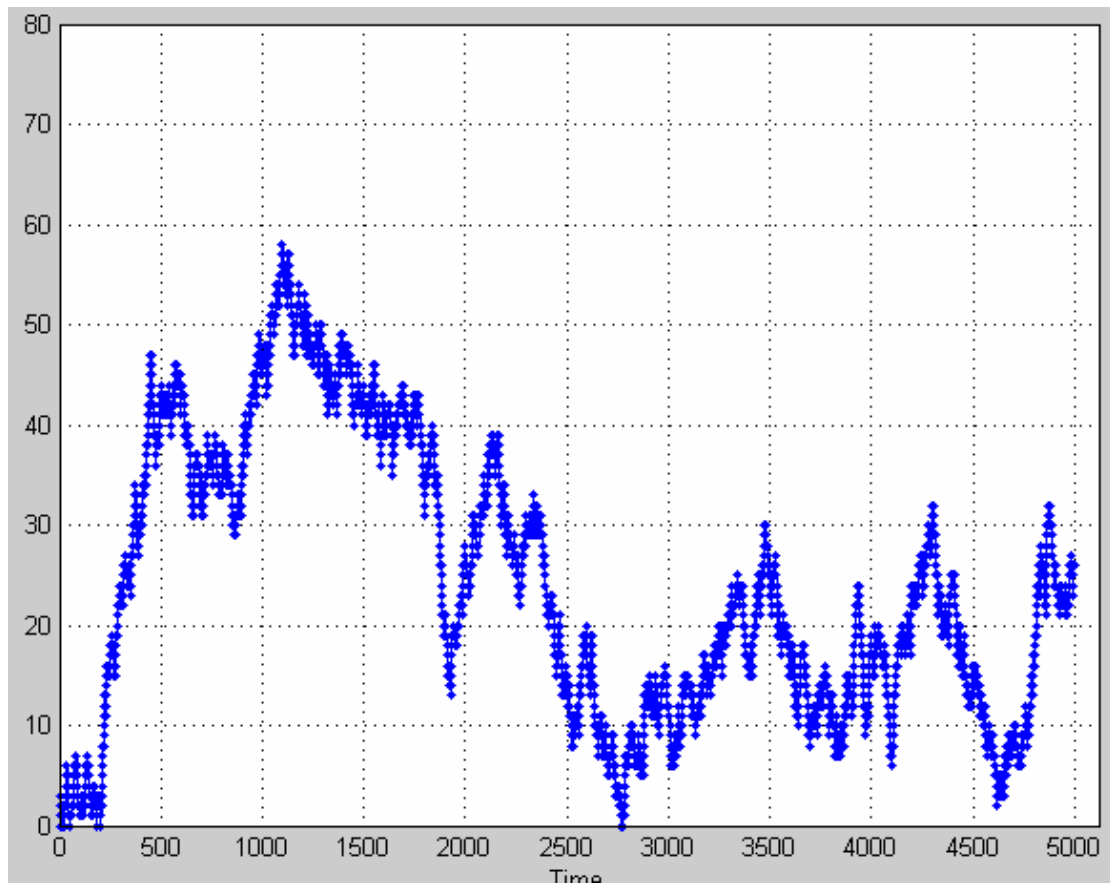


如果把零件到达时间间隔的均值恢复到 3.0, 再把调度规则改为优先生产 B1 和 B3 中在制品较多者, 则系统中在制品总量随时间变化的情况如下图所示.





系统不稳定. 再把零件到达时间间隔的均值增加到 3.3, 则系统又恢复稳定



由上述仿真分析可见: 对于可重入制造系统, 零件释放进系统的规则与每台机器上零件的调度规则对于系统的稳定性具有重要影响.

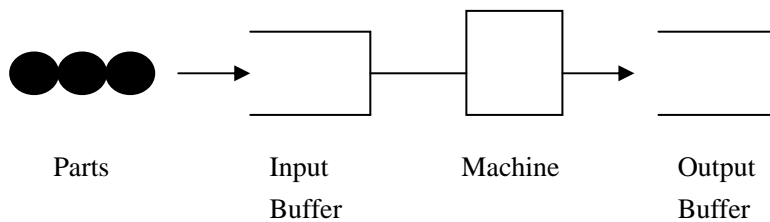
# 第 5 章

## 基于 Petri 网的制造系统分析与控制(1)

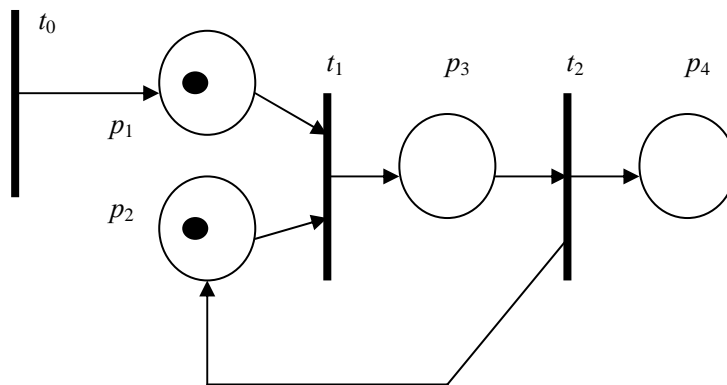
Petri 网(Petri Nets)是在描述和分析离散事件动态系统状态变化的逻辑关系的数学工具,它具有很强的建模能力,已经被广泛地应用于制造系统、计算机网络系统、通讯系统等离散事件动态系统的建模、仿真、分析与控制中.本章和下两章将在介绍 Petri 网的基本概念的基础上,讨论用 Petri 网对制造系统进行建模、分析与控制的方法.

### §5.1 Petri 网的基本概念

首先给出一个用 Petri 网为制造系统建模的简单的例子.设有一单机器制造系统,如图所示:



可建立 Petri 网模型如下:



其中, 库所  $p_1$  表示零件处于等待加工的状态; 库所  $p_2$  表示机器处于空闲状态; 变迁  $t_0$  表示产生一待加工的零件; 变迁  $t_1$  表示零件加工开始这一事件; 库所  $p_3$  表示机器处于正在加工零件的状态; 变迁  $t_2$  表示零件完成加工这一事件; 库所  $p_4$  表示零件处于加工已经完成的状态. 令牌(token, 用黑点表示)表示系统的当前状态.

Petri 网的定义:

一个 Petri 网可表示成一个六元组  $PN = (P, T, F, W, M, M_0)$ , 其中

$P = \{p_1, \dots, p_m\}$  为有限库所(place)集;  
 $T = \{t_1, \dots, t_n\}$  为有限变迁(transition)集; 并且  $P \cap T = \emptyset$ ,  $P \cup T \neq \emptyset$ ;  
 $F \subseteq (P \times T) \cup (T \times P)$  为有向弧集, 表示库所与变迁之间的流关系;  
 $W: F \rightarrow \{1, 2, \dots\}$  为有向弧上的权函数;  
 $M: P \rightarrow \{0, 1, 2, \dots\}$  为状态标识(marking), 用“令牌(token)”表示;  
 $M_0: P \rightarrow \{0, 1, 2, \dots\}$  为初始状态标识(initial marking).

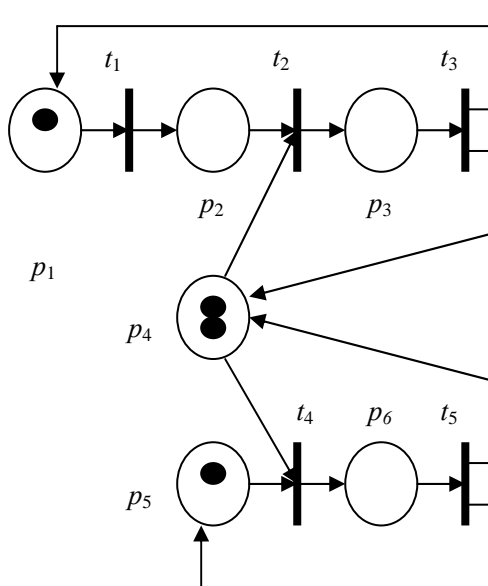
变迁的输入集合:  $I(t) = \{p | (p, t) \in F\}$ ; 变迁的输出集合:  $O(t) = \{p | (t, p) \in F\}$ ;  
 库所的输入集合:  $I(p) = \{t | (t, p) \in F\}$ ; 库所的输出集合:  $O(p) = \{t | (p, t) \in F\}$ ;  
 系统状态的表示:  $M = [M(p_1), \dots, M(p_n)]$ ;  
 系统的初始状态:  $M_0 = [M_0(p_1), \dots, M_0(p_n)]$ .

变迁的使能(enable)和触发(fire): 如果一个变迁满足触发的条件, 则称它被“使能(to be enabled)”, 如果被使能的变迁代表的事件发生了, 则称它被“触发(to be fired)”. 变迁被使能的条件:

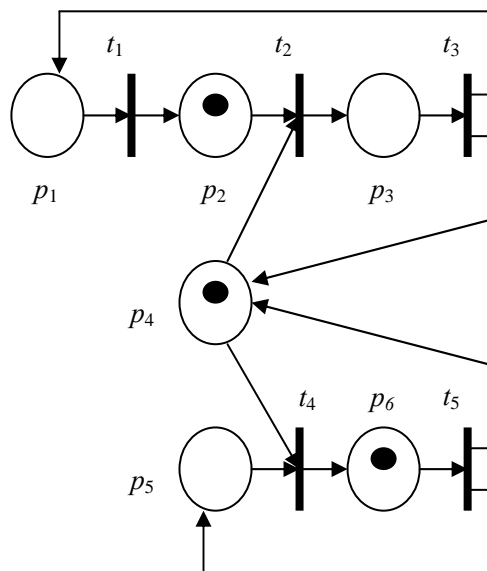
- (1)  $\forall p_i \in I(t)$ ,  $M(p_i) \geq W(p_i, t)$  必须成立;
- (2)  $\forall p_j \in O(t)$ ,  $K(p_j) \geq M(p_j) + W(t, p_j)$  必须成立;
- (3)  $\forall p_r \in I(t) \cap O(t)$ ,  $M(p_r) \geq W(p_r, t)$  和  $K(p_r) \geq M(p_r) - W(p_r, t) + W(t, p_r)$  必须同时成立.

变迁触发后, Petri 网中标识的状态是:

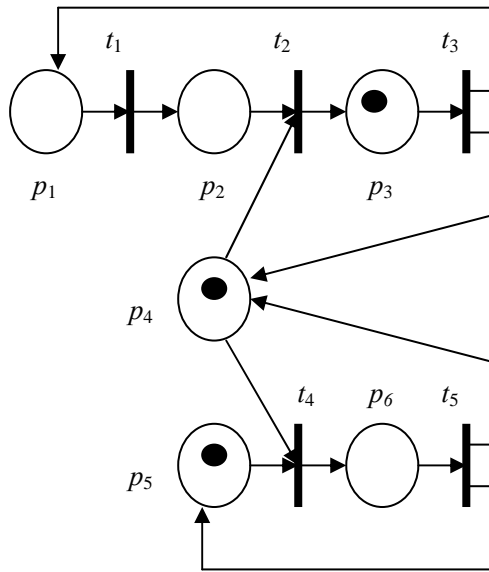
$$M'(p) = \begin{cases} M(p) - W(p, t), & \forall p \in I(t) \text{ and } p \notin O(t); \\ M(p) + W(t, p), & \forall p \in O(t) \text{ and } p \notin I(t); \\ M(p) - W(p, t) + W(t, p), & \forall p \in O(t) \cap I(t); \\ M(p), & \text{otherwise} \end{cases}$$



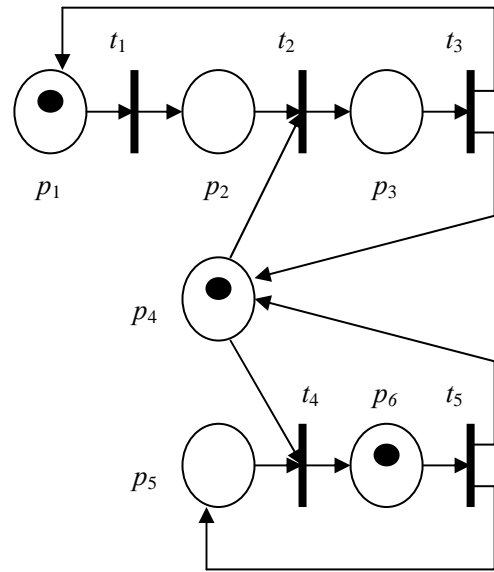
$$M_0 = [1, 0, 0, 2, 1, 0]$$



$$M_1 = [0, 1, 0, 1, 0, 1]$$



$$M_2 = [0, 0, 1, 1, 1, 0]$$



$$M_3 = [1, 0, 0, 1, 0, 1]$$

**Petri 网的基本性质:**

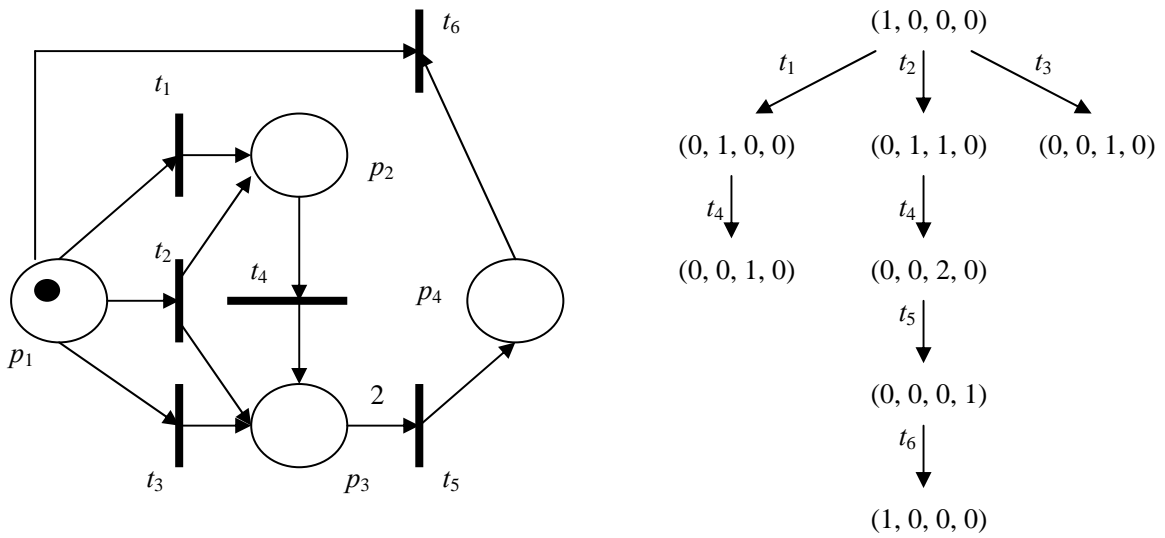
1. 可达性(reachability): 对于给定初始标识的一个 Petri 网  $(N, M_0)$ , 其可达集  $R(N, M_0)$  定义为此 Petri 网在初始标识  $M_0$  下, 按照变迁触发规则, 经过有限次变迁触发可以达到的所有可能状态标识的集合.
  2. 有界性(boundness): 对于给定初始标识的一个 Petri 网  $(N, M_0)$ , 若存在一正整数  $K$ , 使得对于任一可达状态标识  $M \in R(N, M_0)$  及任一库所  $p_i$ , 都有  $M(p_i) \leq K$  成立, 则称该 Petri 网是  $K$ -有界的.
  3. 安全性(safeness): 即 1-有界性.
- 注意: 有界性和安全性实际上刻画了缓冲区的容量有限时制造系统是否会发生阻塞的特性.
4. 活性(liveness): 对于给定初始标识的一个 Petri 网  $(N, M_0)$ , 对于某个变迁  $t$ , 如果对于初始标识  $M_0$  可达的任一状态标识  $M \in R(N, M_0)$ , 都可找到一触发序列, 使得由此到达的新的状态标识  $M'$  可使命变迁  $t$ , 则称此变迁为活的. 若给定初始标识的一个 Petri 网  $(N, M_0)$  中所有变迁都是活的, 则称此 Petri 网是活的.
  5. 死锁(deadlock): 对于给定初始标识的一个 Petri 网  $(N, M_0)$ , 对于某个变迁  $t$ , 如果初始标识  $M_0$  可达的任一状态标识  $M \in R(N, M_0)$  都不能使命该变迁, 则称该变迁为死锁.
  6. 冲突(conflict): 对于给定初始标识的一个 Petri 网  $(N, M_0)$ , 若其中有多于一个已被使能的变迁, 但由于共享某些输入库所, 其中一个变迁的触发会导致其他一个或多个变迁的不能触发, 则称这些变迁冲突.

可达性分析 --- 可达树分析法.

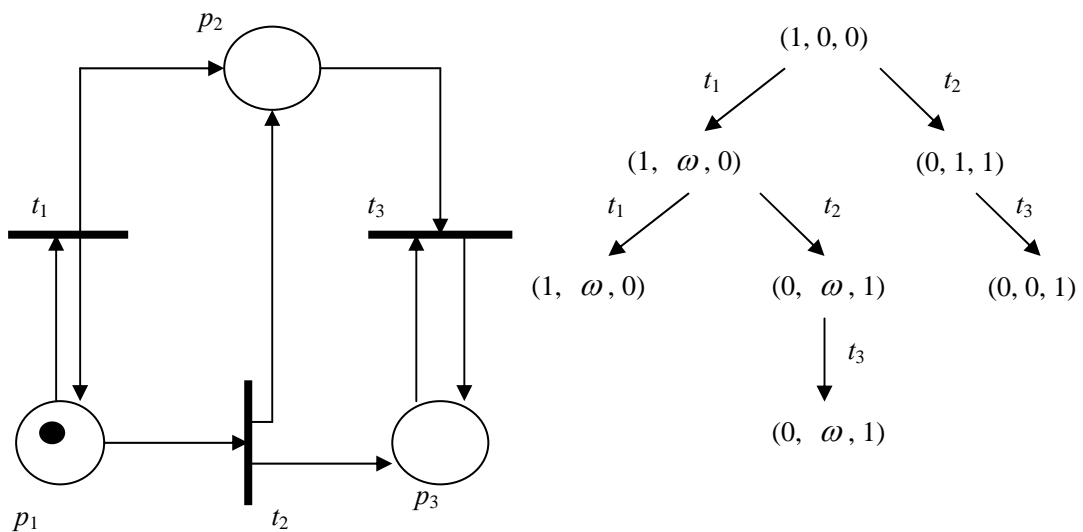
有限可达集对应有限可达树. 对于无限可达集, 可通过引入无限参量  $\omega$ , 用有限树的形式表达, 为此需引入一些规则:

1. 若可达树中某个节点以前层出现过, 则不再往下延伸, 若延伸则会出现重复的结构.

2. 若可达树中某个节点  $M$  比由根节点  $M_0$  到该节点的一条路径上某一中间节点  $M'$  大, 即对于每一库所  $p_i$ , 都有  $M(p_i) \geq M'(p_i)$ , 则对其中“ $>$ ”成立的库所, 以  $\omega$  代替其令牌数.



有限可达树



无限可达树

- 有界性判别: 一个 Petri 网是有界的, 当且仅当其可达树中不出现无限参量  $\omega$ .  
 安全性判别: 一个 Petri 网是安全的, 当且仅当其可达树中的每个节点中只包含 0 和 1 元素.  
 死锁判别: 一个 Petri 网的某个变迁是死锁的, 当且仅当此变迁不出现在可达树任一支上.

扩展 Petri 网:

1. 赋时 Petri 网(Timed Petri Nets)

$$TPN = (P, T, F, W, M, M_0, \theta)$$

其中,  $\theta = (\theta_1, \dots, \theta_n)$  为一组与库所、变迁或有向弧相关联的时延参量.

## 2. 随机 Petri 网(Stochastic Petri Nets)

随机 Petri 网是指时延参量为随机变量的一类赋时 Petri 网, 其定义为

$$SPN = (P, T, F, W, M, M_0, L)$$

其中,  $L = (L_1, \dots, L_n)$  为一组与变迁相关联的触发率参量.

## 3. 有色 Petri 网(Colored Petri Nets)

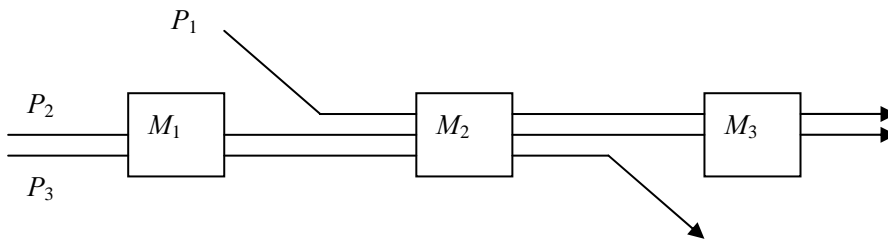
对不同属性的令牌赋予不同的颜色, 同一颜色的有色令牌表示一类有相同行为特性的元素. 在联系库所和变迁的有向弧上, 定义了线性函数, 它确定了变迁的触发条件, 包括不同颜色的令牌的移动.

## 4. 高级 Petri 网(High-level Petri Nets)

包括有色 Petri 网, 谓词/变迁网等. 其共同特征是通过引入具有共同特性的标志(如颜色), 达到简化网的结构和增强描述能力的目的.

用 Petri 网为制造系统建模:

设有一个三台机床和三种零件构成的 FMS, 且加工顺序(调度)和工艺路线都是固定的. 工艺路线如图所示, 每台机器都按照  $P_1, P_2, P_3$  的顺序加工零件.



其 Petri 网模型如下页图所示。

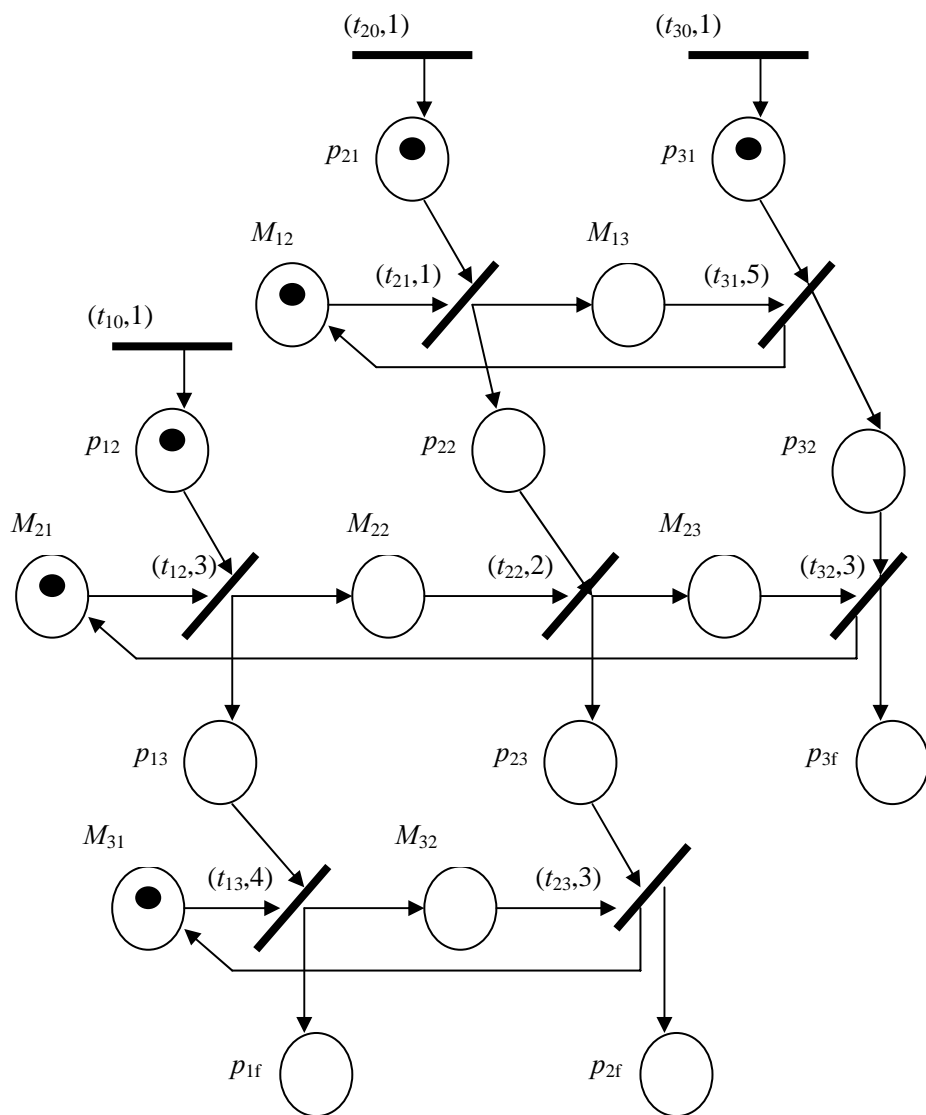
## §5.2 基于 Petri 网的制造系统建模

1. 例: 某装载机总装生产线的 Petri 网模型.

2. 阅读以下文献:

[1] Zhou, M. C. and Jeng M. D., 1998. Modeling, analysis, simulation, scheduling, and control of semiconductor manufacturing systems: a Petri net approach. *IEEE Transaction on Semiconductor Manufacturing*, Vol.11, No.3, pp.333-357

[2] Chaouiya, C. and Dallery, Y. Petri net models of pull control systems for assembly manufacturing systems



附：某装载机总装线的Petri网模型。

库所：

- $p'_i$ ：工件第  $i$  道工序的装配已完成；其中的令牌为  $o_i$ ，表示工件。
- $p_W$ ：空闲工人库所；其中的令牌为  $w_n$ ， $n=1, \dots, N$ ，表示工人， $N$  为工人数量。
- $p_M$ ：空闲设备库所；其中的令牌为  $m_k$ ， $k=1, \dots, K$ ，表示设备， $K$  为设备数量。
- $p_S$ ：可用的空间位置库所；其中的令牌为  $s_{jh}$ ， $j=1, \dots, J$ ， $h=1, \dots, H_j$ ， $s_{jh}$  表示第  $j$  工位的第  $h$  个空间位置。 $J$  为工位数量， $H_j$  为第  $j$  个工位上的空间位置数量。
- $p_i$ ：工件正在进行第  $i$  道工序的装配， $i=1, \dots, 70$ ， $i=1$  时（即  $p_1$ ）表示等待上线的工件源库所， $i$  从 2 到 70 相对应于装配线上 69 道工序。其中的令牌为  $\langle o_i, w_n, m_k, s_{jh} \rangle$ ，表示工人  $w_n$  利用设备  $m_k$  在位置  $s_{jh}$  装配工件  $o_i$ 。工人可有多个，例如  $\langle o_i, w_n, w_q, m_k, s_{jh} \rangle$ ，表示工人  $w_n$  和  $w_q$  利用设备  $m_k$  在位置  $s_{jh}$  装配工件  $o_i$ 。
- $p_i^*$ ：工件进行并完成第  $i$  道工序的装配，它是为简化整个装配线 Petri 网图复杂程度而引入的，是  $p_i$  和  $p'_i$  这两个库所和变迁  $t'_i$  的组合。 $i$  值所表示的意义及相应的令牌的定义见  $p_i$ 。
- $p_{A1} - p_{A16}$ ：16 个装配工位；中的令牌为  $o_i$ 。

$p_{\text{walk}} / p_{\text{pause}}$ ：传送带正在运行/暂停运行；其中的令牌为  $c$ 。

令牌：

- $o_i$ ：表示工件。对于  $p_i$ ，获得  $o_i$  表示工件到达  $p_i$  规定装配的工位处；
- $w_n$ ：表示工人。对于  $p_i$ ，获得  $w_n$  表示装配  $p_i$  所需的工人  $n$  有空；
- $m_k$ ：表示设备。对于  $p_i$ ，获得  $m_k$  表示装配  $p_i$  所需的设备  $k$  有空；
- $s_{jh}$ ：表示空间位置。对于  $p_i$ ，获得  $s_{jh}$  表示  $p_i$  所需的空间位置可用。
- $p_i$  只有在上述条件都满足的条件下（即获得所有的令牌）方可进行装配。
- $c$ ：表示传送带运行状态转换的令牌，当  $p_{\text{walk}}$  获得该令牌传送带趋于移动状态， $p_{\text{pause}}$  获得该令牌时，传送带趋于停止状态。

变迁：

$t_i$ ：此变迁的触发表示工序  $i$  开始进行。它触发的条件是：其前道工序全部完成（库所  $p'_{i-1}, p'_{i-2}$  中存在相应的令牌），工件处于要求的工位上（库所  $p_{Ah}$  中存在符合要求的令牌），所需工人、设备和空间位置等资源满足要求（库所  $p_W, p_M, p_S$  中存在符合要求的令牌）。它触发的结果是：从其所有的输入库所中取出相应的令牌，构成复合令牌  $\langle o_i, w_n, m_k, s_{jh} \rangle$ ，放入其输出库所  $p_i$ 。

$t'_i$ ：此变迁的触发表示工序  $i$  完成。它触发的条件是：复合令牌  $\langle o_i, w_n, m_k, s_{jh} \rangle$  在其输入库所  $p_i$  中停留足够长时间，直至此道工序完成。它触发的结果是：复合令牌  $\langle o_i, w_n, m_k, s_{jh} \rangle$  分解，令牌  $o_i$  被放入库所  $p'_i$ ，令牌  $w_n, m_k, s_{jh}$  分别返回库所  $p_W, p_M, p_S$ 。

$t_{A1} - t_{A17}$ ：这些变迁的触发表示工件从一个工位运行至下一个工位。

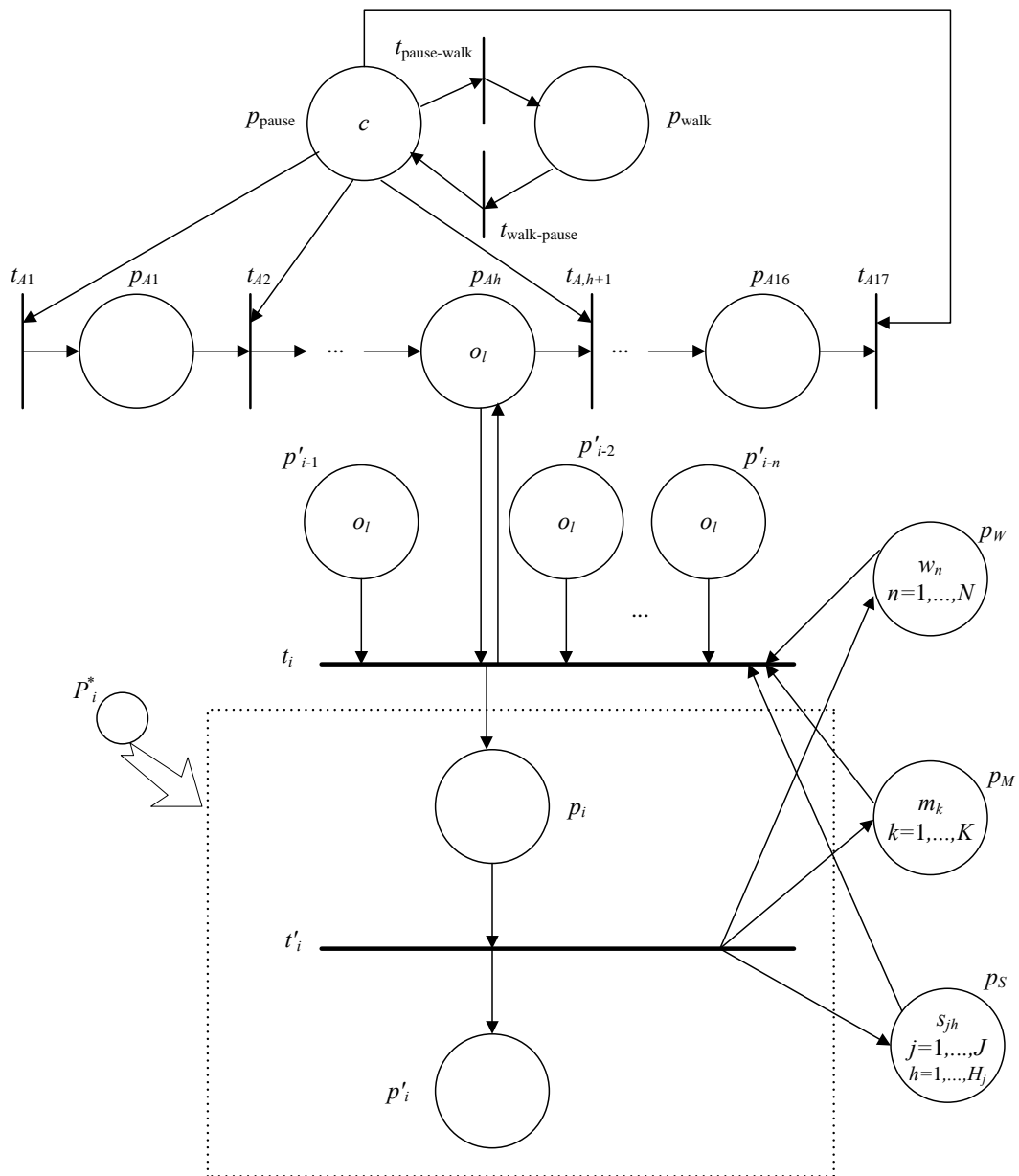
$t_{\text{walk-pause}}$ ：此变迁表示传送带从运行状态改变至暂停状态，其触发由令牌在库所



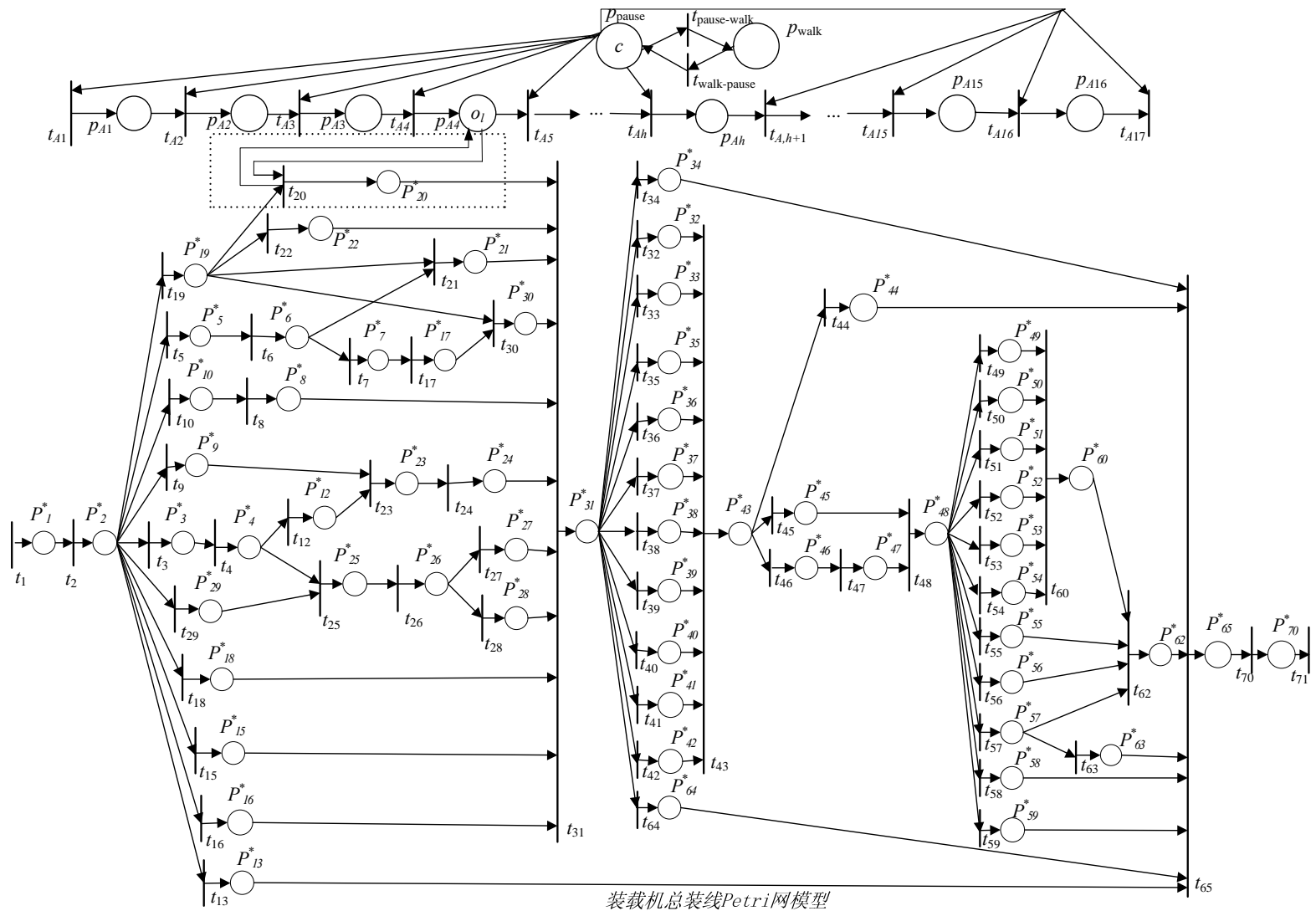
$p_{walk}$  中停留的时间来确定。

$t_{pause-walk}$ ：此变迁表示传送带从暂停状态改变至运行状态，其触发由令牌在库所  $p_{pause}$  中停留的时间来确定。

图中有几个库所在整个装配线的 Petri 网模型中是与每道工序有关的，包括；空闲工人库所  $p_w$ ，空闲设备库所  $p_M$ ，可用的空间位置库所  $p_s$ 。



单道工序状态变化Petri网图



装载机总装线Petri网模型

## 第 6 章

### 基于 Petri 网的制造系统分析与控制(2)

本章分析柔性制造系统中的死锁现象, 并介绍基于 Petri 网的死锁控制方法. 主要内容参见如下文献:

- [1] Banaszak, Z. A. and Krogh, B. H., 1990. Deadlock avoidance in flexible manufacturing systems with concurrently competing process flows. *IEEE Transactions on Robotics and Automation*, Vol.6, No.6, pp. 724–734.
- [2] Viswanadham, N., Narahari, Y. and Johnson, T. L., 1990. Deadlock prevention and deadlock avoidance in flexible manufacturing systems using Petri net models. *IEEE Transactions on Robotics and Automation*, Vol.6, No.6, pp.713–723.
- [3] Uzam, M. and Zhou, M. C., 2007. An iterative synthesis approach to Petri net-based deadlock prevention policy for flexible manufacturing systems. *IEEE Transactions On Systems, Man, and Cybernetics—Part A: Systems And Humans*, Vol.37, No.3, pp.362 -371
- [4] Fanti, M. P. and Zhou, M. C, 2004. Deadlock control methods in automated manufacturing systems. *IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans*, Vol. 34, No.1, pp.5-22.

## 第 7 章

### 基于 Petri 网的制造系统分析与控制(3)

本章介绍基于 Petri 网的监控理论及其在制造系统控制中的应用, 包括基于受控 Petri 网(CtlPN, controlled Petri nets)的监控方法和基于库所不变性的监控方法(SBPI, supervision based on place invariants). 主要内容参见如下文献:

- [1] Krogh, B. H. and Holloway, L. E., 1991. Synthesis of feedback control logic for discrete manufacturing systems. *Automatica*, Vol.27, No.4, pp. 641–651.
- [2] Yamilidou, K., Moody, J., Lemmon, M. and Antsaklis, P., 1996. Feedback control of Petri nets based on place invariants. *Automatica*, Vol.32, No.1, pp. 15–28.
- [3] Moody, J. O. and Antsakli, P. J., 2000. Petri net supervisors for DES with uncontrollable and unobservable transitions. *IEEE Transactions on Automatic Control*, Vol.45, No.3, pp.462-476.

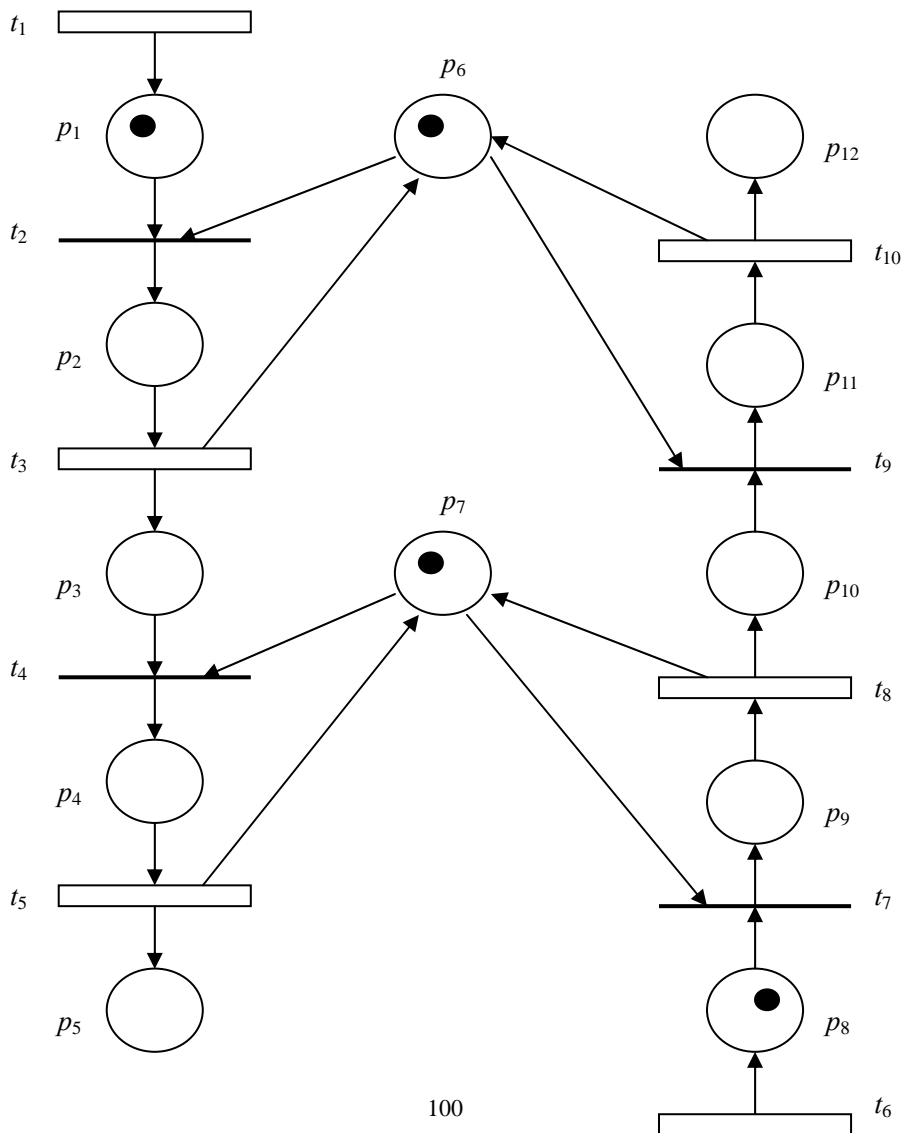
# 第 8 章

## 基于排队网络与随机模型的制造系统分析与控制(1)

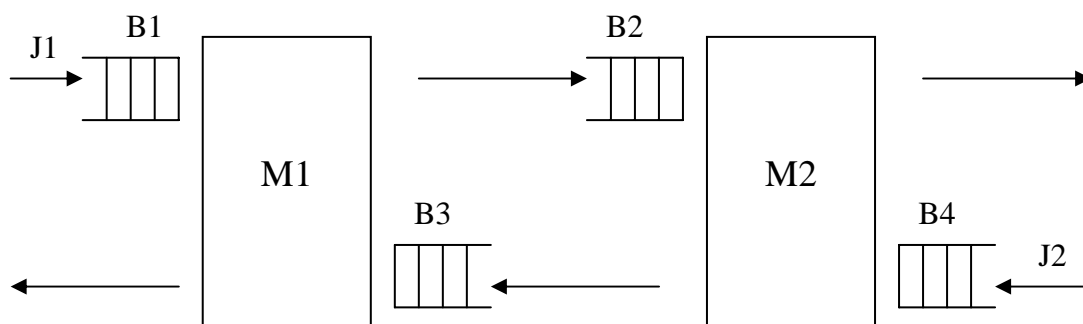
本章介绍基于排队网络(queueing networks)与随机模型制造系统建模、分析与控制理论. 排队网络理论是一种面向性能分析的离散事件动态系统理论, 除了应用于制造系统外, 还广泛地应用于通讯系统, 交通系统和服务系统等领域的分析. 此外, 还将介绍考虑随机因素的制造系统性能分析的一种近似算法: 分解算法.

### §8.1 Petri 网不能做什么?

前面几章介绍了基于 Petri 网的制造系统建模、分析与控制理论. Petri 网可以很细致地为制造系统建模, 在此基础上, 一方面可展开对制造系统逻辑行为的分析, 如死锁的可能性; 另一方面, 可以建立仿真模型. 然而, Petri 网是否可以解决制造系统分析的一切问题呢? 先来看下面的例子.



这是一个如下的 Job Shop 制造系统的 Petri 网模型



对上面的 Petri 网模型, 尝试进行可达性分析. 其中一种情况如下表所示.

变迁	标识	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12
	M0	1					1	1	1				
t1,t2,t6,t7	M1	1	1						1	1			
t1,t3,t6,t8	M2	2		1			1	1	2		1		
t1,t4,t6,t9	M3	3			1				3			1	
...	...	...	...	...	...	...	...	...	...	...	...	...	...

在这种情况下, 所有在逻辑上可以并发的变迁只要一被使能就触发. 而且赋时变迁 t1 和 t6 触发的时间间隔较之其他赋时变迁为短. 因此, 库所 p1 和 p8 内的令牌数量逐渐增加. 如果赋时变迁 t1 和 t6 触发的时间间隔较之其他赋时变迁为长, 则情况有所不同, 如下表所示.

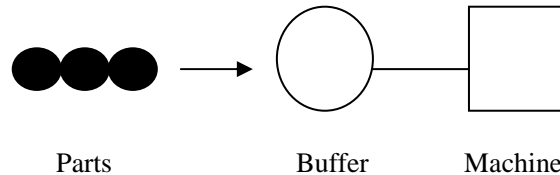
变迁	标识	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12
	M0	1					1	1	1				
t2,t7	M1		1							1			
t3,t8	M2			1			1	1			1		
t4,t9	M3				1							1	
t1,t5,t6,t10	M4	1				1	1	1	1				1
t2,t7	M5		1			1				1			1
...	...	...	...	...	...	...	...	...	...	...	...	...	...

由于这个 Petri 网模型中有赋时变迁, 因此变迁的触发不仅取决于它是否满足使能条件, 还取决于是否到达触发时间. 触发时间的不同造成了事件发生的先后顺序不同, 从而造成系统状态的不同, 最终反映为系统性能的不同. 上面的例子中, 显然第一种情况下系统不稳定, 第二种情况下系统稳定. 由于时间参数的不同, 系统的可达性分析的结果也很不一样. 因此, 在这种情况下, 用 Petri 网来分析系统的性能就比较困难了. 困难存在于三个方面:

1. 系统的状态空间很大, 时序复杂, 逻辑分析困难.
  2. 系统的一些宏观特性难于把握, 例如, 在机器前的缓冲区中等待加工的零件的数量, 即库所  $p1$  中令牌的数量, 它取决于变迁  $t1$  与  $t3$  的触发频率, 或者说零件到达的时间间隔和加工一个零件所需时间.
  3. Petri 网的描述能力强, 其优势在于关于逻辑的系统特性(logical characters)的分析与控制, 如死锁的避免; 而关于时间的系统特性(temporal characters)的分析与控制的能力则较弱.
- 因此, 需要应用其他的模型与方法来分析制造系统的关于时间的特性, 这就是排队网络(queueing networks)

## §8.2 排队网络(queueing networks)概述

首先考虑如下的单机器的制造系统, 它由一台机器, 一个无穷大缓冲区和大量工件构成.



系统具有随机性:

1. 相继到达的两个零件之间的时间间隔为独立同分布的, 服从负指数分布, 分布函数为:

$$F(t) = P\{\tau < t\} = \begin{cases} 1 - e^{-\lambda t}, & \text{若 } t \geq 0 \\ 0, & \text{否则} \end{cases}$$

容易计算出其均值为  $1/\lambda$ . 而  $\lambda\delta t$  则表示  $\delta t$  时间内有零件到达的概率. 在这种情况下, 零件的到达过程为最简单流, 即满足如下条件:

- (1) 平稳性. 在时间区间  $[a, a+t]$  内到达的零件数量为  $n$  的概率与  $a$  无关而仅与  $t$  有关;
- (2) 无后效性. 在不相交的时间区间内到达的零件数量的是相互独立的;
- (3) 有限性. 在有限的时间区间  $[a, a+t]$  中, 只有限多个零件到达;
- (4) 单个性. 在时间区间  $[a, a+t]$  中有超过一个零件到达的概率是区间长度  $t$  的高阶无穷小, 也就是说, 在充分短的时间区间内, 最多只有一个零件到达.

可以证明, 对于最简单流, 在在时间区间  $[a, a+t]$  内到达的零件数量服从 Poisson 分布, 其概率分布为

$$P\{x[a, a+t] = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, n = 0, 1, \dots$$

2. 零件的加工时间服从参数为  $1/\mu$  的负指数分布, 则加工时间的均值为  $1/\mu$ . 而  $\mu\delta t$  则表示  $\delta t$  时间内有零件完成加工的概率.

以  $p_n(t)$  表示在  $t$  时刻系统的在制品数量为  $n$  的概率, 这里  $n$  实际上就是系统的状态. 于是可以得到如下的方程:

当  $n > 0$  时,

$$p_n(t + \delta t) = p_n(t)(1 - \lambda\delta t)(1 - \mu\delta t) + p_{n+1}(t)(1 - \lambda\delta t)\mu\delta t + p_{n-1}(t)\lambda\delta t(1 - \mu\delta t) + o(\delta t)$$

$$\text{即 } \frac{dp_n(t)}{dt} = -p_n(t)(\lambda + \mu) + p_{n+1}(t)\mu + p_{n-1}(t)\lambda$$

稳态时,  $dp_n(t)/dt = 0$ , 于是可得

$$0 = -p_n(\lambda + \mu) + p_{n+1}\mu + p_{n-1}\lambda. \quad (1)$$

当  $n=0$  时,

$$p_0(t + \delta t) = p_0(t)(1 - \lambda\delta t) + p_1(t)(1 - \lambda\delta t)\mu\delta t + o(\delta t)$$

$$\text{即 } \frac{dp_0(t)}{dt} = -p_0(t)\lambda + p_1(t)\mu$$

稳态时,  $dp_n(t)/dt = 0$ , 于是可得

$$0 = -p_0\lambda + p_1\mu. \quad (2)$$

令  $\rho = \lambda/\mu$ , 由式(1)和(2), 根据数学归纳法可证得

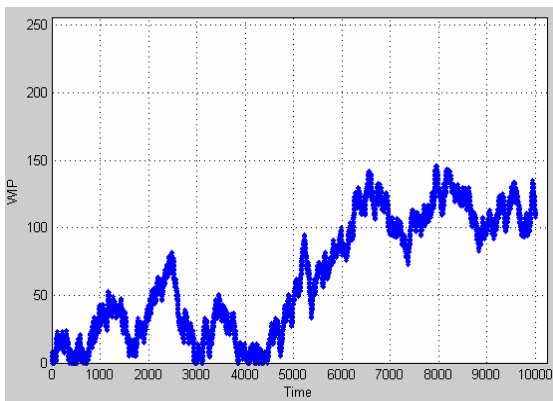
$$p_n = \rho p_{n-1} = \rho^n p_0. \quad (3)$$

概率  $p_n$  应满足如下的归一化条件

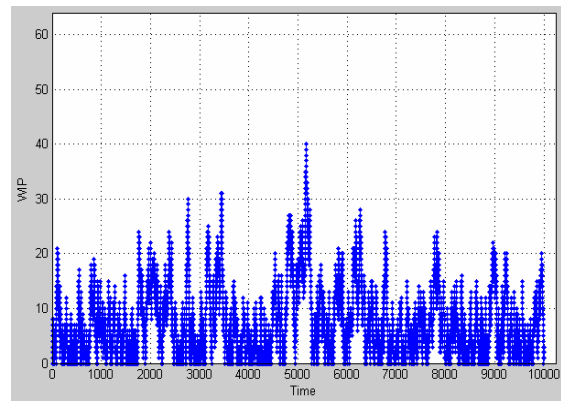
$$1 = \sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} p_0 \rho^n = \lim_{n \rightarrow \infty} \frac{p_0(1 - \rho^n)}{1 - \rho}$$

只有当  $\rho < 1$  时, 上式中的极限存在. 因此可得  $p_0 = 1 - \rho$ , 于是  $p_n = (1 - \rho)\rho^n$ .

可见  $\rho < 1$  是方程(1)和(2) (称为 概率平衡方程) 有解的必要条件, 也是系统稳定的必要条件, 事实上也是充分条件. 回忆 Lecture 3 中的仿真例子, 若零件到达时间间隔服从参数为 1 的指数分布, 零件加工时间也服从参数为 1 的指数分布, 则系统不稳定; 而如果把零件加工时间均值缩短为 0.9, 则系统稳定. 造成这一差异的原因就在于前者  $\rho = \lambda/\mu = 1$ , 而后者  $\rho = \lambda/\mu = 0.9 < 1$ .



$\rho = \lambda/\mu = 1$  时在制品数量



$\rho = \lambda/\mu = 0.9 < 1$  时在制品数量

计算在制品均值

$$WIP = \sum_{n=0}^{\infty} n p_n = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}. \quad (4)$$

继续计算生产周期的均值. 假设某个零件到达时系统的  $WIP=n$ , 按照先到先服务 (FCFS) 原则, 设此零件的生产周期为  $CT(n)$ , 包括等待原来  $n$  个零件完成的时间和其自身加工时间之和. 则  $CT(n) > x$  的充分且必要条件是在  $[0, x]$  这段时间内零



件加工完成的次数  $m(x) = i < n+1$ . 由于加工时间是负指数分布的, 因此加工完成事件序列是一个最简流. 于是加工完成次数  $m(x) = i$  的概率为

$$P\{m(x) = i\} = e^{-\mu x} \frac{(\mu x)^i}{i!}$$

于是,  $P\{CT(n) > x\} = \sum_{i=0}^n e^{-\mu x} \frac{(\mu x)^i}{i!}$ , 即  $P\{CT(n) < x\} = \sum_{i=n+1}^{\infty} e^{-\mu x} \frac{(\mu x)^i}{i!}$

因此, CT 的分布函数为

$$\begin{aligned} F_{CT}(x) &= P\{CT < x\} = \sum_{n=0}^{\infty} (1-\rho)\rho^n \sum_{i=n+1}^{\infty} e^{-\mu x} \frac{(\mu x)^i}{i!} = \sum_{i=1}^{\infty} \sum_{n=0}^{i-1} (1-\rho)\rho^n e^{-\mu x} \frac{(\mu x)^i}{i!} \\ &= (1-\rho) \sum_{i=1}^{\infty} \frac{1-\rho^i}{1-\rho} e^{-\mu x} \frac{(\mu x)^i}{i!} = \sum_{i=1}^{\infty} (1-\rho^i) \frac{e^{-\mu x} (\mu x)^i}{i!} = \sum_{i=1}^{\infty} \frac{e^{-\mu x} (\mu x)^i}{i!} - \sum_{i=1}^{\infty} \frac{e^{-\mu x} (\rho \mu x)^i}{i!} \\ &= e^{-\mu x} (e^{\mu x} - 1) - e^{-\mu x} (e^{\rho \mu x} - 1) = 1 - e^{-\mu(1-\rho)x} \end{aligned}$$

也为一负指数分布函数. 因此, CT 的均值为

$$CT = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu - \lambda}. \quad (5)$$

由式(4)和(5)可见

$$\frac{WIP}{CT} = \lambda = TH. \quad (6)$$

满足 Little 定律.

实际上, 上述分析过程是将单机器制造系统看作一个 M/M/1 排队系统.

### M/M/c 排队系统

由于有多台机器, 因此服务率与等待加工的零件数及机器数有关, 即  $\mu_n = \mu \min\{n, c\}$ . 同样以  $p_n(t)$  表示在  $t$  时刻系统的在制品数量为  $n$  的概率, 可以得到如下的方程:

当  $n > 0$  时,

$$p_n(t + \delta t) = p_n(t)(1 - \lambda \delta t)(1 - \mu_n \delta t) + p_{n+1}(t)(1 - \lambda \delta t)\mu_{n+1} \delta t + p_{n-1}(t)\lambda \delta t(1 - \mu_{n-1} \delta t) + o(\delta t)$$

$$\text{即 } \frac{dp_n(t)}{dt} = -p_n(t)(\lambda + \mu_n) + p_{n+1}(t)\mu_{n+1} + p_{n-1}(t)\lambda$$

稳态时,  $dp_n(t)/dt = 0$ , 于是可得

$$0 = -p_n(\lambda + \mu_n) + p_{n+1}\mu_{n+1} + p_{n-1}\lambda. \quad (7)$$

当  $n=0$  时,

$$p_0(t + \delta t) = p_0(t)(1 - \lambda \delta t) + p_1(t)(1 - \lambda \delta t)\mu_1 \delta t + o(\delta t)$$

$$\text{即 } \frac{dp_0(t)}{dt} = -p_0(t)\lambda + p_1(t)\mu_1$$

稳态时,  $dp_n(t)/dt = 0$ , 于是可得

$$0 = -p_0\lambda + p_1\mu_n. \quad (8)$$

注意: 若  $n < c$ , 则  $\mu_n = n\mu$ ; 若  $n \geq c$ , 则  $\mu_n = c\mu$ . 于是由式(7)和(8)可得, 当  $n < c$  时,

$$p_1\mu_1 = \lambda p_0 \rightarrow p_1 = \frac{\lambda}{\mu} p_0$$

$$p_2 \mu_2 = p_1 (\lambda + \mu_1) - p_0 \lambda \rightarrow p_2 = \frac{\lambda^2}{2\mu^2} p_0$$

$$p_3 \mu_3 = p_2 (\lambda + \mu_2) - p_1 \lambda \rightarrow p_3 = \frac{\lambda^3}{3!\mu^3} p_0$$

.....

假设  $p_k = \frac{\lambda^k}{k!\mu^k} p_0$ ,  $k=1, \dots, n$ , 于是可得

$$p_{n+1} \mu_{n+1} = p_n (\lambda + \mu_n) - p_{n-1} \lambda = p_0 \frac{\lambda^n}{n!\mu^n} (\lambda + \mu_n) - p_0 \frac{\lambda^{n-1}}{(n-1)!\mu^{n-1}} \lambda \rightarrow$$

$$p_{n+1} (n+1)\mu = p_0 \frac{\lambda^n}{n!\mu^n} (\lambda + n\mu) - p_0 \frac{n\mu\lambda^n}{n!\mu^n} = p_0 \frac{\lambda^{n+1}}{n!\mu^n} \rightarrow p_{n+1} = p_0 \frac{\lambda^{n+1}}{(n+1)!\mu^{n+1}}$$

由数学归纳法原理可知:

$$p_n = \frac{\lambda^n}{n!\mu^n} p_0 \text{ 对于 } n \leq c \text{ 成立.} \quad (9)$$

当  $n=c$  时,  $p_c = \frac{\lambda^c}{c!\mu^c} p_0$ .

当  $n > c$  时, 由式(7)可知,

$$p_{c+1} \mu_{c+1} = p_c (\lambda + \mu_c) - p_{c-1} \lambda \rightarrow$$

$$p_{c+1} c\mu = \frac{\lambda^c}{c!\mu^c} p_0 (\lambda + c\mu) - \frac{\lambda^{c-1}}{(c-1)!\mu^{c-1}} p_0 \lambda = \frac{\lambda^c}{c!\mu^c} p_0 \lambda \rightarrow p_{c+1} = \frac{\lambda^{c+1}}{c \cdot c!\mu^{c+1}} p_0;$$

$$p_{c+2} \mu_{c+2} = p_{c+1} (\lambda + \mu_{c+1}) - p_c \lambda \rightarrow p_{c+2} = \frac{\lambda^{c+2}}{c^2 \cdot c!\mu^{c+2}} p_0$$

$$p_{c+3} \mu_{c+3} = p_{c+2} (\lambda + \mu_{c+2}) - p_{c+1} \lambda \rightarrow p_{c+3} = \frac{\lambda^{c+3}}{c^3 \cdot c!\mu^{c+3}} p_0$$

.....

假设  $p_{c+k} = \frac{\lambda^{c+k}}{c^k \cdot c!\mu^{c+k}} p_0$ ,  $k=1, \dots, n$ , 于是可得

$$p_{c+(n+1)} \mu_{c+(n+1)} = p_{c+n} (\lambda + \mu_{c+n}) - p_{c+(n-1)} \lambda \rightarrow$$

$$p_{c+(n+1)} c\mu = \frac{\lambda^{c+n}}{c^n \cdot c!\mu^{c+n}} p_0 (\lambda + c\mu) - \frac{\lambda^{c+(n-1)}}{c^{n-1} \cdot c!\mu^{c+(n-1)}} p_0 \lambda = \frac{\lambda^{c+n}}{c^n \cdot c!\mu^{c+n}} p_0 \lambda \rightarrow$$

$$p_{c+(n+1)} = \frac{\lambda^{c+(n+1)}}{c^{n+1} \cdot c!\mu^{c+(n+1)}} p_0$$

于是同样由数学归纳法原理可得  $p_{c+n} = \frac{\lambda^{c+n}}{c^n \cdot c!\mu^{c+n}} p_0$  对于  $n=1, 2, \dots$  成立, 即

$$p_n = \frac{\lambda^n}{c^{n-c} \cdot c!\mu^n} p_0 \text{ 对于 } n > c \text{ 成立.} \quad (10)$$

综合式(9)和(10)可得, M/M/c 排队系统的 WIP 的概率分布为

$$p_n = \begin{cases} \frac{\lambda^n}{n! \mu^n} p_0 = \frac{\rho^n}{n!} p_0, & \text{if } n \leq c \\ \frac{\lambda^n}{c^{n-c} \cdot c! \mu^n} p_0 = \frac{\rho^n}{c^{n-c} \cdot c!} p_0, & \text{if } n > c \end{cases}. \quad (11)$$

同时,  $\sum_{n=0}^{\infty} p_n = 1$  应成立, 于是可得  $p_0 = \left[ \sum_{n=0}^c \frac{\rho^n}{n!} + \frac{\rho^c}{c!(c-\rho)} \right]^{-1}$ .

在制品均值

$$\begin{aligned} WIP &= \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^c n \frac{\rho^n}{n!} p_0 + \sum_{n=c+1}^{\infty} n \frac{\rho^n}{c^{n-c} \cdot c!} p_0 = \rho p_0 \sum_{n=1}^c \frac{\rho^{n-1}}{(n-1)!} + p_0 \frac{c^c}{c!} \sum_{n=1}^{\infty} (c+n) \left( \frac{\rho}{c} \right)^{c+n} \\ &= \left[ \rho \sum_{n=1}^c \frac{\rho^{n-1}}{(n-1)!} + \frac{\rho^{c+1}(c-\rho+1)}{(c-1)!(c-\rho)^2} \right] p_0. \end{aligned}$$

Job Shop 制造系统的分析:

设有一 Job Shop 制造系统, 由  $m$  个加工中心构成, 每个加工中心又由  $c_i$  台相同的机器构成,  $i=1, \dots, m$ . 生产  $r$  种零件, 每种零件的工艺路线各不相同, 设加工第  $l$  种零件共有  $n_l$  道工序, 其中的第  $k$  道工序的加工中心为  $C_{lk}$ , 加工时间服从参数为  $\mu_{lk}$  的负指数分布. 令  $N_l(t)$  表示时间区间  $[0, t]$  内第  $l$  种零件进入此 Job Shop 的数量, 令  $\lambda(l) = \lim_{t \rightarrow \infty} N_l(t)/t$ . 则  $\lambda_l$  为第  $l$  种零件的到达率. 若把各种零件看作一个整体,

则总的到达率为  $\lambda = \sum_{l=1}^r \lambda(l)$ . 令指示函数(indicator function)  $I(P) = 1$ , 若命题  $P$  成立;  $I(P) = 0$ , 若命题  $P$  不成立. 则所有零件中第一道工序在加工中心  $i$  上完成的概率为

$$\gamma_i = \frac{1}{\lambda} \sum_{l=1}^r \lambda(l) I(C_{l1} = i)$$

从加工中心  $i$  到加工中心  $j$  的转移率为  $\lambda_{ij} = \sum_{l=1}^r \lambda(l) \sum_{k=1}^{n_l-1} I(C_{lk} = i, C_{l,k+1} = j)$ , 而零件

进入加工中心  $i$  的到达率为  $\lambda_i = \sum_{l=1}^r \lambda(l) \sum_{k=1}^{n_l} I(C_{lk} = i)$ , 于是从加工中心  $i$  到加工中心  $j$  的转移的概率为

$$p_{ij} = \frac{\lambda_{ij}}{\lambda_i}.$$

$p_{ij}$  构成状态转移概率矩阵  $P = (p_{ij})_{m \times m}$ . 下面来考察  $\lambda_j$ ,  $\gamma_j$  与  $p_{ij}$  之间的关系.

$$\begin{aligned} \lambda \gamma_i + \sum_{j=1, j \neq i}^m \lambda_j p_{ji} &= \sum_{l=1}^r \lambda(l) I(C_{l1} = i) + \sum_{j=1, j \neq i}^m \sum_{l=1}^r \lambda(l) \sum_{k=1}^{n_l-1} I(C_{lk} = j, C_{l,k+1} = i) \\ &= \sum_{l=1}^r \lambda(l) \left[ I(C_{l1} = i) + \sum_{j=1, j \neq i}^m \sum_{k=1}^{n_l-1} I(C_{lk} = j, C_{l,k+1} = i) \right] = \sum_{l=1}^r \lambda(l) \left[ \sum_{k=1}^{n_l} I(C_{lk} = i) \right] = \lambda_i \end{aligned}$$

于是有下式成立

$$\lambda_i = \lambda \gamma_i + \sum_{j=1, j \neq i}^m \lambda_j p_{ji}. \quad (12)$$

令  $\mathbf{n}(t) = (n_1(t), n_2(t), \dots, n_m(t))$  表示  $t$  时刻每个加工中心的 WIP 数量, 这是 Job Shop 系统的状态, 则  $\mathbf{n}(t + \delta t) = (n_1(t + \delta t), n_2(t + \delta t), \dots, n_m(t + \delta t))$  可能由状态  $\mathbf{n}(t) = (n_1(t), n_2(t), \dots, n_m(t))$ ,  $\mathbf{n}(t) + \mathbf{e}_j = (n_1(t), \dots, n_j(t) + 1, \dots, n_m(t))$ ,  $\mathbf{n}(t) - \mathbf{e}_i = (n_1(t), \dots, n_i(t) - 1, \dots, n_m(t))$ ,  $\mathbf{n}(t) + \mathbf{e}_j - \mathbf{e}_i = (n_1(t), \dots, n_j(t) + 1, \dots, n_i(t) - 1, \dots, n_m(t))$  演变而来. 假定  $p[\mathbf{n}(t)]$  为系统在  $t$  时刻处于状态  $\mathbf{n}(t)$  的概率, 则下式成立:

$$\begin{aligned}
p[\mathbf{n}(t + \delta t)] &= p[\mathbf{n}(t)] \prod_{i=1}^m (1 - \lambda \gamma_i \delta t) (1 - \mu_i(n_i) \delta t) + \sum_{j=1}^m p[\mathbf{n}(t) + \mathbf{e}_j] \mu_j(n_j + 1) \delta t \left(1 - \sum_{i=1}^m p_{ji}\right) \\
&+ \sum_{i=1}^m p[\mathbf{n}(t) - \mathbf{e}_i] \lambda \gamma_i \delta t + \sum_{j=1}^m \sum_{i=1, i \neq j}^m p[\mathbf{n}(t) + \mathbf{e}_j - \mathbf{e}_i] \mu_j(n_j + 1) \delta t p_{ji} + o(\delta t) \\
&= p[\mathbf{n}(t)] \left[1 - \sum_{i=1}^m (\lambda \gamma_i + \mu_i(n_i)) \delta t\right] + \sum_{j=1}^m p[\mathbf{n}(t) + \mathbf{e}_j] \mu_j(n_j + 1) \left(1 - \sum_{i=1}^m p_{ji}\right) \delta t \\
&+ \sum_{i=1}^m p[\mathbf{n}(t) - \mathbf{e}_i] \lambda \gamma_i \delta t + \sum_{j=1}^m \sum_{i=1, i \neq j}^m p[\mathbf{n}(t) + \mathbf{e}_j - \mathbf{e}_i] \mu_j(n_j + 1) p_{ji} \delta t + o(\delta t) \\
\Rightarrow \frac{dp[\mathbf{n}(t)]}{dt} &= -p[\mathbf{n}(t)] \sum_{i=1}^m (\lambda \gamma_i + \mu_i(n_i)) + \sum_{j=1}^m p[\mathbf{n}(t) + \mathbf{e}_j] \mu_j(n_j + 1) \left(1 - \sum_{i=1}^m p_{ji}\right) \\
&+ \sum_{i=1}^m p[\mathbf{n}(t) - \mathbf{e}_i] \lambda \gamma_i + \sum_{j=1}^m \sum_{i=1, i \neq j}^m p[\mathbf{n}(t) + \mathbf{e}_j - \mathbf{e}_i] \mu_j(n_j + 1) p_{ji}
\end{aligned}$$

稳态下,  $dp[\mathbf{n}(t)]/dt = 0$ ,  $p[\mathbf{n}(t)] = p(\mathbf{n})$ , 于是上式成为:

$$\begin{aligned}
&p(\mathbf{n}) \sum_{i=1}^m (\lambda \gamma_i + \mu_i(n_i)) \\
&= \sum_{j=1}^m p(\mathbf{n} + \mathbf{e}_j) \mu_j(n_j + 1) \left(1 - \sum_{i=1}^m p_{ji}\right) + \sum_{i=1}^m p(\mathbf{n} - \mathbf{e}_i) \lambda \gamma_i + \sum_{j=1}^m \sum_{i=1, i \neq j}^m p(\mathbf{n} + \mathbf{e}_j - \mathbf{e}_i) \mu_j(n_j + 1) p_{ji}
\end{aligned} \tag{13}$$

上式中, 等号左侧是离开状态  $\mathbf{n}$  的概率, 等号右侧是进入状态  $\mathbf{n}$  的概率, 式(13)为系统状态的 概率平衡方程, 系统的稳态概率分布  $p(\mathbf{n})$  应满足上述方程.

假定系统的稳态概率分布具有如下的乘积形式:

$$p(\mathbf{n}) = \prod_{i=1}^m p_i(n_i) \tag{14}$$

现在来验证它是否满足稳态概率平衡方程(13). 把式(14)代入上式并以之同时除等式两侧, 可得

$$\begin{aligned}
&\sum_{i=1}^m (\lambda \gamma_i + \mu_i(n_i)) \\
&= \sum_{j=1}^m \frac{p_j(n_j + 1)}{p_j(n_j)} \mu_j(n_j + 1) \left(1 - \sum_{i=1}^m p_{ji}\right) + \sum_{i=1}^m \frac{p_i(n_i - 1)}{p_i(n_i)} \lambda \gamma_i \\
&+ \sum_{j=1}^m \sum_{i=1, i \neq j}^m \frac{p_j(n_j + 1) p_i(n_i - 1)}{p_j(n_j) p_i(n_i)} \mu_j(n_j + 1) p_{ji}
\end{aligned}$$

令  $p_{ii} = 0$ , 即不存在一个零件重复经过同一个加工中心的情形, 则由上式可得

$$\sum_{i=1}^m (\lambda \gamma_i + \mu_i(n_i))$$

$$\begin{aligned}
&= \sum_{i=1}^m \frac{p_i(n_i+1)}{p_i(n_i)} \mu_i(n_i+1) - \sum_{i=1}^m \left[ \sum_{j=1}^m \frac{p_j(n_j+1)}{p_j(n_j)} \mu_j(n_j+1) p_{ji} \right] \\
&\quad + \sum_{i=1}^m \frac{p_i(n_i-1)}{p_i(n_i)} \lambda \gamma_i + \sum_{i=1}^m \left[ \sum_{j=1}^m \frac{p_j(n_j+1) p_i(n_i-1)}{p_j(n_j) p_i(n_i)} \mu_j(n_j+1) p_{ji} \right] \\
&= \sum_{i=1}^m \left[ \frac{p_i(n_i+1)}{p_i(n_i)} \mu_i(n_i+1) - \sum_{j=1}^m \frac{p_j(n_j+1)}{p_j(n_j)} \mu_j(n_j+1) p_{ji} \right. \\
&\quad \left. + \frac{p_i(n_i-1)}{p_i(n_i)} \left( \lambda \gamma_i + \sum_{j=1}^m \frac{p_j(n_j+1)}{p_j(n_j)} \mu_j(n_j+1) p_{ji} \right) \right] \tag{15}
\end{aligned}$$

现在来寻求使下式成立的条件:

$$\begin{aligned}
\lambda \gamma_i + \mu_i(n_i) &= \frac{p_i(n_i+1)}{p_i(n_i)} \mu_i(n_i+1) - \sum_{j=1}^m \frac{p_j(n_j+1)}{p_j(n_j)} \mu_j(n_j+1) p_{ji} \\
&\quad + \frac{p_i(n_i-1)}{p_i(n_i)} \left( \lambda \gamma_i + \sum_{j=1}^m \frac{p_j(n_j+1)}{p_j(n_j)} \mu_j(n_j+1) p_{ji} \right). \tag{16}
\end{aligned}$$

若式(16)成立, 则式(15)必然成立.

令  $w_i = \frac{p_i(n_i+1)}{p_i(n_i)} \mu_i(n_i+1)$ ,  $w_j = \frac{p_j(n_j+1)}{p_j(n_j)} \mu_j(n_j+1)$ , 于是式(15)成为

$$\lambda \gamma_i + \mu_i(n_i) = w_i - \sum_{j=1}^m w_j p_{ji} + \frac{p_i(n_i-1)}{p_i(n_i)} \left( \lambda \gamma_i + \sum_{j=1}^m w_j p_{ji} \right)$$

若有  $w_i = \lambda \gamma_i + \sum_{j=1}^m w_j p_{ji}$  成立, 则上式成为

$$\lambda \gamma_i + \mu_i(n_i) = \lambda \gamma_i + \frac{p_i(n_i-1)}{p_i(n_i)} w_i \rightarrow p_i(n_i) = p_i(n_i-1) \frac{w_i}{\mu_i(n_i)}. \tag{17}$$

于是式(16)成立的条件是  $w_i = \lambda \gamma_i + \sum_{j=1}^m w_j p_{ji}$  及  $p_i(n_i) = p_i(n_i-1) \frac{w_i}{\mu_i(n_i)}$ . 由式(12)

可知, 加工中心的零件到达率之间的关系  $\lambda_i = \lambda \gamma_i + \sum_{j=1, j \neq i}^m \lambda_j p_{ji}$  恰好满足

$w_i = \lambda \gamma_i + \sum_{j=1}^m w_j p_{ji}$ . 而式(17)成为

$$p_i(n_i) = p_i(n_i-1) \frac{\lambda_i}{\mu_i(n_i)} = \dots = p_i(0) \frac{\lambda_i^{n_i}}{\mu_i(n_i) \mu_i(n_i-1) \dots \mu_i(1)}$$

若当  $n_i \leq c_i$  时,  $\mu_i(n_i) = n_i \mu_i$ ; 当  $n_i > c_i$  时,  $\mu_i(n_i) = c_i \mu_i$ , 则可得

$$p_i(n_i) = \begin{cases} \frac{\lambda_i^{n_i}}{n_i! \mu_i^{n_i}} p_i(0) = \frac{\rho_i^{n_i}}{n_i!} p_i(0), & \text{if } n_i \leq c_i \\ \frac{\lambda_i^{n_i}}{c_i^{n_i-c_i} c_i! \mu_i^c} p_i(0) = \frac{\rho_i^{n_i}}{c_i^{n_i-c_i} c_i!} p_i(0), & \text{if } n_i > c_i \end{cases} \tag{18}$$

这里,  $\rho_i = \lambda_i / \mu_i$ . 这是 M/M/c 排队系统的状态概率分布.

这里讨论的 Job Shop 可以看作一个开环 Jackson 排队网络(open Jackson queueing

networks), 其系统状态的概率分布具有乘积的形式(14), 这就使得系统中的每个加工中心都可被看作是一个 M/M/c 排队系统, 且它们之间是可以被看作是相互独立的, 其各自的稳态概率分布具有式(18)的形式; 它们之间的物流关系满足式(12). 于是 M/M/c 系统的 WIP 的均值的求法可用于此开 Jackson 排队网络系统. 每个加工中心的生产率即为  $\lambda_j$ , 它是方程组

$$\lambda_i = \lambda\gamma_i + \sum_{j=1, j \neq i}^m \lambda_j p_{ji}, \quad i = 1, \dots, m$$

的解.

对于柔性制造系统(flexible manufacturing systems, FMS), 可用闭环 Jackson 排队网络(closed Jackson queueing networks)来描述. 这是因为在 FMS 中, 被加工的零件通常是固定在托盘(pallet)上的, 而托盘的数量通常是固定的, 当一个零件完成所有加工工序离开 FMS 时, 它所占用的托盘就释放出来, 从而可以安装一个新的毛坯, 让它进入 FMS. 也就是说, FMS 中零件的总量是不变的. 闭环 Jackson 排队网络与开环 Jackson 排队网络的区别也正在于此, 即系统的状态必须满足如下条件

$$\sum_{i=1}^m n_i = N, \quad (19)$$

这里, N 为 FMS 中零件的总量.

闭环 Jackson 排队网络(closed Jackson queueing networks)的稳态概率分布如下:

$$p(n_1, \dots, n_m) = \frac{1}{G(N)} \prod_{i=1}^m \rho_i^{n_i} / A_i(n_i), \quad (20)$$

其中,

$\rho_i = v_i / \mu_i$ ,  $\mu_i$  是加工中心  $i$  中单台机器的生产率;

$$A_i(n_i) = \begin{cases} 1, & \text{if } n_i = 0 \\ \prod_{j=1}^{n_i} a_i(j), & \text{if } n_i \geq 1 \end{cases};$$

$$G(N) = \sum_{n_1 + \dots + n_m = N} \prod_{i=1}^m \rho_i^{n_i} / A_i(n_i);$$

$$a_i(j) = \begin{cases} j, & \text{if } j \leq c_i \\ c_i, & \text{if } j > c_i \end{cases} \text{ 为加工中心 } i \text{ 的生产率系数, } a_i(n_i)\mu_i \text{ 为生产率};$$

$v_i$  是方程组  $\sum_{i=1}^m p_{ij} v_i = v_j$ ,  $j = 1, \dots, m$  的解.

## 第 9 章

# 基于排队网络与随机模型的制造系统分析与控制(2)

本章介绍基于排队网络模型的制造系统稳定性分析理论. 在上一章中, 给出了开环 Jackson 排队网络和闭环 Jackson 排队网络的稳态概率分布, 它们都具有某种乘积的形式. 既然可以显式地给出稳态概率分布, 则说明系统是稳定的. 然而, 并不是所有的排队网络的稳态概率分布都存在. Jackson 排队网络的稳态概率分布之所以可以显式地写出来, 是基于这样一个假设: 在同一机器上加工不同的工序所需的加工时间是独立同分布的, 并且服从指数分布. 如果在同一机器上加工不同的工序所需的加工时间不是独立同分布的, 则不能获得上述乘积形式的稳态概率分布. 而要判定系统是否稳定就需要采取其它方法. 本章介绍一种类似与 Lyapunov 直接方法的排队网络稳定性分析方法. 本章的内容参见以下文献:

- [1] Kumar, S. and Kumar, P. R., 2001. Queueing network models in the design and analysis of semiconductor wafer fabs. *IEEE Transactions on Robotics and Automation*, Vol.17, No.5, pp.548-561.
- [2] Kumar, P. R. and Meyn, S. P., 1995. Stability of queueing networks and scheduling policies. *IEEE Transactions on Automatic Control*, Vol.40, No.2, pp.251-260.
- [3] Kumar, S. and Kumar, P. R., 1994. Performance bounds for queueing networks and scheduling policies. *IEEE Transactions on Automatic Control*, Vol.39, No.8, pp.1600-1611.

## 第 10 章

# 基于排队网络与随机模型的制造系统分析与控制(3)

上面两章中的制造系统的排队网络模型具有如下的一些特征:

- (1) 系统的随机性存在于零件的到达时间间隔和机器对零件的加工时间;
- (2) 每个加工单元前的缓冲区容量是无限的.

事实上, 系统的随机性还可能来自与机器的不可靠性, 即机器可能发生故障; 同时, 缓冲区的容量也不可能是无限的. 因此, 本章将介绍基于另外一类制造系统随机模型, 即机器不可靠且缓冲区有限的制造系统的性能分析方法---近似分解方法.

### §10.1 加工时间确定时串行生产线的近似分解分析

本节内容参考如下文献:

- [1] Gershwin, S. B., 1987. An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking. *Operations Research*, Vol.35, No.2, pp.291-305.

### §10.2 加工时间为随机变量时串行生产线的近似分解分析

先考虑两台机器构成的串行短生产线, 加工一种零件, 两机器之间有一有限的缓冲区, 容量为  $N$ . 第一台机器不会发生缺货(*be starved*), 但由于缓冲区是有限的, 可能会被阻塞(*be blocked*); 第二台机器不会发生阻塞, 但有可能发生缺货. 加工时间分别服从参数为  $\mu_1$  和  $\mu_2$  的负指数分布. 机器仅在加工零件时发生故障, 故障时间间隔均值(MTTF)分别为  $1/p_1$  和  $1/p_2$ , 修复时间均值(MTTR)分别为  $1/r_1$  和  $1/r_2$ . 系统的状态:  $(n, \alpha_1, \alpha_2)$ , 其中,  $n = 0, 1, \dots, N$ , 为在制品数量;  $\alpha_i = 1, 0$ ,  $i = 1, 2$  分别表示两台机器的正常或故障状态.

当  $n = 1, \dots, N-1$  时, 稳态概率平衡方程为

$$p(n,1,1)(\mu_1 + \mu_2 + p_1 + p_2) = p(n,1,0)r_2 + p(n,0,1)r_1 + p(n-1,1,1)\mu_1 + p(n+1,1,1)\mu_2$$

$$p(n,1,0)(\mu_1 + p_1 + r_2) = p(n,1,1)p_2 + p(n,0,0)r_1 + p(n-1,1,0)\mu_1$$

$$p(n,0,1)(\mu_2 + r_1 + p_2) = p(n,1,1)p_1 + p(n,0,0)r_2 + p(n+1,0,1)\mu_2$$

$$p(n,0,0)(r_1 + r_2) = p(n,1,0)p_1 + p(n,0,1)p_2$$

当  $n = 0$  时, 稳态概率平衡方程(边界条件 1)为



$$p(0,1,1)(\mu_1 + p_1) = p(0,1,0)r_2 + p(0,0,1)r_1 + p(1,1,1)\mu_2$$

$$p(0,1,0)(\mu_1 + p_1 + r_2) = p(0,0,0)r_1$$

$$p(0,0,1)r_1 = p(0,1,1)p_1 + p(0,0,0)r_2 + p(1,0,1)\mu_2$$

$$p(0,0,0)(r_1 + r_2) = p(0,1,0)p_1$$

当  $n = N$  时, 稳态概率平衡方程(边界条件 2)为

$$p(N,1,1)(\mu_2 + p_2) = p(N,1,0)r_2 + p(N,0,1)r_1 + p(N-1,1,1)\mu_1$$

$$p(N,1,0)r_2 = p(N,1,1)p_2 + p(N,0,0)r_1 + p(N-1,1,0)\mu_1$$

$$p(N,0,1)(\mu_2 + r_1 + p_2) = p(N,0,0)r_2$$

$$p(N,0,0)(r_1 + r_2) = p(N,0,1)p_2$$

上述方程组的解为:

$$p(0,0,0) = 0, \quad p(0,0,1) = \frac{1}{r_1} \sum_{j=1}^4 C_j (p_1 Y_{1j} + \mu_2 X_j) Y_{2j}, \quad p(0,1,0) = 0,$$

$$p(0,1,1) = \sum_{j=1}^4 C_j Y_{1j} Y_{2j}, \quad p(n, \alpha_1, \alpha_2) = \sum_{j=1}^4 C_j X_j^n Y_{1j}^{\alpha_1} Y_{2j}^{\alpha_2}, \quad n = 1, \dots, N-1,$$

$$p(N,0,0) = 0, \quad p(N,0,1) = 0, \quad p(N,1,0) = \frac{1}{r_2} \sum_{j=1}^4 C_j X_j^{N-1} Y_{1j} (p_2 X_j Y_{2j} + \mu_1),$$

$$p(N,1,1) = \sum_{j=1}^4 C_j X_j^N Y_{1j} Y_{2j}$$

其中  $X_j, Y_{1j}, Y_{2j}, j = 1, \dots, 4$  是方程组  $Y_1 p_1 + Y_2 p_2 - r_1 - r_2 = 0$ ,

$$\mu_1 (X^{-1} - 1) + Y_1^{-1} r_1 - Y_1 p_1 + r_1 - p_1 = 0, \quad \mu_2 (X - 1) + Y_2^{-1} r_2 - Y_2 p_2 + r_2 - p_2 = 0$$

的解.  $C_j, j = 1, \dots, 4$  是方程组  $C_1 = 0, \quad \sum_{j=2}^4 C_j Y_{1j} = 0, \quad \sum_{j=2}^4 C_j X_j^N Y_{2j} = 0,$

$$\sum_{n=0}^N \sum_{\alpha_1=0}^1 \sum_{\alpha_2=0}^1 p(n, \alpha_1, \alpha_2) = 1 \text{ 的解.}$$

性能指标计算如下:

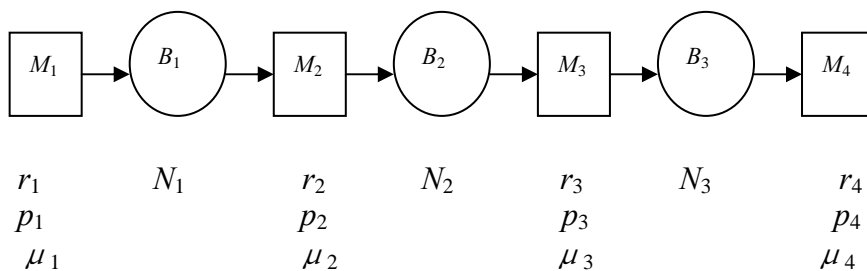
$$\text{生产率 } P = \mu_1 e_1 (1 - p_b) = \mu_2 e_2 (1 - p_s), \quad e_i = \frac{r_i}{p_i + r_i}$$

缺货概率  $p_s = p(0,0,1)$ , 阻塞概率  $p_b = p(N,1,0)$ , 在制品均值

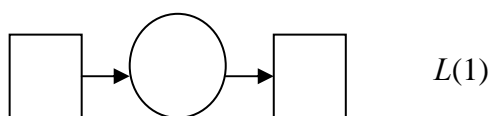
$$\bar{n} = \sum_{n, \alpha_1, \alpha_2} n p(n, \alpha_1, \alpha_2)$$

串行长生产线的分析:

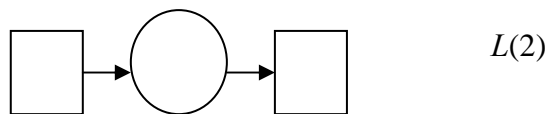
设有一串行生产线, 由  $k$  台机器构成, 相邻两机器之间有一缓冲区, 缓冲区容量是有限的.



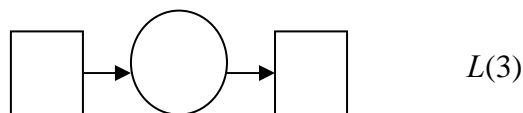
这一系统的状态数量为:  $2^k \prod_{i=1}^{k-1} (N_i + 1)$ . 分析起来相当困难, 因此采用近似计算方法. 把上述长生产线划分成  $k-1$  个由两台机器和一个缓冲区构成的短生产线.



$M_u(1)$      $B(1)$      $M_d(1)$   
 $r_u(1)$      $N_1$      $r_d(1)$   
 $p_u(1)$         $p_d(1)$   
 $\mu_u(1)$         $\mu_d(1)$



$M_u(2)$      $B(2)$      $M_d(2)$   
 $r_u(2)$      $N_1$      $r_d(2)$   
 $p_u(2)$         $p_d(2)$   
 $\mu_u(2)$         $\mu_d(2)$



$M_u(3)$      $B(3)$      $M_d(3)$   
 $r_u(3)$      $N_1$      $r_d(3)$   
 $p_u(3)$         $p_d(3)$   
 $\mu_u(3)$         $\mu_d(3)$

在短生产线  $L(i)$  中, 缓冲区  $B(i)$  的容量等于长生产线中相应的缓冲区容量  $N_i$ , 上游虚拟机器  $M_u(i)$  模拟长生产线中缓冲区  $B_i$  的上游部分, 下游虚拟机器  $M_d(i)$  模拟长生产线中缓冲区  $B_i$  的下游部分. 要保证分解前后的生产线等价, 为此要计算出  $r_u(i), p_u(i), \mu_u(i), r_d(i), p_d(i), \mu_d(i), i = 1, \dots, k-1$ .

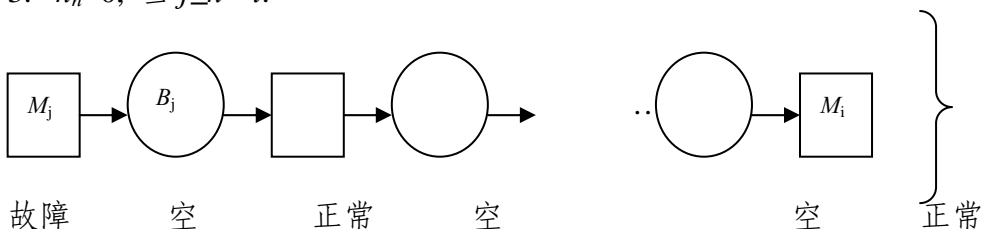
- $r_u(i)\delta t$ : 在  $(t, t + \delta t)$  时间区间内  $M_u(i)$  由故障状态恢复正常的概率;
- $p_u(i)\delta t$ : 在  $(t, t + \delta t)$  时间区间内  $M_u(i)$  在未被阻塞的情况下发生故障的概率;
- $\mu_u(i)\delta t$ : 在  $(t, t + \delta t)$  时间区间内  $M_u(i)$  处于正常状态且未被阻塞的情况有一个零件到达  $B(i)$  的概率;

$r_d(i)\delta t$ : 在  $(t, t + \delta t)$  时间区间内  $M_d(i)$  由故障状态恢复正常的概率;  
 $p_d(i)\delta t$ : 在  $(t, t + \delta t)$  时间区间内  $M_d(i)$  在没有缺货的情况下发生故障的概率;  
 $\mu_d(i)\delta t$ : 在  $(t, t + \delta t)$  时间区间内  $M_d(i)$  处于正常状态且没有缺货的情况有一个零件离开  $B(i)$  的概率.

关于虚拟机器的正常与故障状态的定义(up and down):

$M_u(i)$  发生故障, 若对于某  $j \leq i$ , 有

1.  $M_j$  发生故障; 且
2.  $M_h$  正常工作, 当  $j < h \leq i$ ; 且
3.  $n_h = 0$ , 当  $j \leq h < i$ .



等价的定义:  $M_u(i)$  发生故障, 若

1.  $M_i$  发生故障; 或
2.  $n_{i-1} = 0$ , 且  $M_u(i-1)$  发生故障.

$M_u(i)$  正常工作, 若

1.  $M_i$  正常工作且  $n_{i-1} > 0$ ; 或
2.  $M_i$  正常工作,  $n_{i-1} = 0$ , 且  $M_u(i-1)$  正常工作.

$M_d(i)$  发生故障, 若

1.  $M_{i+1}$  发生故障; 或
2.  $n_{i+1} = N_{i+1}$ , 且  $M_d(i+1)$  发生故障.

$M_d(i)$  正常工作, 若

1.  $M_{i+1}$  正常工作且  $n_{i+1} < N_{i+1}$ ; 或
2.  $M_{i+1}$  正常工作,  $n_{i+1} = N_{i+1}$ , 且  $M_d(i+1)$  正常工作.

一共有  $6(k-1)$  个未知数;  $r_u(i), p_u(i), \mu_u(i), r_d(i), p_d(i), \mu_d(i), i = 1, \dots, k-1$ , 为此, 需构造  $6(k-1)$  个方程.

令  $\alpha_u(i; t)$  和  $\alpha_d(i; t)$  分别表示虚拟机器  $M_u(i)$  和  $M_d(i)$  在  $t$  时刻的状态, 1: 正常; 0: 故障.  $\alpha(i; t)$  表示机器  $M_i$  在  $t$  时刻的状态, 1: 正常; 0: 故障.

物流中断方程(Interruption of Flow):

$$\begin{aligned}
 p_u(i)\delta t &= \text{prob}[\alpha_u(i; t + \delta t) = 0 | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] \\
 &= \text{prob}[\alpha(i; t + \delta t) = 0 \text{ or } n_{i-1}(t + \delta t) = 0 \text{ and } \alpha_u(i-1; t + \delta t) = 0 \\
 &\quad | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] \\
 &= \text{prob}[\alpha(i; t + \delta t) = 0 | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i]
 \end{aligned}$$

$$+ \text{prob}[n_{i-1}(t + \delta t) = 0 \text{ and } \alpha_u(i-1; t + \delta t) = 0 | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] \quad (1)$$

上式中，第一项  $\text{prob}[\alpha(i; t + \delta t) = 0 | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] = p_i \delta t$ ，第二项

$$\begin{aligned} & \text{prob}[n_{i-1}(t + \delta t) = 0 \text{ and } \alpha_u(i-1; t + \delta t) = 0 | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] \\ = & \text{prob}[n_{i-1}(t + \delta t) = 0 \text{ and } \alpha_u(i-1; t + \delta t) = 0 \text{ and } \alpha_d(i-1; t + \delta t) = 1 | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] \\ = & \text{prob}[n_{i-1}(t) = 1 \text{ and } \alpha_u(i-1; t) = 0 \text{ and } \alpha_d(i-1; t) = 1 \\ & | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] \mu_d(i-1) \delta t \\ + & \text{prob}[n_{i-1}(t) = 0 \text{ and } \alpha_u(i-1; t) = 1 \text{ and } \alpha_d(i-1; t) = 1 \\ & | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] p_u(i-1) \delta t \\ + & \text{prob}[n_{i-1}(t) = 0 \text{ and } \alpha_u(i-1; t) = 0 \text{ and } \alpha_d(i-1; t) = 1 \\ & | \alpha_u(i; t) = 1 \text{ and } n_i(t) < N_i] [1 - r_u(i-1)] \delta t \end{aligned} \quad (2)$$

最后一项应为 0，因为若  $\alpha_u(i; t) = 1$ ，则  $n_{i-1}(t) > 0$  或  $n_{i-1}(t) = 0$  且  $\alpha_u(i-1; t) = 1$  二者必有一成立，而最后一项却是  $n_{i-1}(t) = 0$  且  $\alpha_u(i-1; t) = 0$ 。式(2)的第一项等于

$$\frac{1}{E_u(i)} \text{prob}[n_{i-1}(t) = 1 \text{ and } \alpha_u(i-1; t) = 0 \text{ and } \alpha_d(i-1; t) = 1] \mu_d(i-1) \delta t \rightarrow$$

$$\frac{p(i-1; 101)}{E_u(i)} \mu_d(i-1) \delta t, \text{ 这里 } p(i-1; 101) \text{ 是短生产线 } L(i-1) \text{ 处于状态 } (1, 0, 1) \text{ 的稳态}$$

概率；

$$E_u(i) = \text{prob}[\alpha_u(i) = 1 \text{ and } n_i < N_i] = e_u(i) [1 - p_b(i)] = \frac{r_u(i)}{p_u(i) + r_u(i)} [1 - \text{prob}(n_i = N_i)]$$

类似地，式(2)的第二项等于

$$\frac{p(i-1; 011)}{E_u(i)} p_u(i-1) \delta t.$$

于是可得式(1)的第二项等于

$$\begin{aligned} & \frac{p(i-1; 101)}{E_u(i)} \mu_d(i-1) \delta t + \frac{p(i-1; 011)}{E_u(i)} p_u(i-1) \delta t \\ = & \frac{1}{E_u(i)} [p(i-1; 101) \mu_d(i-1) + p(i-1; 011) p_u(i-1)] \delta t = \frac{1}{E_u(i)} p(i-1; 001) r_u(i-1) \delta t \end{aligned}$$

代入式(1)可得

$$p_u(i) = p_i + \frac{r_u(i-1) p(i-1; 001)}{E_u(i)}, \quad i = 2, \dots, k-1. \quad (3)$$

类似地，可得

$$p_d(i) = p_{i+1} + \frac{r_d(i+1) p(i+1; N10)}{E_d(i)}, \quad i = 1, \dots, k-2. \quad (4)$$

$$E_d(i) = \text{prob}[\alpha_d(i) = 1 \text{ and } n_i > 0] = e_d(i) [1 - p_s(i)] = \frac{r_d(i)}{p_d(i) + r_d(i)} [1 - \text{prob}(n_i = 0)]$$

物流恢复方程(Resumption of Flow):

$$r_u(i)\delta t = A(i-1)X(i) + B(i)[1 - X(i)] \quad (5)$$

其中  $X(i)$  是由  $M_u(i-1)$  发生故障而导致  $M_u(i)$  发生故障的概率,  $A(i-1)$  是  $M_u(i-1)$  修复的概率, 显然  $A(i-1) = r_u(i-1)\delta t$ ;  $1 - X(i)$  是由  $M_i$  发生故障而导致  $M_u(i)$  发生故障的概率,  $B(i)$  是  $M_i$  修复的概率, 显然  $B(i) = r_i\delta t$ .

而  $X(i) = \text{prob}[n_{i-1}(t) = 0 \text{ and } \alpha_u(i-1;t) = 0 \mid \alpha_u(i;t) = 0 \text{ and } n_i(t) < N_i]$

$$X(i) = \frac{\text{prob}[\{n_{i-1}(t) = 0 \text{ and } \alpha_u(i-1;t) = 0\} \text{ and } \alpha_u(i;t) = 0 \text{ and } n_i(t) < N_i]}{\text{prob}[\alpha_u(i;t) \text{ and } n_i(t) < N_i]} \quad (6)$$

式(6)中分子为短生产线  $L(i-1)$  处于状态  $(0,0,1)$  的概率, 即  $p(i-1;001)$ . 由于  $r_u(i)\text{prob}[\alpha_u(i;t) = 0 \text{ and } n_i(t) < N_i] = p_u(i)\text{prob}[\alpha_u(i;t) = 1 \text{ and } n_i(t) < N_i]$ , 所以  $\text{prob}[\alpha_u(i;t) = 0 \text{ and } n_i(t) < N_i] = \frac{p_u(i)}{r_u(i)}\text{prob}[\alpha_u(i;t) = 1 \text{ and } n_i(t) < N_i] = \frac{p_u(i)}{r_u(i)}E_u(i)$

于是

$$X(i) = \frac{p(i-1;001)r_u(i)}{p_u(i)E_u(i)} \quad (7)$$

另外,  $A(i-1) = r_u(i-1)\delta t$ ,  $B(i) = r_i\delta t$ , 代入式(6), 可得

$$r_u(i) = r_u(i-1)X(i) + r_i(1 - X(i)), \quad i = 2, \dots, k-1 \quad (8)$$

类似地可得

$$r_d(i) = r_d(i+1)Y(i) + r_{i+1}(1 - Y(i)), \quad i = 1, \dots, k-2 \quad (9)$$

$$Y(i) = \frac{p(i+1;N10)r_d(i)}{p_d(i)E_d(i)}. \quad (10)$$

物流守恒方程(Conseration of Flow):

所有短生产线  $L(i)$  的生产率都是相同的

$$P_i = P_k = P(i) = P_u(i) = P_d(i), \quad i = 1, \dots, k-1. \quad (11)$$

生产率与空闲时间之间的关系(Flow Rate - Idle Time Relationship):

$$P_i = \mu_i e_i (1 - \text{prob}[n_{i-1} = 0] - \text{prob}[n_i = N_i]) = \mu_i e_i (1 - p_s(i-1) - p_b(i)) \quad (12)$$

这里,  $e_i = \frac{r_i}{r_i + p_i}$ ,  $p_s(i-1) = \text{prob}[n_{i-1} = 0]$  为短生产线  $L(i-1)$  发生缺货的概率,

$p_b(i) = \text{prob}[n_i = N_i]$  为短生产线  $L(i)$  发生阻塞的概率. 对于两台机器构成的短生产线, 其生产率与空闲时间之间的关系如下:

$$P_d(i-1) = \mu_d(i-1)e_d(i-1)[1 - p_s(i-1)] = \mu_d(i-1)E_d(i-1),$$

$$P_u(i) = \mu_u(i)e_u(i)[1 - p_b(i)] = \mu_u(i)E_u(i).$$

把上两式代入式(12), 可得

$$\frac{1}{e_i \mu_i} + \frac{1}{P} = \frac{1}{e_d(i-1)\mu_d(i-1)} + \frac{1}{e_u(i)\mu_u(i)}, \quad i = 2, \dots, k-1. \quad (13)$$

边界条件(Boundary Conditions)

$$r_u(1) = r_1, \quad p_u(1) = p_1, \quad \mu_u(1) = \mu_1, \quad r_d(k-1) = r_k, \quad p_d(k-1) = p_k, \quad \mu_d(k-1) = \mu_k.$$

Dallery-David-Xie 算法 (DDX Algorithm):

1. 初始化. 对于每个两机器短生产线  $L(i)$ , 猜测其参数的初始值:

$$p_u(i) = p_i, \quad r_u(i) = r_i, \quad \mu_u(i) = \mu_i \quad \text{for } i = 1, \dots, k-1;$$

$$p_d(i) = p_{i+1}, \quad r_d(i) = r_{i+1}, \quad \mu_d(i) = \mu_{i+1} \quad \text{for } i = 1, \dots, k-1.$$

2. 交替执行下列步骤 1 和步骤 2, 直至终止条件满足.

(a) 步骤 1. 令  $i$  从 2 依次变化到  $k-1$ , 对于两机器短生产线  $L(i-1)$ , 用最近获得的  $r_u(i-1), p_u(i-1), \mu_u(i-1), r_d(i-1), p_d(i-1), \mu_d(i-1)$ , 计算  $P(i-1)$  和  $p(i-1;001)$ . 用前一次迭代获得的  $r_u(i), p_u(i), \mu_u(i), r_d(i), p_d(i), \mu_d(i)$  的值, 计算本次迭代中的  $L(i)$  的参数  $r_u(i), p_u(i), \mu_u(i)$ :

$$p_u(i) = p_i + \frac{r_u(i-1)p(i-1;001)\mu_u(i)}{P(i-1)}$$

$$X(i) = \frac{\mu_u(i)p(i-1;001)r_u(i)}{P(i-1)p_u(i)}$$

$$r_u(i) = r_u(i-1)X(i) + r_i(1-X(i))$$

$$e_u(i) = \frac{r_u(i)}{r_u(i) + p_u(i)}$$

$$\mu_u(i) = \frac{1}{e_u(i)} \left\{ \frac{1}{\frac{1}{P(i-1)} + \frac{1}{e_i\mu_i} - \frac{1}{e_d(i-1)\mu_d(i-1)}} \right\}$$

(b) 步骤 2. 令  $i$  从  $k-2$  依次变化到 1, 对于两机器短生产线  $L(i+1)$ , 用最近获得的  $r_u(i+1), p_u(i+1), \mu_u(i+1), r_d(i+1), p_d(i+1), \mu_d(i+1)$ , 计算  $P(i+1)$  和  $p(i+1;N10)$ . 用前一次迭代获得的  $r_u(i), p_u(i), \mu_u(i), r_d(i), p_d(i), \mu_d(i)$ , 计算本次迭代中的  $L(i)$  的参数  $r_d(i), p_d(i), \mu_d(i)$

$$p_d(i) = p_{i+1} + \frac{r_d(i+1)p(i+1;N10)\mu_d(i+1)}{P(i+1)}$$

$$Y(i) = \frac{\mu_d(i)p(i+1;N10)r_d(i)}{P(i+1)p_d(i)}$$

$$r_d(i) = r_d(i+1)Y(i) + r_{i+1}(1-Y(i))$$

$$e_d(i) = \frac{r_d(i)}{r_d(i) + p_d(i)}$$

$$\mu_d(i) = \frac{1}{e_d(i)} \left\{ \frac{1}{\frac{1}{P(i+1)} + \frac{1}{e_{i+1}\mu_{i+1}} - \frac{1}{e_u(i+1)\mu_u(i+1)}} \right\}$$

3. 终止条件:

$$\max\{|P(i) - P(1)|, i = 2, \dots, k-1\} < \varepsilon$$

$\varepsilon$  为一给定正数.

# 第 11 章

## 基于连续系统模型的制造系统分析与控制(1)

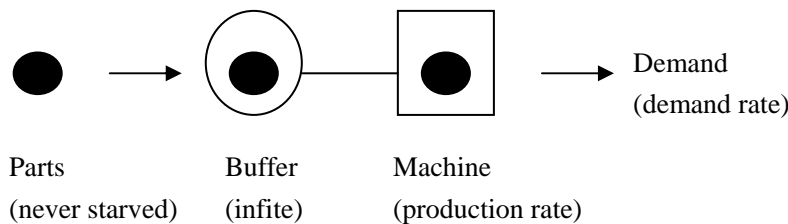
本章介绍基于连续系统模型的制造系统分析与控制，重点在控制。本章对制造系统控制问题的讨论基于如下假设：

1. 把制造系统看作一个连续变量动态系统；
2. 考虑随机因素，例如机器的故障或需求的不确定性；
3. 生产能力是有限制的。

控制的目标是使生产成本最小化，控制策略是状态反馈策略。本章将从最简单的单机器单品种制造系统出发，逐步深入研究基于连续模型的制造系统控制问题。

### § 11.1 单机器单品种制造系统的最优控制

设有如下的单机器单品种(single machine and single part-type)制造系统：



假设不会发生缺货；缓冲区容量无限；机器在正常状态下的最大生产率是  $\mu$ ，为一常数；需求率为  $d$ ，也为一常数；机器可能发生故障，以  $\alpha(t)$  表示  $t$  时刻机器的状态， $\alpha(t) = 1$  表示机器是正常的， $\alpha(t) = 0$  表示机器发生故障，单位时间内的故障率为  $p$ ，修复率为  $q$ ；以  $x(t)$  表示  $t$  时刻系统内的在制品状态，若  $x(t) \geq 0$ ，则  $x(t)$  表示实际的在制品数量，若  $x(t) < 0$ ，则  $x(t)$  表示欠产数量；以  $u(t)$  表示  $t$  时刻机器的实际生产率，显然  $u(t) \leq \mu$ 。对于这样一个简单的制造系统，可以容易地写出它作为一个连续系统的动态方程：

$$\frac{dx(t)}{dt} = u(t) - d. \tag{1}$$

其中，控制变量  $u(t)$  是有约束的：

$$u(t) \leq \alpha(t)\mu. \tag{2}$$

现在考虑控制的目标：最小化生产成本。这里仅考虑两类成本：在制品库存成本和欠产惩罚。以  $c^+$  表示单位时间单位数量的在制品的成本，以  $c^-$  表示单位时间单位数量欠产惩罚，另  $x^+(t) = \max\{x(t), 0\}$  表示实际在制品数量，以  $x^-(t) = \max\{-x(t), 0\}$  表示实际欠产数量。于是在  $[0, T]$  这段时间内的生产成本的均值为：

$$\frac{1}{T} E \int_0^T [c^+ x^+(t) + c^- x^-(t)] dt \tag{3}$$

于是这样一类单机器单品种制造系统的最优控制问题可以表述为：

$$\begin{aligned} \min \quad & \lim_{T \rightarrow \infty} \frac{1}{T} E \int_0^T [c^+ x^+(t) + c^- x^-(t)] dt \\ \text{s.t.} \quad & \frac{dx(t)}{dt} = u(t) - d \\ & u(t) \leq \alpha(t) \mu \\ & u(t) \geq 0 \end{aligned}$$

可以用动态规划法来解此最优控制问题. 动态规划的最优原理:

对于某个系统的最优控制问题, 假定该系统的初始状态是  $x(t_0)$ , 在一最优控制  $u(t)$  下在  $t_1$  时刻经过状态  $x(t_1)$ . 则对于以  $x(t_1)$  为初始条件的该系统而言,  $u(t)$  仍然是最优控制.

定义下述最优价值函数(或称未来成本函数, cost-to-go function):

$$J(x(t), \alpha(t), t) = \min_u E \left\{ \int_t^T g(x(s)) ds \middle| x(t), \alpha(t) \right\}. \quad (4)$$

这里,  $g(x(s)) = c^+ x^+(s) + c^- x^-(s)$ , 最小化是针对  $u(s)$  且  $s \in [t, T]$  而言的, 并且满足  $u(s) \in \Omega(\alpha(s))$ ,  $\Omega(\alpha(t)) = \{u(s) | u(s) \leq \alpha(s) \mu, u(s) \geq 0\}$  为  $u(s)$  的可行解空间.

根据最优原理, 对于一个很短的时间推移  $\delta t$ , 有

$$J(x(t), \alpha(t), t) = \min_u E \left\{ \int_t^{t+\delta t} g(x(s)) ds + J(x(t+\delta t), \alpha(t+\delta t), t+\delta t) \middle| x(t), \alpha(t) \right\} \quad (5)$$

注意, 这时最小化是针对  $u(s)$  且  $s \in [t, t+\delta t]$  而言的, 并且满足  $u(s) \in \Omega(\alpha(s))$ , 于是上式成为

$$\begin{aligned} J(x(t), \alpha(t), t) = \min_u E \left\{ g(x(t)) \delta t + J(x(t), \alpha(t), t) + \sum_j J(x(t), j, t) \lambda_{j\alpha} \delta t \right. \\ \left. + \frac{\partial J}{\partial x}(x(t), \alpha(t), t) \delta x(t) + \frac{\partial J}{\partial t}(x(t), \alpha(t), t) \delta t \right\} + o(\delta t) \end{aligned}$$

这里  $\lambda_{j\alpha}$  是机器状态从  $j$  到  $\alpha$  的转移概率,  $\lambda_{j\alpha} \delta t = P\{\alpha(t+\delta t) = \alpha | \alpha(t) = j\}$ . 对于单机器系统有  $\lambda_{10} \delta t = p \delta t$ ,  $\lambda_{01} \delta t = q \delta t$ ,  $\lambda_{11} \delta t = 1 - p \delta t$ ,  $\lambda_{00} \delta t = 1 - q \delta t$ . 在上式中,

$\delta x(t) = \frac{dx(t)}{dt} \delta t = (u(t) - d) \delta t$  成立. 在上式两侧同时除以  $\delta t$  并令  $\delta t \rightarrow 0$ , 于是有

$$-\frac{\partial J}{\partial t}(x, \alpha, t) = \min_{u \in \Omega(\alpha)} E \left\{ g(x(t)) + \sum_j J(x(t), j, t) \lambda_{j\alpha} + \frac{\partial J}{\partial x}(x, \alpha, t)(u - d) \right\} \quad (6)$$

式(6)叫做“Bellman 方程”或 Hamilton-Jacob-Bellman (HJB) 方程. 从上式可见, 只有最后一项与控制变量  $u$  有关. 于是最优控制问题实际上转化为

$$\min_{u \in \Omega(\alpha)} \frac{\partial J}{\partial x}(x, \alpha, t)(u - d) \quad (7)$$

注意:  $\frac{\partial J}{\partial x}(x, \alpha, t)$  实际上是已经确定的, 尽管我们未必知道其具体值. 于是, 为了

最小化  $\frac{\partial J}{\partial x}(x, \alpha, t)(u - d)$ , 当  $\alpha = 1$ , 即机器正常工作时, 可采取如下的控制:

若  $\frac{\partial J}{\partial x}(x, \alpha, t) > 0$ , 则  $u$  应尽可能小, 即  $u = 0$ ;



若  $\frac{\partial J}{\partial x}(x, \alpha, t) < 0$ , 则  $u$  应尽可能大, 即  $u = \mu$ ;

若  $\frac{\partial J}{\partial x}(x, \alpha, t) = 0$ , 则保持生产率与需求率一致, 即  $u = d$ ;

当  $\alpha = 0$ , 即机器发生故障时, 显然  $u = 0$ .

这是一种根据不同状态来切换控制变量值的控制策略, 称为 **bang-bang** 控制. 如果 **cost-to-go** 函数具有下凸的形式, 即存在一极小值点  $Z$ , 当  $x < Z$  时,  $\frac{\partial J}{\partial x}(x, \alpha, t) < 0$ ; 当  $x > Z$  时,  $\frac{\partial J}{\partial x}(x, \alpha, t) > 0$ , 则上述控制策略具有如下的形式:

$$u = \begin{cases} 0, & \text{if } x > Z \text{ and } \alpha = 1 \\ d, & \text{if } x = Z \text{ and } \alpha = 1 \\ \mu, & \text{if } x < Z \text{ and } \alpha = 1 \\ 0, & \text{if } \alpha = 0 \end{cases} \quad (8)$$

式(8)称为单机器单品种制造系统的 **Hedging Point** 控制,  $Z$  即为 **Hedging Point**. 从上式可以看出, **Hedging Point** 控制是一种状态反馈控制, 具有简单明确的形式, 因而很容易实现. 控制参数只有一个, 即 **Hedging Point**  $Z$ , 剩下的问题是: 如何确定  $Z$  的值?

由系统的动态方程和控制策略可知, 在稳态(steady state)下, 系统的在制品状态只能是  $x = Z$  或  $x \in (-\infty, Z)$  而系统的机器状态是  $\alpha = 1$  或  $\alpha = 0$ . 因此, 可用  $(x, \alpha)$  来表示系统的整个状态. 由于机器状态  $\alpha$  为一随机变量, 因此需要分析系统状态  $(x, \alpha)$  的概率分布. 令  $f(x, \alpha)$  表示  $x \in (-\infty, Z)$  时系统状态的概率分布密度函数, 于是有

$$\begin{aligned} f(x, 1) &= f(x - (\mu - d)\delta t, 1)(1 - p\delta t) + f(x + d\delta t, 0)q\delta t + o(\delta t) \\ &= \left[ f(x, 1) - \frac{\partial f(x, 1)}{\partial x}(\mu - d)\delta t + o(\delta t) \right] (1 - p\delta t) + \left[ f(x, 0) + \frac{\partial f(x, 0)}{\partial x}d\delta t + o(\delta t) \right] q\delta t + o(\delta t) \\ &= f(x, 1) - f(x, 1)p\delta t - \frac{\partial f(x, 1)}{\partial x}(\mu - d)\delta t + f(x, 0)q\delta t + o(\delta t) \end{aligned}$$

在上式两侧同时除以  $\delta t$  并令  $\delta t \rightarrow 0$ , 可得

$$f(x, 1)p + \frac{\partial f(x, 1)}{\partial x}(\mu - d) = f(x, 0)q \quad (9)$$

类似地

$$\begin{aligned} f(x, 0) &= f(x + d\delta t, 0)(1 - q\delta t) + f(x - (\mu - d)\delta t, 1)p\delta t + o(\delta t) \\ &= \left[ f(x, 0) + \frac{\partial f(x, 0)}{\partial x}d\delta t + o(\delta t) \right] (1 - q\delta t) + \left[ f(x, 1) - \frac{\partial f(x, 1)}{\partial x}(\mu - d)\delta t + o(\delta t) \right] p\delta t + o(\delta t) \\ &= f(x, 0) - f(x, 0)q\delta t + \frac{\partial f(x, 0)}{\partial x}d\delta t + f(x, 1)p\delta t + o(\delta t) \end{aligned}$$

在上式两侧同时除以  $\delta t$  并令  $\delta t \rightarrow 0$ , 可得

$$f(x, 0)q - \frac{\partial f(x, 0)}{\partial x}d = f(x, 1)p \quad (10)$$

实际上, 式(9)和(10)是系统稳态概率平衡方程.

假定稳态概率分布具有如下的形式:

$$f(x, \alpha) = A_\alpha \exp\{ax\} \quad (11)$$

代入式(9)和(10), 可得

$$pA_1 \exp\{ax\} + (\mu - d)A_1 a \exp\{ax\} = qA_0 \exp\{ax\}$$

$$qA_0 \exp\{ax\} - dA_0 a \exp\{ax\} = pA_1 \exp\{ax\}$$

即

$$pA_1 + (\mu - d)A_1 a = qA_0 \quad (12)$$

$$qA_0 - dA_0 a = pA_1 \quad (13)$$

解上述两个方程, 可得

$$A_1 = \frac{d}{\mu - d} A_0 \quad (14)$$

$$a = \frac{q}{d} - \frac{p}{\mu - d} \quad (15)$$

注意: 根据  $\mu \frac{q}{p+q} > d$  可得  $a > 0$ . 根据式(11), 系统状态在  $x \in (-\infty, Z)$  时的稳态

概率分布密度函数的具体形式就可以获得了, 只是参数  $A_0$  待定.

对于在制品状态  $x = Z$ , 令  $P(Z, \alpha)$  表示系统状态处于  $(Z, \alpha)$  的概率, 于是

$$\begin{aligned} P(Z, 1) &= P(Z, 1)(1 - p\delta t) + (1 - p\delta t) \int_{Z-(\mu-d)\delta t}^Z f(x, 1) dx + o(\delta t) \\ &= P(Z, 1) - P(Z, 1)p\delta t + f(Z, 1)(\mu - d)\delta t + o(\delta t) \rightarrow \\ pP(Z, 1) &= f(Z, 1)(\mu - d) \quad (\text{令 } \delta t \rightarrow 0) \rightarrow \\ P(Z, 1) &= \frac{\mu - d}{p} f(Z, 1) = \frac{\mu - d}{p} A_1 \exp\{aZ\} = \frac{d}{p} A_0 \exp\{aZ\}. \end{aligned} \quad (16)$$

对于状态  $(Z, 0)$ , 由于此时的控制  $u = 0$ , 所以系统状态  $(x, \alpha)$  一旦达到  $(Z, 0)$  就会立即离开, 也就是说  $(Z, 0)$  不是一个稳态, 因此  $P(Z, 0) = 0$ .

根据概率分布的归一化(normalization)原则, 下式必须成立:

$$P(Z, 1) + P(Z, 0) + \int_{-\infty}^Z [f(x, 1) + f(x, 0)] dx = 1 \rightarrow \quad (17)$$

$$\frac{d}{p} A_0 \exp\{aZ\} + \int_{-\infty}^Z (A_1 + A_0) \exp\{ax\} dx = 1 \rightarrow$$

$$\frac{d}{p} A_0 \exp\{aZ\} + \frac{A_1 + A_0}{a} \exp\{aZ\} = 1 \quad (\text{把式(14)代入}) \rightarrow$$

$$\left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right] A_0 \exp\{aZ\} = 1 \rightarrow$$

$$A_0 = \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} \exp\{-aZ\} \quad (18)$$

目标函数(3)实际上等价于稳态时生产成本的均值, 即  $Eg(x) + Eg(Z)$ ,  $x \in (-\infty, Z)$ . 首先考虑  $Z \geq 0$  的情况. 此时

$$\begin{aligned} Eg(x) &= \int_{-\infty}^Z g(x) [f(x, 1) + f(x, 0)] dx = \int_{-\infty}^Z (c^+ x^+ + c^- x^-) [f(x, 1) + f(x, 0)] dx \\ &= \int_{-\infty}^0 (-c^- x) [f(x, 1) + f(x, 0)] dx + \int_0^Z (c^+ x) [f(x, 1) + f(x, 0)] dx \\ &= -c^- (A_1 + A_0) \int_{-\infty}^0 x \exp\{ax\} dx + c^+ (A_1 + A_0) \int_0^Z x \exp\{ax\} dx \end{aligned}$$

$$= \left[ \frac{c^+ Z \exp\{aZ\}}{a} - \frac{c^+ \exp\{aZ\}}{a^2} + \frac{c^+ + c^-}{a^2} \right] \frac{\mu}{\mu - d} A_0. \quad (19)$$

$$Eg(Z) = c^+ ZP(Z,1) = \frac{c^+ d}{p} A_0 Z \exp\{aZ\}. \quad (20)$$

于是, 生产成本均值

$$\begin{aligned} C(Z) &= Eg(x) + Eg(Z) \\ &= \left[ \frac{c^+ Z \exp\{aZ\}}{a} - \frac{c^+ \exp\{aZ\}}{a^2} + \frac{c^+ + c^-}{a^2} \right] \frac{\mu}{\mu - d} A_0 + \frac{c^+ d}{p} A_0 Z \exp\{aZ\} \\ &= \frac{(c^+ + c^-)\mu}{a^2(\mu - d)} \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} \exp\{-aZ\} + c^+ Z - \frac{c^+ \mu}{a^2(\mu - d)} \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} \end{aligned} \quad (21)$$

令  $\frac{dC(Z)}{dZ} = 0$ , 可得

$$Z = \frac{1}{a} \ln \left\{ \frac{(c^+ + c^-)p\mu}{c^+ [p\mu + ad(\mu - d)]} \right\}$$

把式(15)代入上式, 可得

$$Z = \frac{d(\mu - d)}{q(\mu - d) - pd} \ln \left\{ \frac{(c^+ + c^-)p\mu}{c^+ (p + q)(\mu - d)} \right\} \quad (22)$$

要保证  $Z > 0$ , 必须  $\frac{(c^+ + c^-)p\mu}{c^+ (p + q)(\mu - d)} > 1$  成立.

又  $\frac{d^2 C(Z)}{dZ^2} = \frac{(c^+ + c^-)\mu}{(\mu - d)} \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} \exp\{-aZ\}$ , 由于

$$\frac{d}{p} + \frac{\mu}{a(\mu - d)} = \frac{p + q}{p \left[ \frac{q}{d} - \frac{p}{\mu - d} \right]} = \frac{p + q}{ap} > 0$$

因此, 式(22)为生产成本的极小值点, 即 Hedging Point 的最优解.

特别地, 当  $\frac{(c^+ + c^-)p\mu}{c^+ (p + q)(\mu - d)} = 1$  时, Hedging Point 的最优解为  $Z = 0$ .

再考虑  $Z < 0$  的情况, 此时

$$\begin{aligned} Eg(x) &= \int_{-\infty}^Z g(x)[f(x,1) + f(x,0)]dx = \int_{-\infty}^Z (c^+ x^+ + c^- x^-)[f(x,1) + f(x,0)]dx \\ &= \int_{-\infty}^Z (-c^- x)[f(x,1) + f(x,0)]dx = -c^- (A_1 + A_0) \int_{-\infty}^Z x \exp\{ax\}dx \\ &= \left( \frac{1}{a^2} - \frac{Z}{a} \right) \frac{c^- \mu}{\mu - d} \exp\{aZ\} A_0. \end{aligned}$$

$$Eg(Z) = -c^- ZP(Z,1) = -\frac{c^- d}{p} A_0 Z \exp\{aZ\}$$

$$A_0 = \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} \exp\{-aZ\}$$

$$C(Z) = Eg(x) + Eg(Z) = \left( \frac{1}{a^2} - \frac{Z}{a} \right) \frac{c^- \mu}{\mu - d} \exp\{aZ\} A_0 - \frac{c^- d}{p} A_0 Z \exp\{aZ\}$$

$$\begin{aligned}
&= \frac{1}{a^2} \frac{c^- \mu}{\mu - d} \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} - c^- \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right] \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} Z \\
&= \frac{1}{a^2} \frac{c^- \mu}{\mu - d} \left[ \frac{d}{p} + \frac{\mu}{a(\mu - d)} \right]^{-1} - c^- Z
\end{aligned}$$

因此  $\frac{dC(Z)}{dZ} = -c^-$ . 即当  $Z < 0$ , 生产成本函数  $C(Z)$  关于  $Z$  是单调下降的. 于是有

$C(Z) > C(0)$ , 当  $Z < 0$  时.

总结: 对于单机器单零件品种的制造系统, 在满足  $\mu q / (p + q) > d$  的条件下,

1. Hedging Point 的最优值不可能小于 0;

2. 当  $\frac{(c^+ + c^-) p \mu}{c^+ (p + q)(\mu - d)} \leq 1$  时, Hedging Point 的最优解为  $Z = 0$ ;

3. 当  $\frac{(c^+ + c^-) p \mu}{c^+ (p + q)(\mu - d)} > 1$  时, Hedging Point 的最优解为

$$Z = \frac{d(\mu - d)}{q(\mu - d) - pd} \ln \left\{ \frac{(c^+ + c^-) p \mu}{c^+ (p + q)(\mu - d)} \right\}$$

## § 11.2 单机器多品种制造系统的最优控制

设有一单机器多品种(single machine and multiple part-types)制造系统, 生产  $n$  种零件. 对于第  $i$  种零件, 若机器只生产这一种零件, 则在正常状态下的最大生产率是  $\mu_i$ , 为一常数. 对于第  $i$  种零件的需求率为  $d_i$ , 也为一常数. 以  $x_i(t)$  表示  $t$  时刻系统内第  $i$  种零件的在制品状态,  $x_i(t) \geq 0$  表示实际的在制品数量,  $x_i(t) < 0$  表示欠产数量. 以  $u_i(t)$  表示  $t$  时刻第  $i$  种零件的实际生产率, 当然  $u_i(t) \leq \mu_i$ . 这样一个系统的动态方程为:

$$\frac{dx_i(t)}{dt} = u_i(t) - d_i, \quad i = 1, \dots, n. \quad (1)$$

这里  $n$  表示产品品种数量. 控制变量  $u_i(t)$  满足如下约束:

$$\sum_{i=1}^n \frac{u_i(t)}{\mu_i} \leq \alpha(t). \quad (2)$$

$\alpha(t) \in \{0, 1\}$  表示机器的状态, 为一随机变量.

以  $c_i^+$  表示单位时间单位数量第  $i$  种零件的在制品的成本, 以  $c_i^-$  表示单位时间单位数量第  $i$  种零件的欠产惩罚, 另  $x_i^+(t) = \max\{x_i(t), 0\}$  表示实际的第  $i$  种零件在制品数量, 以  $x_i^-(t) = \max\{-x_i(t), 0\}$  表示实际的第  $i$  种零件欠产数量. 于是在  $[0, T]$  这段时间内的生产成本均值为:

$$\frac{1}{T} E \int_0^T \sum_{i=1}^n [c_i^+ x_i^+(t) + c_i^- x_i^-(t)] dt \quad (3)$$

于是单机器多品种制造系统的最优控制问题可以表述为:

$$\begin{aligned}
& \min \lim_{T \rightarrow \infty} \frac{1}{T} E \int_0^T \sum_{i=1}^n [c_i^+ x_i^+(t) + c_i^- x_i^-(t)] dt \\
& \text{s.t. } \frac{dx_i(t)}{dt} = u_i(t) - d_i, \quad i = 1, \dots, n \\
& \sum_{i=1}^n \frac{u_i(t)}{\mu_i} \leq \alpha(t) \\
& u_i(t) \geq 0, \quad i = 1, \dots, n
\end{aligned}$$

对于单机器多品种制造系统, 一种在工程上便于实现的控制是具有优先级的 Hedging Point (Prioritized Hedging Point, PHP) 控制: 为系统中所有  $n$  种零件设定优先级, 不失一般性, 假定第 1 种零件优先级最高, 第 2 种零件次之, ..., 第  $n$  种零件的优先级最低. 零件的优先级可根据其各自的需求率由大到小的顺序排定, 即需求率最大者优先级最高. 令  $Z_i (i=1, \dots, n)$  表示处于第  $i$  种的零件生产剩余量的 Hedging Point, 于是有

$$u_1(t) = \begin{cases} \mu_1, & \text{if } \alpha = 1 \text{ and } x_1(t) < Z_1 \\ d_1, & \text{if } \alpha = 1 \text{ and } x_1(t) = Z_1 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$u_i(t) = \begin{cases} \left(1 - \sum_{k=1}^{i-1} \frac{d_k}{\mu_k}\right) \mu_i, & \text{if } \alpha = 1 \text{ and } x_k = Z_k \quad \forall k = 1, \dots, i-1, \\ & \text{and } x_i < Z_i; \\ d_i, & \text{if } \alpha = 1 \text{ and } x_k = Z_k \quad \forall k = 1, \dots, i-1, \quad \forall i = 2, \dots, n \\ & \text{and } x_i = Z_i; \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

式(4)和(5)意味着对于在这一机器上加工的所有零件, 当机器正常工作时, 先给优先级最高的零件分配生产能力, 若其在制品数量小于其 Hedging Point 值, 则把所有的生产能力分配给它, 即生产率尽可能大; 若其在制品数量等于其 Hedging Point 值, 则生产率等于其需求率. 对于某个优先级不是最高的零件, 若所有优先级高于它的零件的在制品数量都达到各自的 Hedging Point, 且该种零件的在制品数量小于其 Hedging Point, 则应把机器除为维持优先级更高的零件数量处于 Hedging Point 而投入的生产能力之外剩余的生产能力全部分配给该种零件; 若所有优先级高于它的零件的在制品数量都达到各自的 Hedging Point, 且该种零件的在制品数量也达到 Hedging Point, 则该种零件的生产率应等于其需求率, 以使得在制品数量维持在 Hedging Point; 在其余情形下, 生产率为零.

### § 11.3 多机器单品种制造系统的最优控制

设有一串行生产线, 由  $m$  台机器构成. 该系统只生产一种零件, 每道工序在一台机器上加工, 依次经过这  $m$  台机器. 令  $x_i(t)$  表示  $t$  时刻第  $i$  台机器上该种零件的在制品状态,  $i = 1, \dots, m$ ;  $\alpha_i(t) = 1$  或  $0$  分别表示  $t$  时刻机器  $i$  处于正常运行或发生故障状态,  $i = 1, \dots, m$ ;  $u_i(t)$  表示  $t$  时刻第  $i$  台机器的生产率,  $i = 1, \dots, m$ ;  $\mu_i$  表示

零件在机器  $i$  上的最大生产率,  $i=1, \dots, m$ ;  $d$  表示对零件的需求率; 机器  $i$  与  $i+1$  之间的缓冲区容量为  $N_i$ . 于是这一制造系统动态方程为

$$\frac{dx_i(t)}{dt} = u_i(t) - d, \quad i=1, \dots, m.$$

$x_i(t)$  又称生产剩余量(production surplus). 令  $g(x_1(t), \dots, x_m(t))$  表示此制造系统的性能指标, 则其最优控制问题可表述为

$$\min \lim_{T \rightarrow \infty} \frac{1}{T} E \left\{ \int_0^T g(x_1(t), \dots, x_m(t)) dt \mid x_1(0), \dots, x_m(0), \alpha_1(0), \dots, \alpha_m(0) \right\} \quad (6)$$

$$\text{s.t.} \quad \frac{dx_i(t)}{dt} = u_i(t) - d, \quad i=1, \dots, m. \quad (7)$$

$$\frac{u_i(t)}{\mu_i} \leq \alpha_i(t), \quad i=1, \dots, m. \quad (8)$$

与单机器的情形不同的是, 在串行生产线中, 除第一台机器外, 任何一台机器的生产率都受到前一台机器生产率的制约, 即当该机器与其前一台机器之间的在制品数量为零时, 其生产率不可能超过前一台机器的生产率; 同时也受到后一台机器的制约, 即当该机器与其后一台机器之间的在制品数量等于缓冲区容量时, 其生产率不可能超过后一台机器的生产率. 由于存在这样的制约, 单零件品种串行制造系统的 Hedging Point 控制常采取如下形式:

$$u_1(t) = \begin{cases} 0, & \text{if } x_1(t) > Z_1 \\ d, & \text{if } x_1(t) = Z_1 \text{ and } x_1(t) - x_2(t) < N_1 \\ \min\{d, u_2(t)\}, & \text{if } x_1(t) = Z_1 \text{ and } x_1(t) - x_2(t) = N_1. \\ \mu_1, & \text{if } x_1(t) < Z_1 \text{ and } x_1(t) - x_2(t) < N_1 \\ \min\{\mu_1, u_{i+1}(t)\}, & \text{if } x_1(t) < Z_1 \text{ and } x_1(t) - x_2(t) = N_1 \end{cases} \quad (9)$$

$$u_i(t) = \begin{cases} 0, & \text{if } x_i(t) > Z_i \\ d, & \text{if } x_i(t) = Z_i \text{ and } x_{i-1}(t) - x_i(t) > 0 \text{ and } x_i(t) - x_{i+1}(t) < N_i \\ \min\{d, u_{i-1}(t)\}, & \text{if } x_i(t) = Z_i \text{ and } x_{i-1}(t) - x_i(t) = 0 \text{ and } x_i(t) - x_{i+1}(t) < N_i \\ \min\{d, u_{i+1}(t)\}, & \text{if } x_i(t) = Z_i \text{ and } x_{i-1}(t) - x_i(t) > 0 \text{ and } x_i(t) - x_{i+1}(t) = N_i \\ \min\{d, u_{i-1}(t), u_{i+1}(t)\}, & \\ & \text{if } x_i(t) = Z_i \text{ and } x_{i-1}(t) - x_i(t) = 0 \text{ and } x_i(t) - x_{i+1}(t) = N_i, \\ \mu_i, & \text{if } x_i(t) < Z_i \text{ and } x_{i-1}(t) - x_i(t) > 0 \text{ and } x_i(t) - x_{i+1}(t) < N_i \\ \min\{\mu_i, u_{i-1}(t)\}, & \text{if } x_i(t) < Z_i \text{ and } x_{i-1}(t) - x_i(t) = 0 \text{ and } x_i(t) - x_{i+1}(t) < N_i \\ \min\{\mu_i, u_{i+1}(t)\}, & \text{if } x_i(t) < Z_i \text{ and } x_{i-1}(t) - x_i(t) > 0 \text{ and } x_i(t) - x_{i+1}(t) = N_i \\ \min\{\mu_i, u_{i-1}(t), u_{i+1}(t)\}, & \\ & \text{if } x_i(t) < Z_i \text{ and } x_{i-1}(t) - x_i(t) = 0 \text{ and } x_i(t) - x_{i+1}(t) = N_i \end{cases}, \quad (10)$$

$i = 2, \dots, m.$

## § 11.4 多机器多品种制造系统的最优控制

考虑一多机器多品种制造系统, 由  $m$  台机器构成, 生产  $n$  种零件. 令  $\alpha_i(t) = 1$  或  $0$  分别表示  $t$  时刻机器  $i$  处于正常运行或发生故障状态,  $i = 1, \dots, m$ ;  $x_j(t)$  表示  $t$  时刻第  $j$  种零件的生产剩余量(production surplus),  $j = 1, \dots, n$ ;  $u_j(t)$  表示  $t$  时刻第  $j$  种零件的生产率,  $j = 1, \dots, n$ ;  $\mu_{ij}$  表示零件  $j$  在机器  $i$  上的最大生产率,  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ ;  $d_j$  表示对零件  $j$  的需求率. 于是这一制造系统动态方程为

$$\frac{dx_j(t)}{dt} = u_j(t) - d_j, \quad j = 1, \dots, n. \quad (1)$$

令  $g(x_1(t), \dots, x_n(t))$  表示此制造系统的性能指标, 则其最优控制问题可表述为:

$$\min \lim_{T \rightarrow \infty} \frac{1}{T} E \left\{ \int_0^T g(x_1(t), \dots, x_n(t)) dt \middle| x_1(0), \dots, x_n(0), \alpha_1(0), \dots, \alpha_n(0) \right\} \quad (2)$$

$$\text{s.t.} \quad \frac{dx_j(t)}{dt} = u_j(t) - d_j, \quad j = 1, \dots, n, \quad (3)$$

$$\sum_{j=1}^n u_j(t) / \mu_{ij} \leq \alpha_i(t), \quad i = 1, \dots, m. \quad (4)$$

仍然用动态规划法解此问题. 定义下述 cost-to-go 函数:

$$J(\mathbf{x}(t), \boldsymbol{\alpha}(t), t) = \min_{\mathbf{u}} E \left\{ \int_t^T g(\mathbf{x}(s)) ds \middle| \mathbf{x}(t), \boldsymbol{\alpha}(t) \right\}. \quad (5)$$

这里,  $\mathbf{u}(t) = (u_1(t), \dots, u_n(t))$ ,  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$ ,  $g(\mathbf{x}(s)) = \sum_{j=1}^n [c_j^+ x_j^+(s) + c_j^- x_j^-(s)]$ .

根据动态规划最优原理, 对于一个很短的时间推移  $\delta t$ , 有

$$J(\mathbf{x}(t), \boldsymbol{\alpha}(t), t) = \min_{\mathbf{u}} E \left\{ \int_t^{t+\delta t} g(\mathbf{x}(s)) ds + J(\mathbf{x}(t+\delta t), \boldsymbol{\alpha}(t+\delta t), t+\delta t) \middle| \mathbf{x}(t), \boldsymbol{\alpha}(t) \right\} \quad (6)$$

注意, 这时最小化是针对  $\mathbf{u}(s)$  且  $s \in [t, t+\delta t]$  而言的, 并且满足  $\mathbf{u}(s) \in \Omega(\boldsymbol{\alpha}(s))$ , 这

里  $\Omega(\boldsymbol{\alpha}(s)) = \left\{ \mathbf{u}(s) \middle| \sum_{j=1}^m \frac{u_j(s)}{\mu_{ij}} \leq \alpha_i(s), u_j(s) \geq 0 \right\}$  为  $\mathbf{u}(s)$  的可行解空间. 于是上式成

为

$$J(\mathbf{x}(t), \boldsymbol{\alpha}(t), t) = \min_{\mathbf{u}} E \left\{ g(\mathbf{x}(t)) \delta t + J(\mathbf{x}(t), \boldsymbol{\alpha}(t), t) + \sum_{\beta} J(\mathbf{x}(t), \boldsymbol{\beta}, t) \lambda_{\beta\alpha} \delta t \right. \\ \left. + \sum_{j=1}^n \frac{\partial J}{\partial x_j}(\mathbf{x}(t), \boldsymbol{\alpha}(t), t) \delta x_j(t) + \frac{\partial J}{\partial t}(\mathbf{x}(t), \boldsymbol{\alpha}(t), t) \delta t \right\} + o(\delta t)$$

在上式中,  $\delta x_j(t) = (u_j(t) - d_j) \delta t$  成立, 令  $\delta t \rightarrow 0$ , 于是可得 Bellman 方程

$$-\frac{\partial J}{\partial t}(\mathbf{x}, \boldsymbol{\alpha}, t) = \min_{\mathbf{u} \in \Omega(\boldsymbol{\alpha})} E \left\{ g(\mathbf{x}(t)) + \sum_{\beta} J(\mathbf{x}(t), \boldsymbol{\beta}, t) \lambda_{\beta\alpha} + \sum_{j=1}^n \frac{\partial J}{\partial x_j}(\mathbf{x}, \boldsymbol{\alpha}, t) (u_j - d_j) \right\}. \quad (7)$$

于是最优控制问题实际上转化为线性规划(Linear Programming)问题

$$\min \sum_{j=1}^n \frac{\partial J}{\partial x_j}(\mathbf{x}, \mathbf{a}, t)[u_j(t) - d_j] \quad (8)$$

$$\text{s.t.} \quad \sum_{j=1}^n u_j(t) / \mu_{ij} \leq \alpha_i(t), \quad i = 1, \dots, m. \quad (9)$$

$$u_j(t) \geq 0, \quad j = 1, \dots, n$$

这一问题解起来有些困难, 原因在于  $J(\mathbf{x}, \mathbf{a}, t)$  虽然是确定的(与控制变量无关), 但我们并不知道其具体值. 为此, 假设 cost-to-go 函数具有正定二次型的形式

$$J(\mathbf{x}, \mathbf{a}, t) = \frac{1}{2} \mathbf{x}^T A(\mathbf{a}) \mathbf{x} - \mathbf{b}^T(\mathbf{a}) \mathbf{x} + C(\mathbf{a}). \quad (10)$$

这里,  $A(\mathbf{a}) = \text{diag}\{a_1(\mathbf{a}), \dots, a_n(\mathbf{a})\}$  为  $n$  阶正定对角阵;  $\mathbf{b} = [b_1(\mathbf{a}), \dots, b_n(\mathbf{a})]^T$  为  $n$  维列向量,  $C$  为标量, 于是有

$$\frac{\partial J}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}, t) = A\mathbf{x} - \mathbf{b}$$

令  $\frac{\partial J}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}, t) = 0$ , 可解得  $\mathbf{x} = A^{-1}\mathbf{b}$ .

同时又有  $\frac{\partial^2 J}{\partial \mathbf{x}^2}(\mathbf{x}, \mathbf{a}, t) = A$ , 由于  $A$  是正定的, 因此  $A^{-1}\mathbf{b}$  为  $J(\mathbf{x}, \mathbf{a}, t)$  的一个极小值点, 令  $Z = A^{-1}\mathbf{b}$ , 即  $(Z_1, \dots, Z_n) = (b_1/a_1, \dots, b_n/a_n)$ . 于是可得

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}, t) &= A(\mathbf{x} - A^{-1}\mathbf{b}) = [a_1(x_1 - Z_1), \dots, a_n(x_n - Z_n)], \text{ 即} \\ \frac{\partial J}{\partial x_j}(\mathbf{x}, \mathbf{a}, t) &= a_j(x_j - Z_j), \quad j = 1, \dots, n. \end{aligned} \quad (10)$$

代入式(8), 可得

$$\sum_{j=1}^n \frac{\partial J}{\partial x_j}(\mathbf{x}, \mathbf{a}, t)(u_j - d_j) = \sum_{j=1}^n a_j(x_j - Z_j)(u_j - d_j) = \sum_{j=1}^n a_j(x_j - Z_j)u_j - \sum_{i=1}^n a_j(x_j - Z_j)d_j$$

上式第二项与控制变量无关, 因此原来的线性规划的目标函数成为  $\sum_{j=1}^n a_j(x_j - Z_j)u_j$ , 则原来的线性规划问题成为

$$\min \sum_{j=1}^n a_j(x_j - Z_j)u_j \quad (11)$$

$$\text{s.t.} \quad \sum_{j=1}^n u_j / \mu_{ij} \leq \alpha_i, \quad i = 1, \dots, m. \quad (12)$$

$$u_j \geq 0, \quad j = 1, \dots, n \quad (13)$$

### 线性规划(Linear Programming):

目标函数是线性的, 约束条件也是线性的, 决策变量是非负的, 这样一类有约束的优化问题叫做线性规划问题. 注意, 线性规划问题存在不等式约束, 即决策变量的非负性约束, 这使得它不能用 Lagrange 乘子法解决.

线性规划问题的标准形式:



$$\begin{aligned} \min \quad & f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & D\mathbf{x} = \boldsymbol{\theta} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

系数矩阵  $D$  可以分解为两部分, 基矩阵  $D_B$  与非基矩阵  $D_N$ , 即  $D = (D_B, D_N)$  其中基矩阵是一个非奇异方阵, 其阶数等于系数矩阵  $D$  的行数, 实际上  $D_B$  是由  $D$  的若干线性无关的列向量构成的. 与这些列向量对应的决策变量  $\mathbf{x}$  的分量叫做基变量, 它们构成基向量  $\mathbf{x}_B$ , 其余的分量构成非基向量  $\mathbf{x}_N$ , 于是  $\mathbf{x} = (\mathbf{x}_B, \mathbf{x}_N)$ . 目标函数中与基向量和非基向量对应的成本系数向量可分别记为  $\mathbf{c}_B^T$  和  $\mathbf{c}_N^T$ , 于是  $\mathbf{c} = (\mathbf{c}_B^T, \mathbf{c}_N^T)$ . 因此上述线性规划问题可写作:

$$\begin{aligned} \min \quad & f(\mathbf{x}_B, \mathbf{x}_N) = \mathbf{c}_B^T \mathbf{x}_B + \mathbf{c}_N^T \mathbf{x}_N \\ \text{s.t.} \quad & D_B \mathbf{x}_B + D_N \mathbf{x}_N = \boldsymbol{\theta} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

由第一个约束条件可得  $\mathbf{x}_B = D_B^{-1}\boldsymbol{\theta} - D_B^{-1}D_N\mathbf{x}_N$ , 代入目标函数, 可得

$$f(\mathbf{x}_B, \mathbf{x}_N) = \mathbf{c}_B^T D_B^{-1}\boldsymbol{\theta} - \mathbf{c}_B^T D_B^{-1}D_N\mathbf{x}_N + \mathbf{c}_N^T \mathbf{x}_N = \mathbf{c}_B^T D_B^{-1}\boldsymbol{\theta} + (\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1}D_N)\mathbf{x}_N$$

若令  $\mathbf{x}_N = \mathbf{0}$ , 则  $\mathbf{x}_B = D_B^{-1}\boldsymbol{\theta}$ , 此时若  $\mathbf{x}_B \geq \mathbf{0}$ , 则称其为基可行解. 此时,  $f(\mathbf{x}_B, \mathbf{x}_N) = \mathbf{c}_B^T D_B^{-1}\boldsymbol{\theta}$ . 显然, 若  $\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1}D_N \geq \mathbf{0}$ , 则基可行解为最优解. 线性规划的约束条件数总是不大于决策变量数的, 由于约束条件的存在, 其可行解空间是以决策变量数为维数的线性空间的一个线性子空间. 而由于其线性特性, 最优解只可能存在于可行解空间的边界上.

回到线性规划问题(11)-(13), 把它改写成标准形式:

$$\min \sum_{i=1}^n a_i(x_i - Z_i)u_i + \sum_{i=n+1}^{n+m} 0 \cdot u_i \quad (14)$$

$$\text{s.t.} \sum_{j=1}^n u_j / \mu_{ij} + u_{n+i} = \alpha_i, \quad i = 1, \dots, m. \quad (15)$$

$$u_j \geq 0, \quad j = 1, \dots, n, n+1, \dots, n+m. \quad (16)$$

其中,  $u_{n+1}, \dots, u_{n+m}$  为松弛变量. 不妨设  $(\alpha_1, \dots, \alpha_m) = (1, \dots, 1)$ , 于是式(15)可写为

$$\begin{bmatrix} 1/\mu_{11} & \cdots & 1/\mu_{1n} & 1 & & \\ \vdots & \ddots & \vdots & & \ddots & \\ 1/\mu_{m1} & \cdots & 1/\mu_{mn} & & & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \\ u_{n+1} \\ \vdots \\ u_{n+m} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

$$\text{上式中, 令 } D_B = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}, \quad D_N = \begin{bmatrix} 1/\mu_{11} & \cdots & 1/\mu_{1n} \\ \vdots & \ddots & \vdots \\ 1/\mu_{m1} & \cdots & 1/\mu_{mn} \end{bmatrix}, \quad \mathbf{c}_B^T = (0, \dots, 0),$$

$$\mathbf{c}_N^T = [a_1(x_1 - Z_1), \dots, a_n(x_n - Z_n)], \quad \mathbf{u}_B = (u_{n+1}, \dots, u_{n+m})^T, \quad \mathbf{u}_N = (u_1, \dots, u_n)^T. \text{ 于是}$$

$$\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N = [a_1(x_1 - Z_1), \dots, a_n(x_n - Z_n)]$$

因此  $\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N \geq \mathbf{0}$  即等价于  $x_1 > Z_1, \dots, x_n > Z_n$ . 也就是说, 当



到另一个区域，而这个区域对应于另一个基可行解，此时控制才发生变化。下面以一个例子说明这一点。

有一个 2 零件 3 机器制造系统，根据 Bellman 方程获得的线性规划模型如下：

$$\min c_1(\mathbf{x})u_1 + c_2(\mathbf{x})u_2 \quad (19)$$

$$\text{s.t. } \tau_{11}u_1 + \tau_{12}u_2 + s_1 = 1 \quad (20)$$

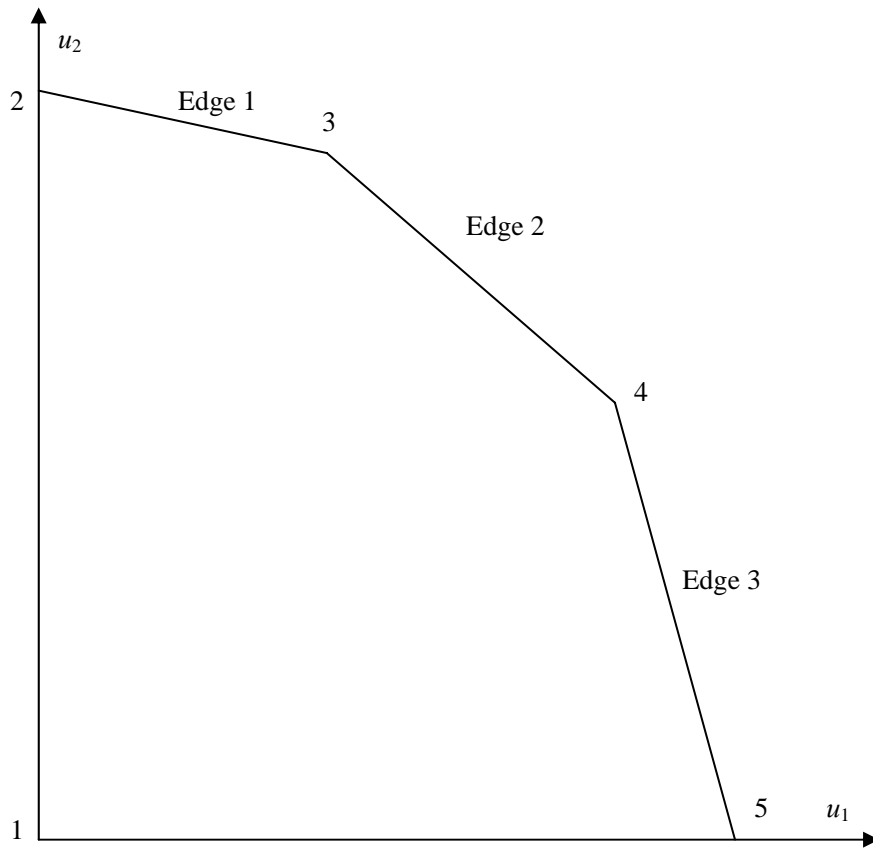
$$\tau_{21}u_1 + \tau_{22}u_2 + s_2 = 1 \quad (21)$$

$$\tau_{31}u_1 + \tau_{32}u_2 + s_3 = 1 \quad (22)$$

其中， $c_1(\mathbf{x}) = a_1(x_1 - Z_1)$ ， $c_2(\mathbf{x}) = a_2(x_2 - Z_2)$ ， $\tau_{ij} = 1/\mu_{ij}$ 。若写成矩阵形式，则有

$$D = \begin{bmatrix} \tau_{11} & \tau_{12} & 1 \\ \tau_{21} & \tau_{22} & 1 \\ \tau_{31} & \tau_{32} & 1 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

假定  $\tau_{12} > \tau_{22} > \tau_{32}$ ， $\tau_{31} > \tau_{21} > \tau_{11}$ 。于是，由式(20)-(22)确定的可行解空间如下图所示，其中各顶点为：



顶点 1:  $u_1 = u_2 = 0$ ， $s_1 = s_2 = s_3 = 1$ 。基变量为  $s_1, s_2, s_3$ 。

顶点 2:  $u_1 = s_1 = 0$ ， $u_2 = 1/\tau_{12}$ ， $s_2 = 1 - \tau_{22}/\tau_{12}$ ， $s_3 = 1 - \tau_{32}/\tau_{12}$ 。基变量为  $u_2, s_2, s_3$ 。

顶点 3:  $s_1 = s_2 = 0$ ， $u_1 = \frac{\tau_{22} - \tau_{12}}{\tau_{11}\tau_{22} - \tau_{21}\tau_{12}}$ ， $u_2 = \frac{\tau_{11} - \tau_{21}}{\tau_{11}\tau_{22} - \tau_{21}\tau_{12}}$ ，

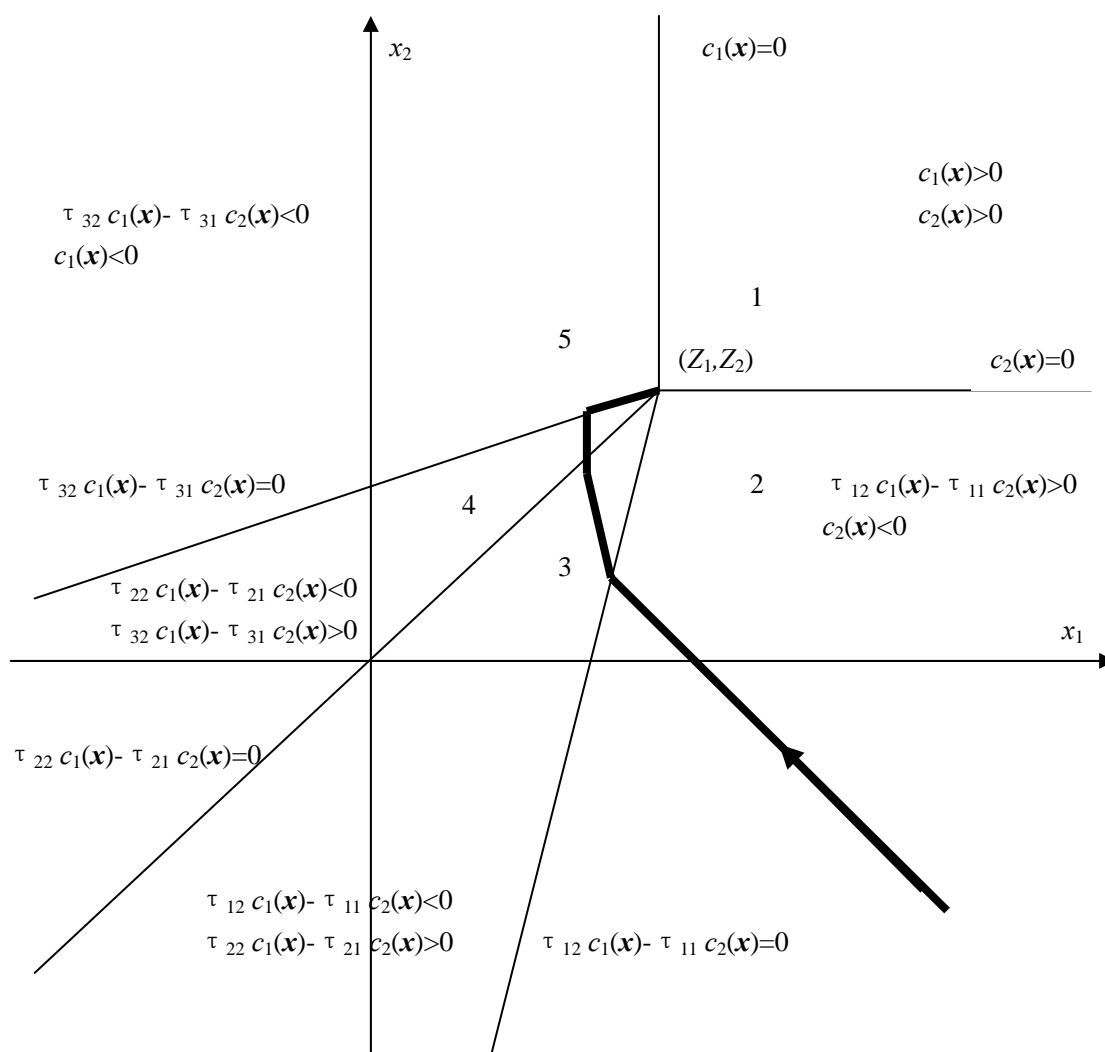
$s_3 = 1 - \frac{\tau_{31}(\tau_{22} - \tau_{12}) - \tau_{32}(\tau_{11} - \tau_{21})}{\tau_{11}\tau_{22} - \tau_{21}\tau_{12}}$ 。基变量为  $u_1, u_2, s_3$ 。

顶点 4:  $s_2 = s_3 = 0$ ,  $u_1 = \frac{\tau_{22} - \tau_{32}}{\tau_{31}\tau_{22} - \tau_{21}\tau_{32}}$ ,  $u_2 = \frac{\tau_{31} - \tau_{21}}{\tau_{31}\tau_{22} - \tau_{21}\tau_{32}}$ ,

$$s_1 = 1 - \frac{\tau_{11}(\tau_{22} - \tau_{32}) - \tau_{12}(\tau_{31} - \tau_{21})}{\tau_{31}\tau_{22} - \tau_{21}\tau_{32}}. \text{ 基变量为 } u_1, u_2, s_1.$$

顶点 5:  $u_2 = s_3 = 0$ ,  $u_1 = 1/\tau_{31}$ ,  $s_1 = 1 - \tau_{11}/\tau_{31}$ ,  $s_2 = 1 - \tau_{21}/\tau_{31}$ . 基变量为  $u_1, s_1, s_2$ .

下面计算各顶点对应的系统状态所处区域:



顶点 1:  $\mathbf{c}_N^T = [c_1(\mathbf{x}), c_2(\mathbf{x})] = [a_1(x_1 - Z_1), a_2(x_2 - Z_2)]$ ,  $\mathbf{c}_B^T = [0, 0, 0]$ ,

$$D_N = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \\ \tau_{31} & \tau_{32} \end{bmatrix}, \quad D_B = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

$\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N = [a_1(x_1 - Z_1), a_2(x_2 - Z_2)] > 0 \Rightarrow x_1 > Z_1, x_2 > Z_2$  (区域 1)

顶点 2:  $\mathbf{c}_N^T = [c_1(\mathbf{x}), 0] = [a_1(x_1 - Z_1), 0]$ ,  $\mathbf{c}_B^T = [c_2(\mathbf{x}), 0, 0] = [a_2(x_2 - Z_2), 0, 0]$ ,

$$D_N = \begin{bmatrix} \tau_{11} & 1 \\ \tau_{21} & 0 \\ \tau_{31} & 0 \end{bmatrix}, \quad D_B = \begin{bmatrix} \tau_{12} & 0 & 0 \\ \tau_{22} & 1 & 0 \\ \tau_{32} & 0 & 1 \end{bmatrix}$$

$$\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N = \left[ c_1(\mathbf{x}) - \frac{\tau_{11}c_2(\mathbf{x})}{\tau_{12}}, -\frac{c_2(\mathbf{x})}{\tau_{12}} \right] > 0 \rightarrow$$

$$\tau_{12}c_1(\mathbf{x}) - \tau_{11}c_2(\mathbf{x}) > 0, \quad c_2(\mathbf{x}) < 0$$

i.e.  $\tau_{12}a_1x_1 - \tau_{11}a_2x_2 > \tau_{12}a_1Z_1 - \tau_{11}a_2Z_2, \quad x_2 < Z_2$  (区域 2)

顶点 3:  $\mathbf{c}_N^T = [0,0], \quad \mathbf{c}_B^T = [c_1(\mathbf{x}), c_2(\mathbf{x}), 0] = [a_1(x_1 - Z_1), a_2(x_2 - Z_2), 0],$

$$D_N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad D_B = \begin{bmatrix} \tau_{11} & \tau_{12} & 0 \\ \tau_{21} & \tau_{22} & 0 \\ \tau_{31} & \tau_{32} & 1 \end{bmatrix}$$

$$\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N = \left[ \frac{\tau_{21}c_1(\mathbf{x}) - \tau_{22}c_2(\mathbf{x})}{\tau_{11}\tau_{22} - \tau_{21}\tau_{12}}, \frac{\tau_{12}c_1(\mathbf{x}) - \tau_{11}c_2(\mathbf{x})}{\tau_{11}\tau_{22} - \tau_{21}\tau_{12}} \right] > 0 \rightarrow$$

$$\tau_{21}c_2(\mathbf{x}) - \tau_{22}c_1(\mathbf{x}) < 0, \quad \tau_{12}c_1(\mathbf{x}) - \tau_{11}c_2(\mathbf{x}) < 0 \quad (\text{由于 } \tau_{11}\tau_{22} - \tau_{21}\tau_{12} < 0)$$

i.e.  $\tau_{21}a_2x_2 - \tau_{22}a_1x_1 < \tau_{21}a_2Z_2 - \tau_{22}a_1Z_1, \tau_{12}a_1x_1 - \tau_{11}a_2x_2 < \tau_{12}a_1Z_1 - \tau_{11}a_2Z_2$  (区域 3)

顶点 4:  $\mathbf{c}_N^T = [0,0], \quad \mathbf{c}_B^T = [c_1(\mathbf{x}), c_2(\mathbf{x}), 0] = [a_1(x_1 - Z_1), a_2(x_2 - Z_2), 0],$

$$D_N = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_B = \begin{bmatrix} \tau_{11} & \tau_{12} & 1 \\ \tau_{21} & \tau_{22} & 0 \\ \tau_{31} & \tau_{32} & 0 \end{bmatrix}$$

$$\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N = \left[ -\frac{\tau_{32}c_1(\mathbf{x}) - \tau_{31}c_2(\mathbf{x})}{\tau_{21}\tau_{32} - \tau_{31}\tau_{22}}, -\frac{\tau_{21}c_2(\mathbf{x}) - \tau_{22}c_1(\mathbf{x})}{\tau_{21}\tau_{32} - \tau_{31}\tau_{22}} \right] > 0 \rightarrow$$

$$\tau_{32}c_1(\mathbf{x}) - \tau_{31}c_2(\mathbf{x}) > 0, \quad \tau_{21}c_2(\mathbf{x}) - \tau_{22}c_1(\mathbf{x}) > 0 \quad (\text{由于 } \tau_{21}\tau_{32} - \tau_{31}\tau_{22} < 0)$$

i.e.  $\tau_{31}a_2x_2 - \tau_{32}a_1x_1 < \tau_{31}a_2Z_2 - \tau_{32}a_1Z_1, \tau_{21}a_2x_2 - \tau_{22}a_1x_1 < \tau_{21}a_2Z_2 - \tau_{22}a_1Z_1$  (区域 4)

顶点 5:  $\mathbf{c}_N^T = [c_2(\mathbf{x}), 0] = [a_2(x_2 - Z_2), 0], \quad \mathbf{c}_B^T = [c_1(\mathbf{x}), 0, 0] = [a_1(x_1 - Z_1), 0, 0],$

$$D_N = \begin{bmatrix} \tau_{12} & 1 \\ \tau_{22} & 0 \\ \tau_{32} & 0 \end{bmatrix}, \quad D_B = \begin{bmatrix} \tau_{11} & 0 & 0 \\ \tau_{21} & 1 & 0 \\ \tau_{31} & 0 & 1 \end{bmatrix}$$

$$\mathbf{c}_N^T - \mathbf{c}_B^T D_B^{-1} D_N = \left[ c_2(\mathbf{x}) - \frac{\tau_{32}c_1(\mathbf{x})}{\tau_{31}}, -\frac{c_1(\mathbf{x})}{\tau_{31}} \right] > 0 \rightarrow$$

$$\tau_{32}c_1(\mathbf{x}) - \tau_{31}c_2(\mathbf{x}) < 0, \quad c_1(\mathbf{x}) < 0$$

i.e.  $\tau_{32}a_1x_1 - \tau_{31}a_2x_2 < \tau_{32}a_1Z_1 - \tau_{31}a_2Z_2, \quad x_1 < Z_1$ . (区域 5)

设  $\tau_{11} = 1, \tau_{12} = 3, \tau_{21} = 2.5, \tau_{22} = 2.5, \tau_{31} = 4, \tau_{32} = 1, Z_1 = 1, Z_2 = 1, a_1 = 1, a_2 = 1, d_1 = 0.2, d_2 = 0.05, x_1(0) = 2, x_1(0) = -1$ . 则在各个区域内的控制如下:

区域 2:  $(u_1, u_2) = (0, 0.333)$ , 在此控制下, 系统状态在  $t_1 = 5.66$  时到达区域 2 与 3 的边界直线  $\tau_{12}a_1x_1 - \tau_{11}a_2x_2 = \tau_{12}a_1Z_1 - \tau_{11}a_2Z_2$  (即  $3x_1 - x_2 = 2$ ) 上的点  $(x_1, x_2) = (0.87, 0.60)$ .

区域 2 与区域 3 的边界: 令  $h = 3x_1 - x_2 - 2$ , 则在区域 2 内,  $h > 0$ , 在区域 3 内,  $h < 0$ . 而  $dh/dt = 3(u_1 - d_1) - (u_2 - d_2)$ , 在区域 2 内,  $dh/dt < 0$ , 在区域 3 内,  $dh/dt < 0$ . 因此系统状态不能在边界上停留, 而是趋向于进入区域 3.

区域 3:  $(u_1, u_2) = (0.1, 0.3)$ . 在此控制下, 系统状态在  $t_2 = 6.41$  时到达区域 3 与 4 的边界直线  $\tau_{21}a_2x_2 - \tau_{22}a_1x_1 = \tau_{21}a_2Z_2 - \tau_{22}a_1Z_1$  (即  $x_1 - x_2 = 0$ ) 上的点  $(x_1, x_2) = (0.79, 0.79)$ .

区域 3 与区域 4 的边界: 令  $h = x_1 - x_2$ , 则在区域 3 内,  $h > 0$ , 在区域 4 内,  $h < 0$ . 而  $dh/dt = (u_1 - d_1) - (u_2 - d_2)$ , 在区域 3 内,  $dh/dt < 0$ , 在区域 4 内,  $dh/dt < 0$ . 因此系统状态不能在边界上停留, 而是趋向于进入区域 4.

区域 4:  $(u_1, u_2) = (0.2, 0.2)$ , 在此控制下, 系统状态在  $t_3 = 12.3$  时到达区域 4 与 5 的边界直线  $\tau_{32}a_1x_1 - \tau_{31}a_2x_2 = \tau_{32}a_1Z_1 - \tau_{31}a_2Z_2$  (即  $x_1 - 4x_2 = -3$ ) 上的点  $(x_1, x_2) = (0.79, 0.95)$ .

区域 4 与区域 5 的边界: 令  $h = x_1 - 4x_2 + 3$ , 则在区域 4 内,  $h > 0$ , 在区域 5 内,  $h < 0$ . 而  $dh/dt = (u_1 - d_1) - 4(u_2 - d_2)$ , 在区域 4 内,  $dh/dt < 0$ , 在区域 5 内,  $dh/dt > 0$  (由于  $(u_1, u_2) = (0.25, 0)$ ). 因此系统状态就趋向于被固定在在边界上. 在这一边界上, 由于系统状态满足  $x_1 - 4x_2 = -3$ , 因此控制变量必须满足  $u_1 - d_1 = 4(u_2 - d_2)$ . 同时, 最优控制还应该是在区域 4 和 5 中的控制的凸组合, 即  $(u_1, u_2) = G(0.2, 0.2) + (1 - G)(0.25, 0)$ . 于是可解得  $G = 0.29$ ,  $(u_1, u_2) = (0.235, 0.059)$ . 在这一控制下, 系统状态在  $t_4 = 18.23$  时到达 Hedging Point  $(Z_1, Z_2)$ .

Hedging Point  $(Z_1, Z_2)$ :  $(u_1, u_2) = (0.2, 0.05) = (d_1, d_2)$ . 系统状态被保持在 Hedging Point 上.

几点说明:

1.  $\alpha$  变化时, 平面  $u_1 - u_2$  可行解空间的上的边界直线会发生平移, 从而顶点发生变化. 平移到一定程度, 基可行解会发生变化, 即某些顶点会消失, 另外一些顶点会出现, 这就意味着某些约束会失效. 若  $\alpha$  的变化不足以改变基可行解, 则边界只是与系数矩阵的基向量有关, 而与  $\alpha$  无关. 这里  $\alpha$  表示机器的状态,  $\alpha$  的变化意味着生产能力的变化.
2. Hedging Point  $(Z_1, Z_2)$  发生变化时, 平面  $x_1 - x_2$  上的边界会发生变化, 但仅仅是平移.
3. 若 cost-to-go 函数具有二次型(quadratic)的形式, 则系统状态空间被划分的各区域之间的边界是直线. 在边界两侧控制策略是不同的, 边界成了切换控制策略的条件. 这一类控制策略称为 “switching curve policy”, 边界是直线时, 称为 “linear switching curve policy”, 它们是 hedging point policy 的一种推广的形式.

## 第 12 章

### 基于连续系统模型的制造系统分析与控制(2)

本章继续介绍基于连续系统模型的制造系统分析与控制，主要内容为：

(1) Hedging Point 控制策略的推广 --- Control-Point 控制策略. 参见如下文献：

Gershwin, S. B., 2000. Design and operation of manufacturing systems: the control-point policy. *IIE Transactions*, Vol.32, pp.891-906.

(2) 递阶生产计划. 参见如下文献：

Yan, H. S., 2002. Hierarchical production planning in flexible automated workshops with delay interaction. *International Journal of Advanced Manufacturing Technology*, Vol.19, pp.358-369.

## 第 13 章

### 制造系统设计概述

本章简要介绍基于排队网络模型的柔性制造系统设计方法，主要内容参见如下文献：

Lee, H. F., 1999. Optimal design for flexible manufacturing systems: generalized analytical methods. *IIE Transactions*, Vol.31, pp.965-976.