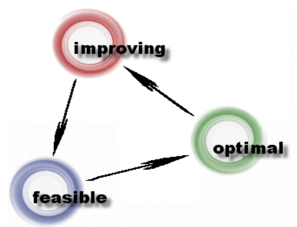
排隊模型與等候時間

政治大學 應用數學系
陸 行

Hsing Luh
Department of Mathematical Sciences

大綱

- 作業研究
- 排隊與等候



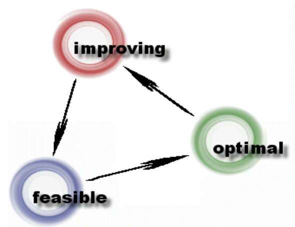
作業研究

所謂「運籌」，就是「調度」，就是將有限資源做有效及最佳調度與配置的意思。

$$f(\mathbf{x}): R^n \rightarrow R$$

目標　配置　資源　有效調度　最佳值

例如：一個企業管理者必須將原料、半成品、成品、存貨、勞動力、資金、時間，做最佳的調度與組合，期使成本降到最低，利潤增到最大。

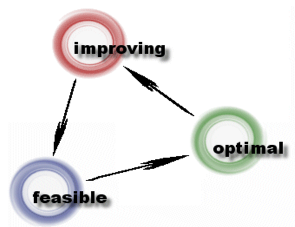


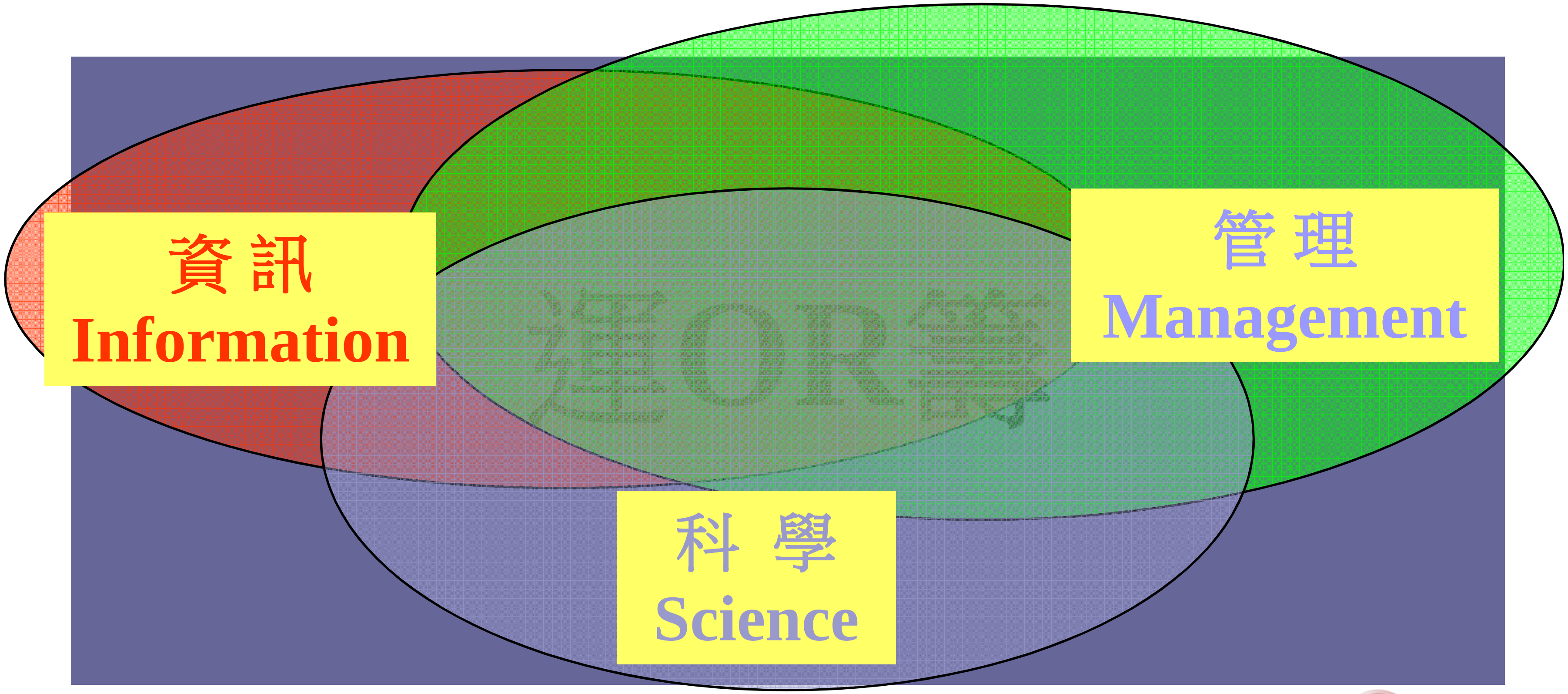
運籌科學與相關領域 OR and Related Fields

- 作業研究
- 運籌學
- Operations Research

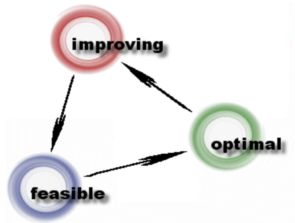
- 數學系、應用數學系
Department of Mathematics, Department of Mathematical Sciences
- 工業工程系、資訊科學系
Department of Industrial Engineering, Department of Computer Science
- 工商管理系、統計系
Department of Business Administration, Department of Statistics

- 線性規劃、非線性規劃、動態規劃
Linear Programming, Nonlinear Programming, Dynamic Programming
- 等候理論、隨機規劃、系統模擬
Queueing Theory, Stochastic Programming, System Simulation





資訊
Information
管理
Management
運OR籌
科學
Science



improving
optimal
feasible

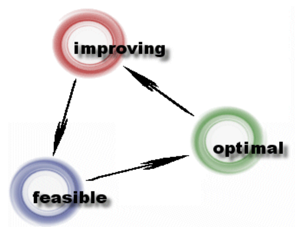
1. **Introduction**
   - History.
   - Methodology. Models.
2. **Linear programming**
   - Applications.
   - Simplex algorithm.
   - Sensitivity analysis.
   - Integer programming.
3. **Network analysis**
   - Transportation problems.
   - Maximum flow problems.
   - CPM and PERT – Project management.
4. **Inventory management**
   - Deterministic EOQ models.
   - Probabilistic models.
   - Production management
5. **Multicriteria decision aid**
   - Methods.
   - Multicriteria decision support systems.
   - Group decision systems.
   - Applications.

模式與方法
Models and Methodologies

	Models deterministic	probabilistic
OR solutions / optimization	✓	✓
OR solutions \ simulation	✓	✓

OR Methodologies

- 事前 / 事後 分析



<u>Some of the most widely used techniques</u> :

- **<u>Linear programming</u>** :
  - Optimization under constraints
  - Production planning, scheduling, financial planning,
- **<u>Simulation methods</u> :**
  - Simulate complex systems
  - Queueing
- **<u>PERT/CPM</u>** :
  - Complex projects management
- **<u>Network models</u>** :
  - Transportation problems
- **<u>Inventory management</u>** :
  - Reduce inventory levels and costs
  - JIT production management

- **Statistical analysis** :

  Summarize and analyze data.

- **Multicriteria decision aid** :

  Identify best compromise solutions

  When conflicting objectives are considered

- **Group decision aid** :

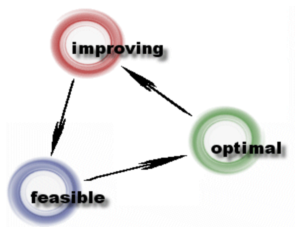
  Achieve consensus efficiently

- **Dynamic programming** :

  Repetitive (w.r.t. time) decision problems.

- **Reliability theory** :

  Reliability of equipments.



供應鏈 Supply Chain Management

The Goal: A Process of Ongoing Improvement

by

Eliyahu M. Goldratt

推銷員問題 Traveling Salesman Problem TSP

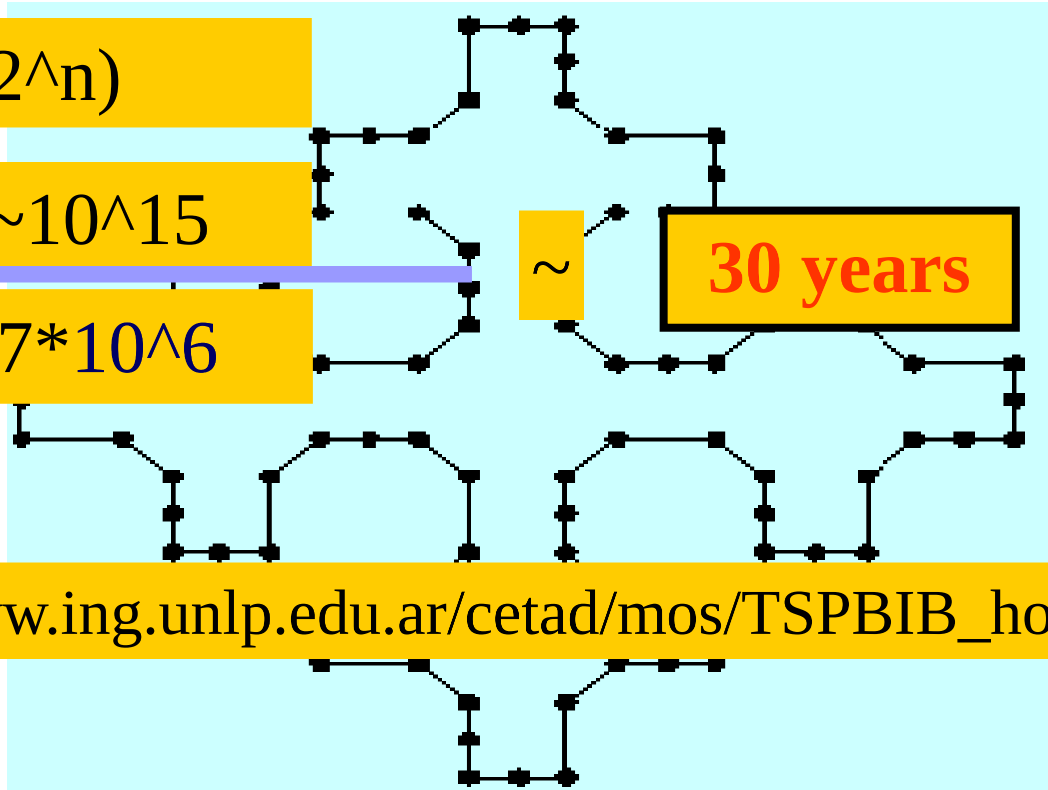
O(2^n)

2^50 ~10^15

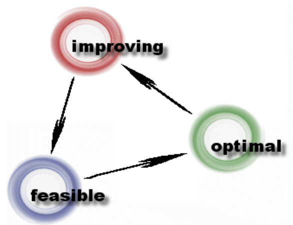
3*10^7*10^6

~

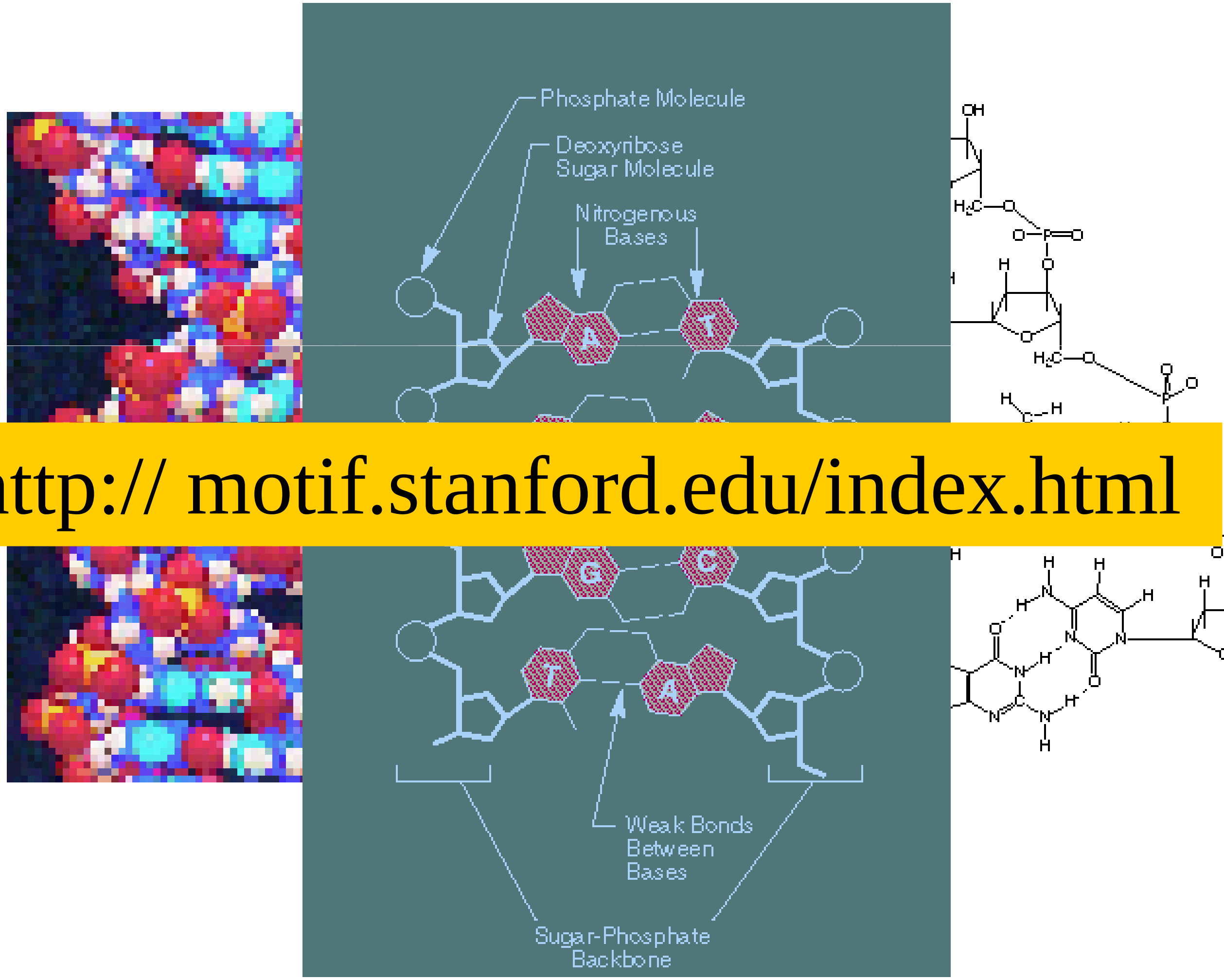
30 years



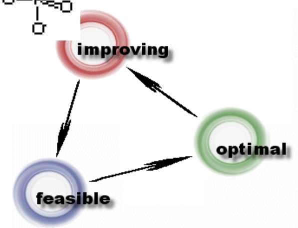
http://www.ing.unlp.edu.ar/cetad/mos/TSPBIB_home.html



BIOINFORMATICS 生物資訊



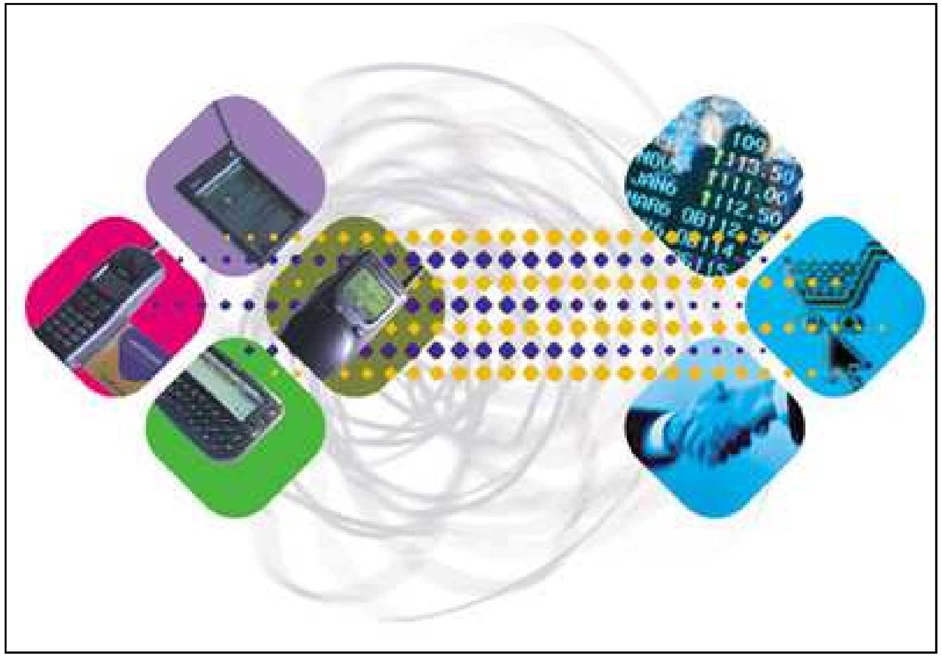
http:// motif.stanford.edu/index.html



CRYTOGRAPHY 密碼學

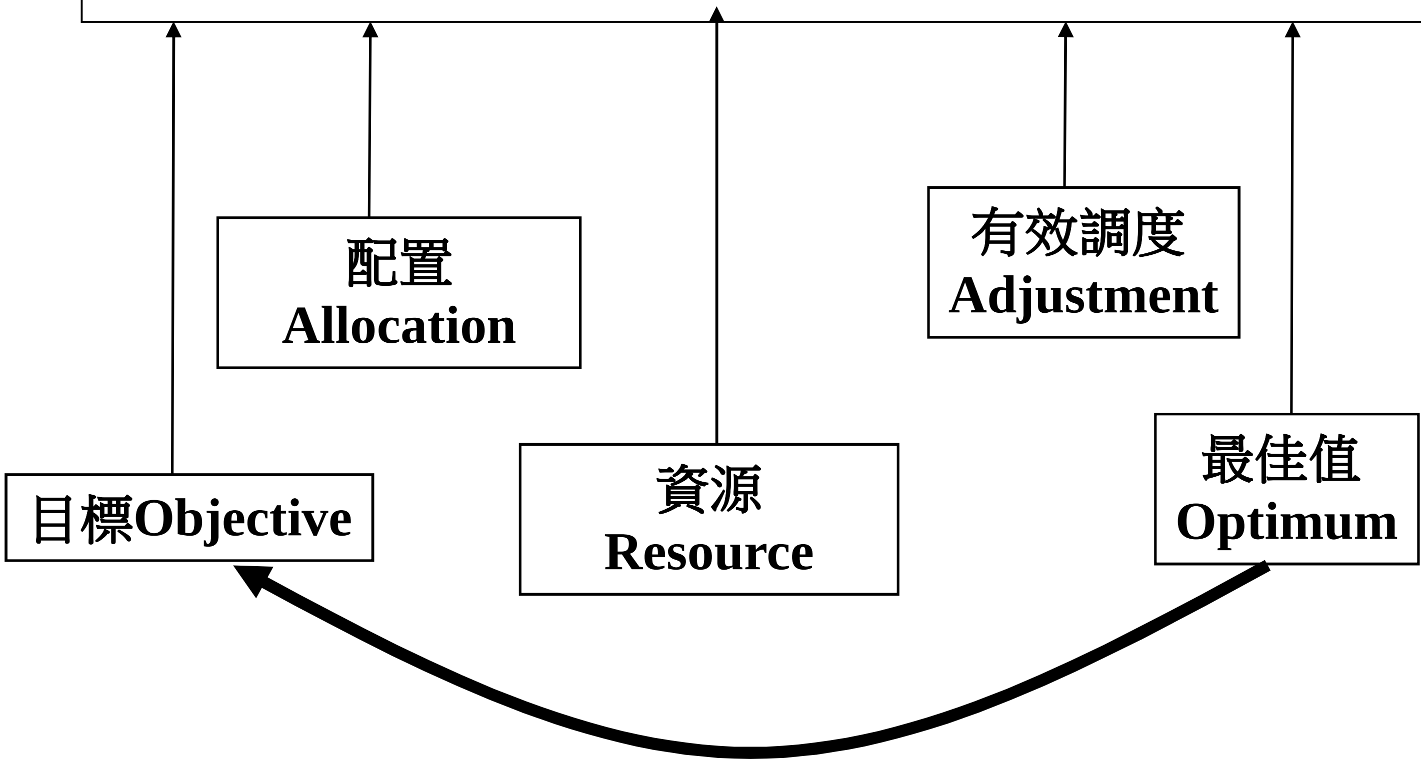
Elliptic Curve

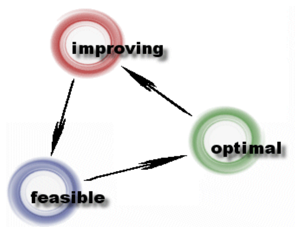
http:// www.certicom.ca



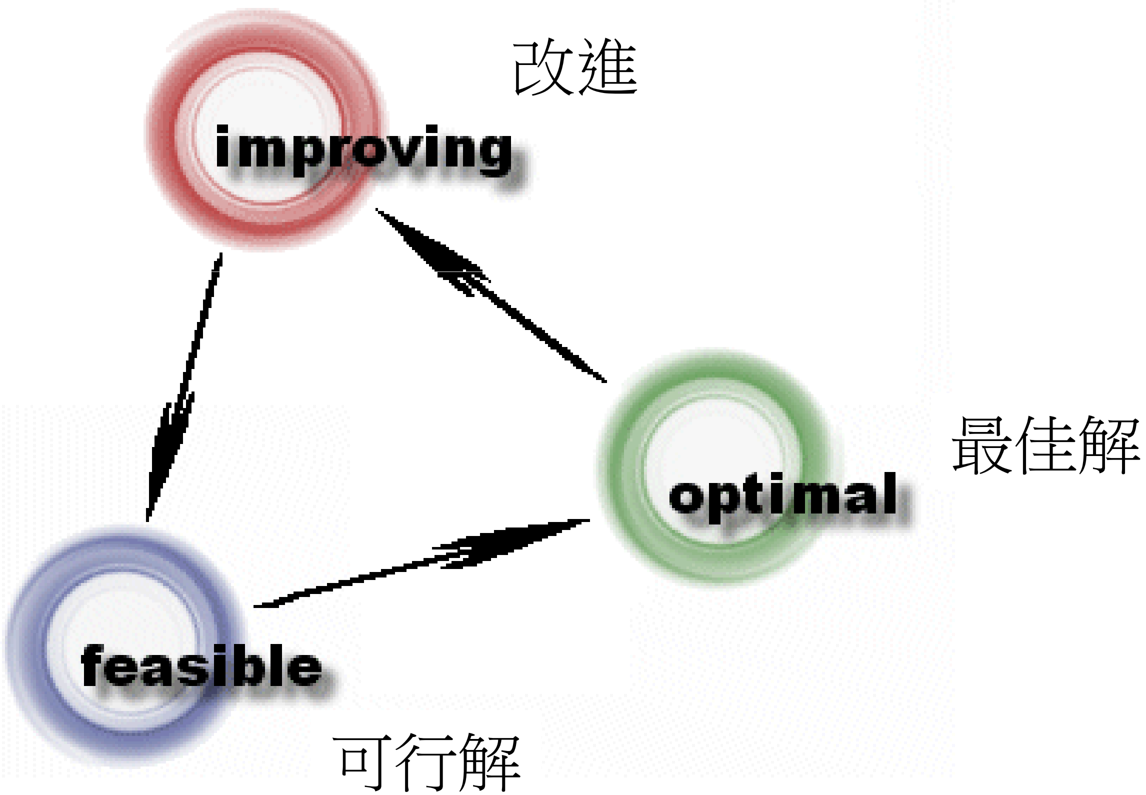
動態系統

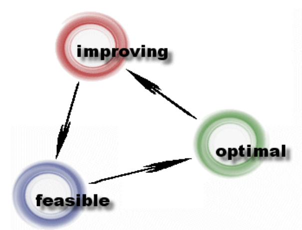
$$f_t(\mathbf{x}) : R^n \rightarrow R^n$$





最佳解


改進
improving
optimal
最佳解
feasible
可行解


improving
optimal
feasible

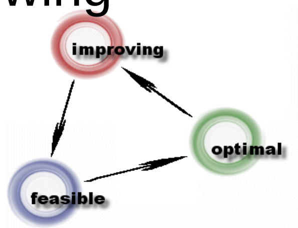
			A	B	C	D
			Cohort/aggregate level/counts		Individual level	
			Expected value, continuous state, deterministic	Markovian, discrete state, stochastic	Markovian, discrete state, individuals	Non-Markovian, discrete-state, individuals
1	No interaction allowed	Untimed	Decision tree rollback	Simulated decision tree (SDT)	Individual sampling model (ISM): Simulated patient-level decision tree (SPLDT)	
2		Timed	Markov model (evaluated deterministically)	Simulated Markov model (SMM)	Individual sampling model (ISM): Simulated patient-level Markov model (SPLMM) (variations as in quadrant below for patient level models with interaction)	

			A	B	C	D
			Cohort/aggregate level/counts		Individual level	
			Expected value, continuous state, deterministic	Markovian, discrete state, stochastic	Markovian, discrete state, individuals	Non-Markovian, discrete-state, individuals
3	Interaction allowed	Discrete time	System dynamics (finite difference equations, FDE)	Discrete time Markov chain model (DTMC)	Discrete-time individual event history model (DT, IEH)	Discrete individual simulation (DT, DES)
4		Continu-ous time	System dynamics (ordinary differential equations, ODE)	Continuous time Markov chain model (CTMC)	Continuous time Individual event history model (CT, IEH)	Discrete event simulation (CT, DES)

Queueing Theory -- Establishing policies for sending in resupply aircraft to small airfields with maximum on ground (MOG) limits. For example, the airfield at Haiti had a MOG of 3 C-5s.

Decision Analysis -- A recent study was completed that looked at methods for considering the counter proliferation effects of weapons of mass destruction when evaluating new systems.

Response Surface Methodology -- Using multiple runs of a large theater-level model, Thunder, to build a response surface allowing quick analysis of effects of changes in air apportionment.

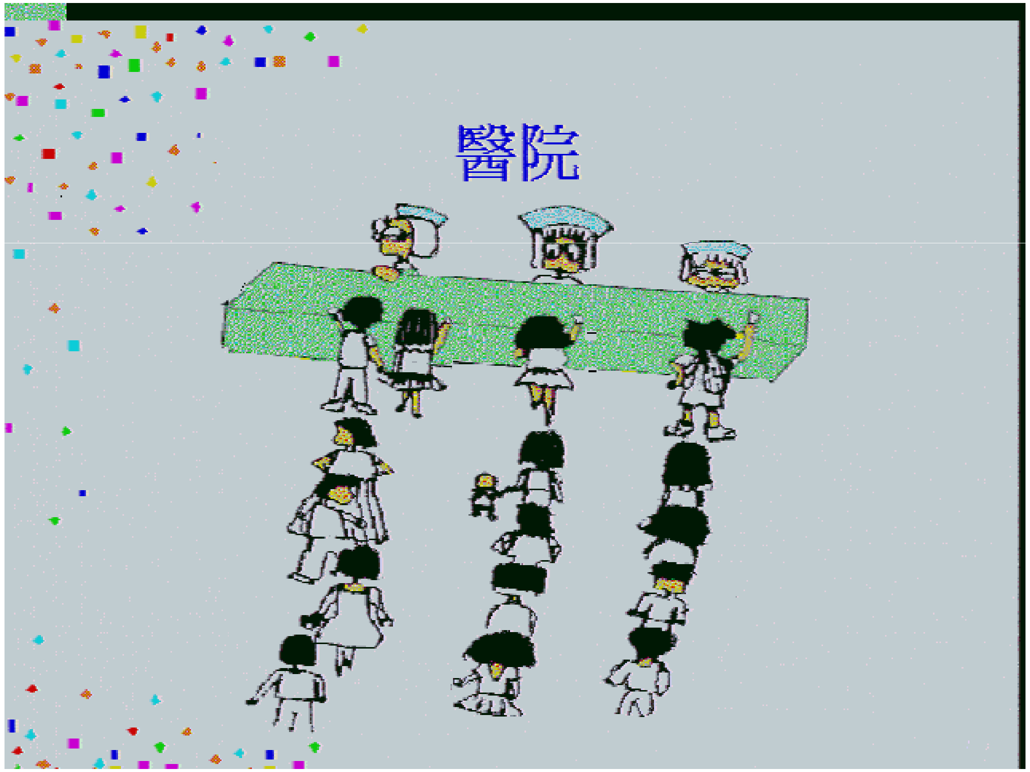


Courtesy to Internet Information

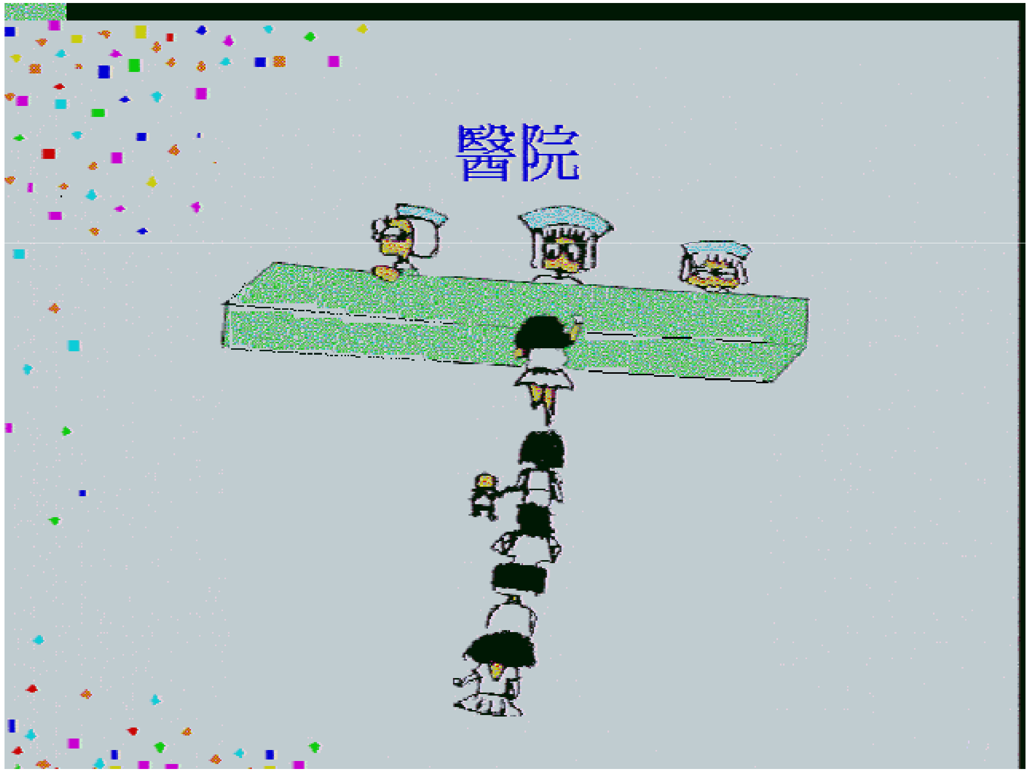
- 貓空纜車假日試辦「進站時間號碼牌」管制
- 「進站時間號碼牌」管制措施是以10分鐘為單位，每單位設定運送遊客200人。民眾只要告知工作人員欲搭乘人數，就可領到一張號碼牌，牌上載明進站時段，在該時段回到入口處就可進站搭車；逾10分鐘以上須重抽牌。

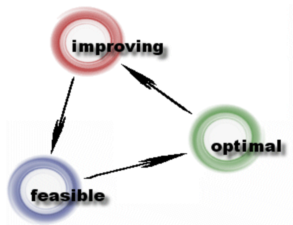
- 全台瘋周年慶 要進**UNIQLO**先排**3**小時
- iwait
- Psychology of Waiting Lines

多線排隊



單線排隊 多人服務





等候時間的計算與應用

- 單線、多線排隊
- 如何描述等候時間

如何描述等候時間

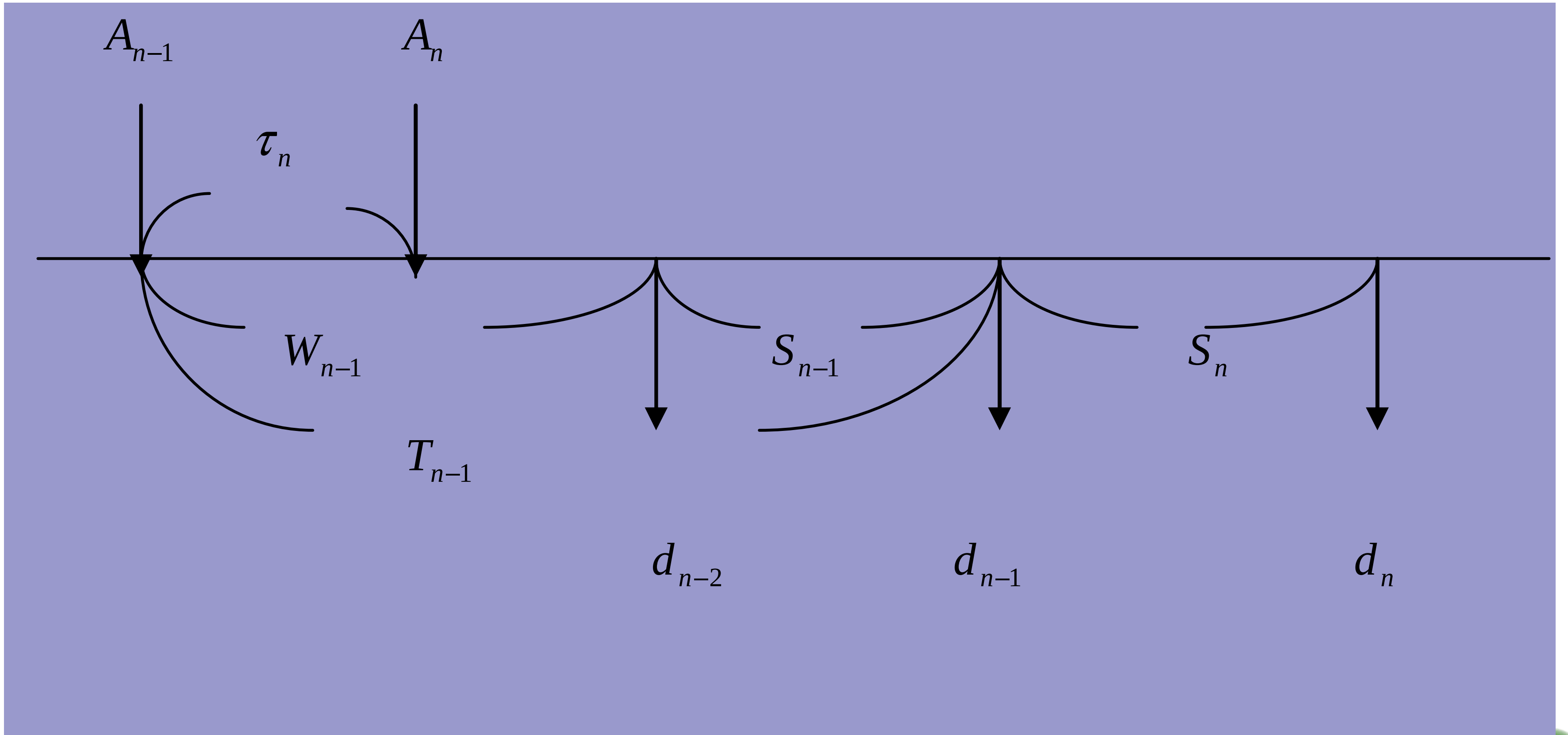
- 平均等候時間

$$\frac{\text{等候時間的總和}}{\text{總人數}}$$

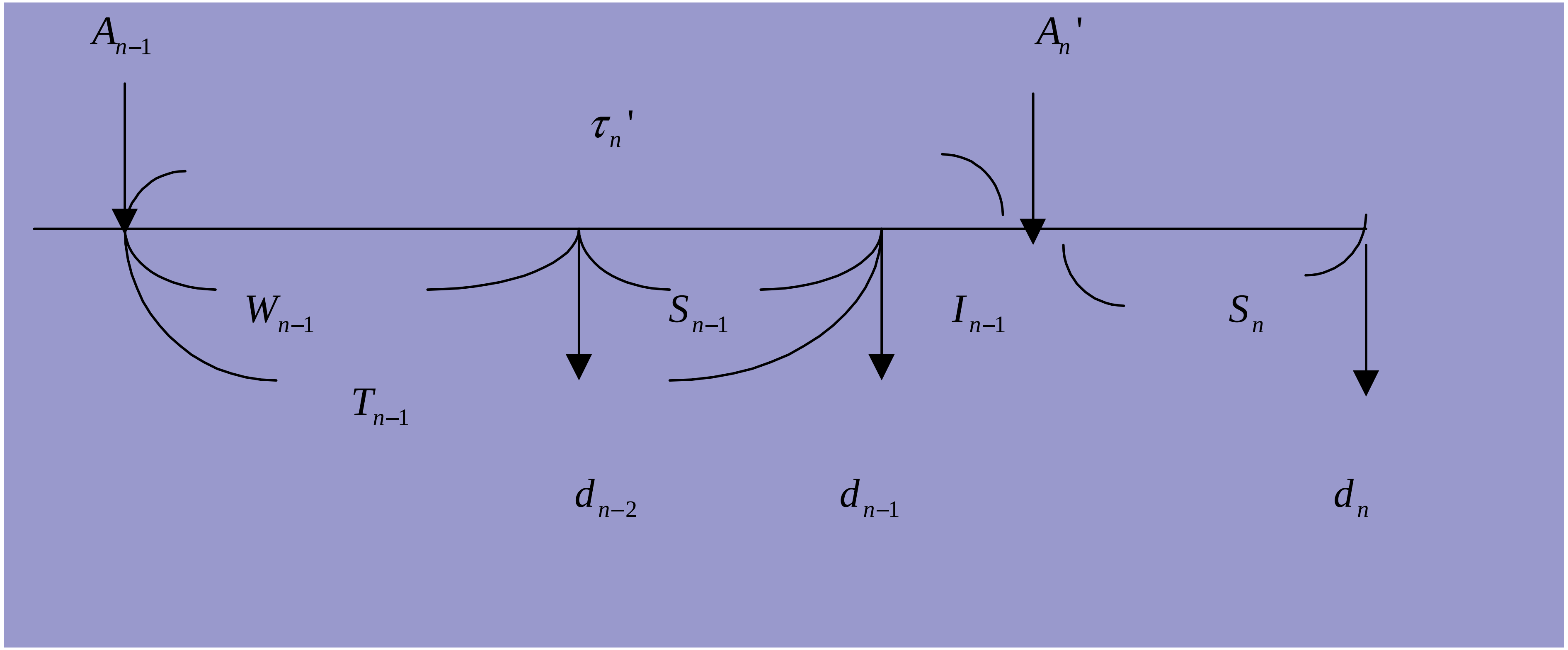
- 等候時間的變異數
- 等候時間的機率分配函數

Agner Krarup Erlang (1878-1929)

到達時間是隨機變數



服務時間是隨機變數



等候時間是隨機變數

$$T_n = S_n + W_n$$

$$\Pr\{W_n < t\}$$

An M/M/1 Queueing system

- P (exact 1 arrival in [t, t+Δt])= $\lambda\Delta t$
- P (no arrivals in [t, t+Δt])= 1-$\lambda\Delta t$
- P (1 service completion in [t, t+Δt])= $\mu\Delta t$
- P (no service completion in [t, t+Δt])= 1-$\mu\Delta t$
- $P_n(t)$ = P (# of **customers** is n at time t)

Courtesy to Internet Information

狀態平衡方程式

$$P_n(t+\Delta t) = P_n(t)a_{n,n}(\Delta t) + P_{n-1}(t)a_{n-1,n}(\Delta t) + P_{n+1}(t)a_{n+1,n}(\Delta t)$$

$$P_0(t+\Delta t) = P_0(t)a_{0,0}(\Delta t) + P_1(t)a_{1,0}(\Delta t)$$

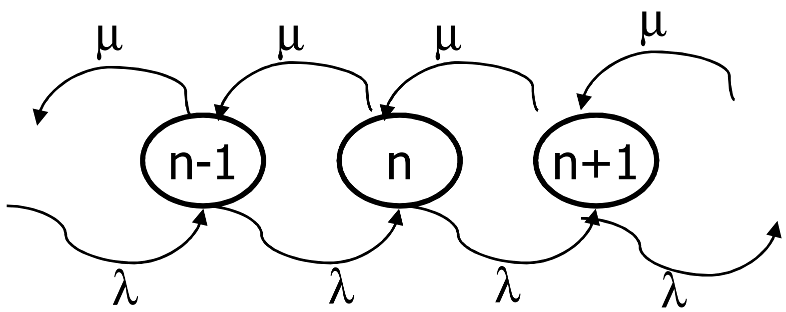
$$P_n(t+\Delta t) = P_n(t)(1-\lambda\Delta t)(1-\mu\Delta t) + P_{n-1}(t)\lambda\Delta t + P_{n+1}(t)\mu\Delta t$$

$$P_0(t+\Delta t) = P_0(t)(1-\lambda\Delta t) + P_1(t)\mu\Delta t$$

$$\frac{dP_n(t)}{dt} = -(\lambda+\mu)P_n(t) + \lambda P_{n-1}(t) + \mu P_{n+1}(t),$$

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t)$$

狀態平衡方程式



$$\frac{dP_n(t)}{dt} = -(\lambda + \mu)P_n(t) + \lambda P_{n-1}(t) + \mu P_{n+1}(t),$$

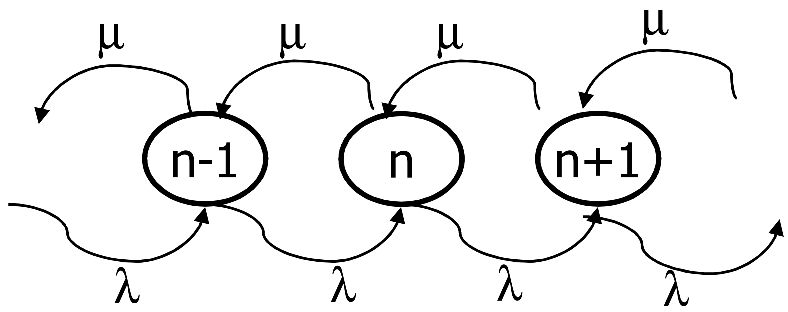
$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t)$$

$$P_0(0) = 1.0$$

$$P_n(0) = 0, n \geq 1$$

Courtesy to Internet Information

長時間的觀察值



$$P_n = \lim_{t \to \infty} P_{i,n}(t)$$

$$0 = -(\lambda + \mu) P_n + \lambda P_{n-1} + \mu P_{n+1},$$

$$0 = -\lambda P_0 + \mu P_1$$

Courtesy to Internet Information

基本假設條件

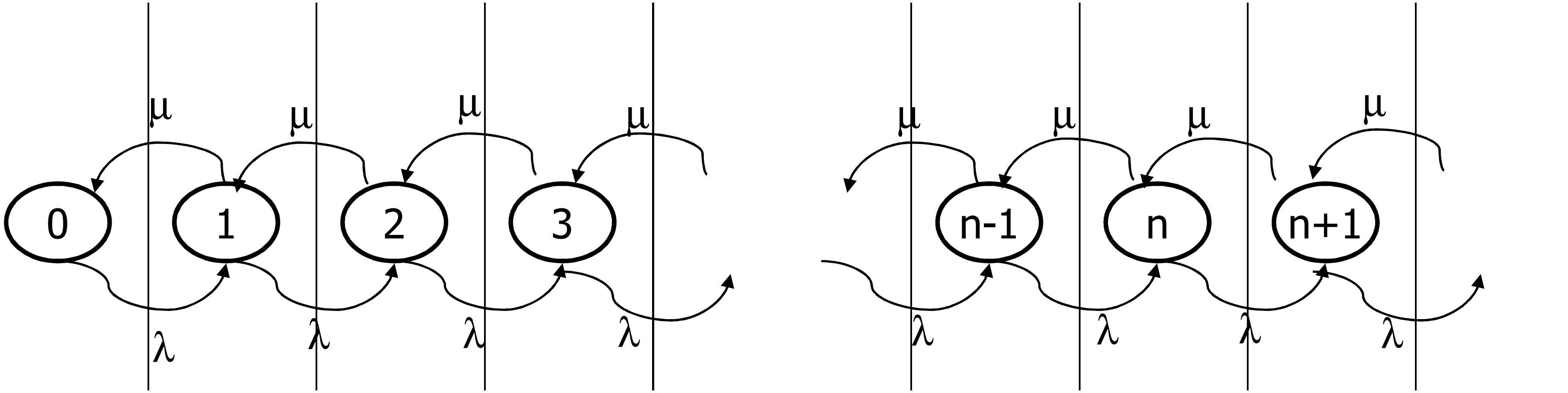
- **Kirchov's current conservation law**

  Flow in = Flow out

- $P_0 + P_1 + \ldots + P_N = 1$

Courtesy to Internet Information

區域狀態平衡方程式



- $\lambda P_0 = \mu\ P_1$
- $\lambda P_1 = \mu\ P_2$
- $\lambda P_2 = \mu\ P_3$
- .....
- $\lambda P_{n-1} = \mu\ P_n$

Courtesy to Internet Information

機率分配

- $P_1 = (\lambda/\mu)P_0$
- $P_2 = (\lambda/\mu)P_1$
- .....
- $P_n = (\lambda/\mu)P_{n-1}$

$$P_n = \left(\frac{\lambda}{\mu}\right)^n P_0$$

$$P_0 = \frac{1}{\sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n}$$

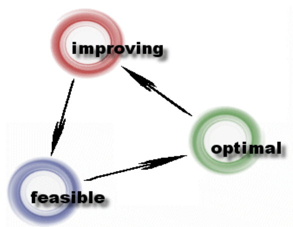
Courtesy to Internet Information

Little's Law

- The mean number of customer in a queueing system

  $L=\lambda W$ W is the mean waiting time

Courtesy to Internet Information



Example of Little Law

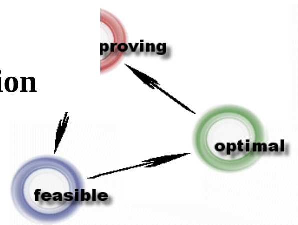
- The average waiting time for M/M/1 queue

$$W = \frac{1/\mu}{(1-\rho)}$$

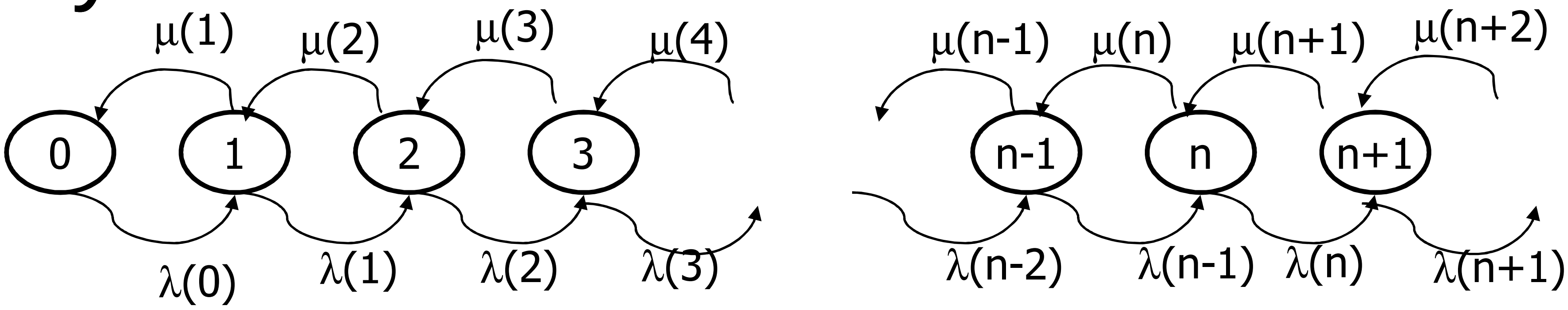
- Little Law also holds if we consider only the queue and not the server.

$$W_q = \frac{1/\mu}{(1-\rho)} - 1/\mu = \frac{\lambda}{(1-\rho)}$$

Courtesy to Internet Information



State dependent M/M/1 Queueing Systems



$$P_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \prod_{i=1}^{n} \frac{\lambda(i-1)}{\mu(i)}}$$

Courtesy to Internet Information

Performance Measure

- Mean Throughput of the queueing system
  - Mean rate at which customers pass through

$$\bar{Y} = \sum_{n=1}^{\infty} \mu(n) P_n$$

- Mean number of customers in the queueing system

$$L = \sum_{n=1}^{\infty} n P_n$$

- Mean time delay (using Little's Law)

$$W = \frac{L}{\bar{Y}} = \frac{\sum_{n=1}^{\infty} n P_n}{\sum_{n=1}^{\infty} \mu(n) P_n}$$

- Utilization

  $U=1-p_0$

Example 1

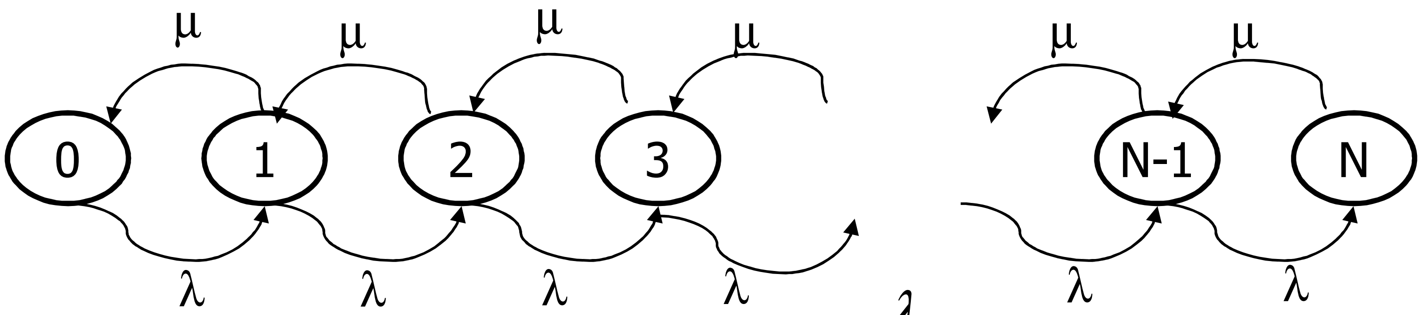
Suppose interarrival time is 12 min and service time is 10 minutes per job.

$$\lambda = \frac{1}{12}\ jobs/\min = 5\ \ jobs/hr \qquad \mu = \frac{1}{10}\ jobs/\min = 6\ \ jobs/hr$$

$$W = \frac{1/\mu}{(1-\rho)} = W_q + \frac{1}{\mu}$$

$$W_q = \frac{\lambda}{(1-\rho)} \qquad p_0 = 1-(\lambda/\mu)$$

M/M/1/N Queueing System
The Finite Buffer Case



$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0, 0 \le n \le N$$

$$p_0 = \frac{1}{\sum_{n=0}^{N} \left(\frac{\lambda}{\mu}\right)^n} = \frac{1 - \frac{\lambda}{\mu}}{1 - \left(\frac{\lambda}{\mu}\right)^{N+1}}$$

- Blocking probability

$$\therefore p_n = \frac{1 - \frac{\lambda}{\mu}}{1 - \left(\frac{\lambda}{\mu}\right)^{N+1}} \left(\frac{\lambda}{\mu}\right)^n$$

Courtesy to Internet Information

Example 2

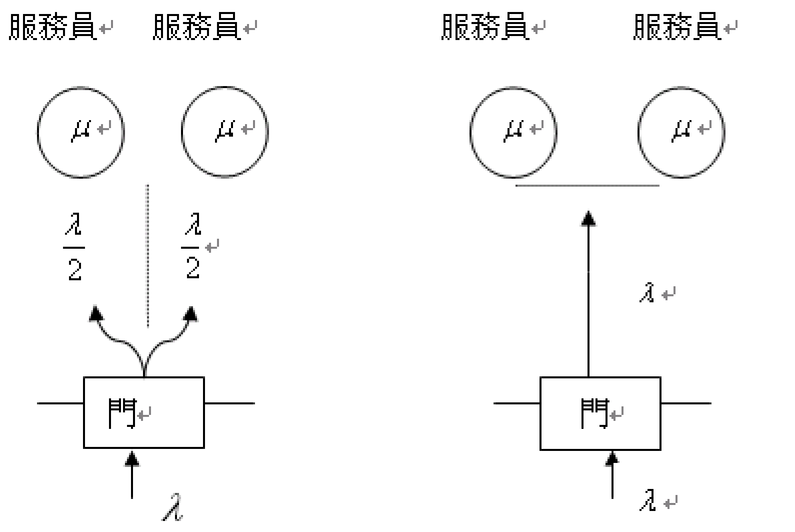
Suppose interarrival time is 12 min and service time is 10 minutes per job.

$$\lambda = \frac{1}{12}\ jobs/\min = 5\ \ jobs/hr \qquad \mu = \frac{1}{10}\ jobs/\min = 6\ \ jobs/hr$$

$$p_0 = \frac{1-\rho}{1+\rho} \qquad W_q = \frac{2\rho^2}{1-\rho^2}$$

$$\rho = \frac{\lambda}{2\mu}$$

比比看



Two M/M/1

One M/M/2

平均等候時間

M/M/1 $\times 2$

平均服務率 μ

平均到達率 $\dfrac{\lambda}{2}$

P{idle time} $P_0 = \dfrac{2\mu - \lambda}{2\mu}$

E(waiting $W_q = \dfrac{\lambda\mu}{(2\mu - \lambda)}$

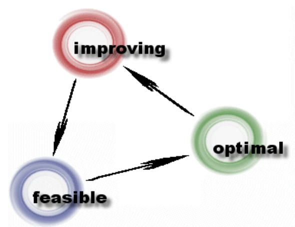
平均等候時間

M/M/2

平均服務率 μ　　　　平均到達率 λ

P{idle time}　　$p_0 = \dfrac{2\mu - \lambda}{2\mu + \lambda}$

E(waiting time}　　$W_q = \dfrac{\lambda^2}{\mu(2\mu + \lambda)(2\mu - \lambda)}$



Example 3

Suppose interarrival time is 12 min and service time is 10 minutes per job.

$$\lambda = \frac{1}{12}\ jobs/\min = 5\ \ jobs/hr \qquad \mu = \frac{1}{10}\ jobs/\min = 6\ \ jobs/hr$$

Two M/M/1:

$$W_q = \frac{\lambda\mu}{(2\mu-\lambda)} = \frac{1/120}{1/5 - 1/12} = \frac{1}{14} \approx 0.0714\,(\text{min}/\text{job})$$

$$L_q = \frac{\lambda}{2}\frac{\lambda\mu}{2\mu-\lambda} = \frac{1}{336} \approx 0.003\ (\text{jobs})$$

$\times 2$

One M/M/2:

$$W_q = \frac{\lambda^2}{\mu(2\mu+\lambda)(2\mu-\lambda)} = \frac{\left(\frac{1}{12}\right)^2}{\frac{1}{10}\left(2*\frac{1}{10}+\frac{1}{12}\right)\left(2*\frac{1}{10}-\frac{1}{12}\right)} = \frac{250}{119} \approx 2.10(\min/job)$$

$$L_q = \lambda * W_q = \frac{1}{12} * \frac{250}{119} = \frac{125}{714} \approx 0.175(jobs)$$

M/M/2與M/M/1 ×2 的比較

	M/M/2	v.s	M/M/1 ×2
E(wainting_time)in queue)	$\frac{\lambda^2}{\mu(2\mu+\lambda)(2\mu-\lambda)}$	$<$	$\frac{\lambda}{\mu(2\mu-\lambda)}$
P{idle_time}	$\frac{2\mu-\lambda}{2\mu+\lambda}$	$<$	$\frac{2\mu-\lambda}{2\mu}$
Var(waiting_time)	$\frac{\lambda^2(4\mu^2+2\lambda\mu-\lambda^2)}{\mu^2(2\mu+\lambda)^2(2\mu-\lambda)^2}$	$<$	$\frac{\lambda(4\mu-\lambda)}{\mu^2(2\mu-\lambda)^2}$
Var(system_time)	$\frac{16\mu^3-4\mu\lambda^2+2\lambda^3}{\mu(2\mu+\lambda)^2(2\mu-\lambda)^2}$	$<$	$\frac{4}{(2\mu-\lambda)^2}$

等候時間的機率分配函數

M/M/2 $\mu_n = \begin{cases} n\mu, 1 \le n \le 2 \\ 2\mu, n \ge 2 \end{cases}$, λ , $\gamma = \frac{\lambda}{\mu}$ $\rho = \frac{\gamma}{c} = \frac{\lambda}{2\mu}$

$$W_q(t) = W_q(0) + \sum_{n=c}^{\infty} P\{n-c+1 \text{ completions in } \le t \mid \text{ arrival found n in system}\} * p_n$$

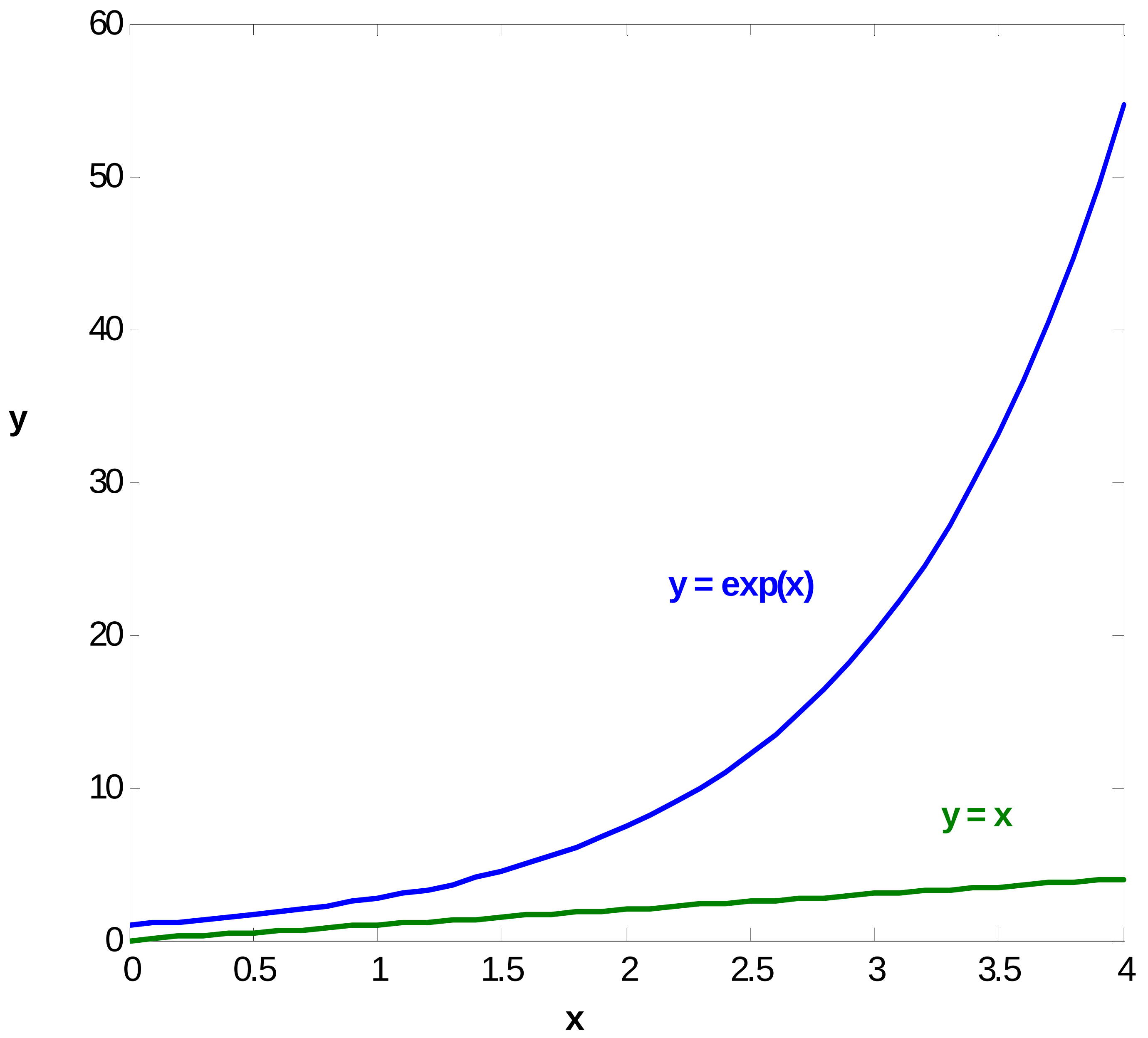
$$= W_q(0) + P_0 \sum_{n=c}^{\infty} \frac{\gamma^n}{c^{n-c} c!} \int_0^t \frac{c\mu(c\mu x)^{n-c}}{(n-c)!} e^{-c\mu x} dx$$

$$= W_q(0) + \frac{\gamma^c P_0}{(c-1)!} \int_0^t \mu e^{-c\mu x} \sum_{n=c}^{\infty} \frac{(\mu \gamma x)^{n-c}}{(n-c)!} dx$$

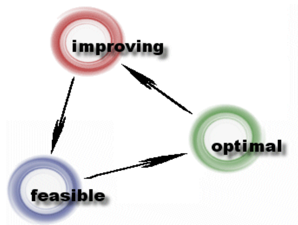
$$= W_q(0) + \frac{\gamma^c P_0}{(c-1)!} \int_0^t \mu e^{-\mu x(c-\gamma)} dx$$

$$= W_q(0) + \frac{\gamma^c P_0}{c!(1-\rho)} \left(1 - e^{-(c\mu-\lambda)t}\right)$$

$$= 1 - \frac{\gamma^c P_0}{c!(1-\rho)} e^{-(c\mu-\lambda)t}$$



60
50
40
30
20
10
0
y
y = exp(x)
y = x
0
0.5
1
1.5
2
2.5
3
3.5
4
x



improving
optimal
feasible

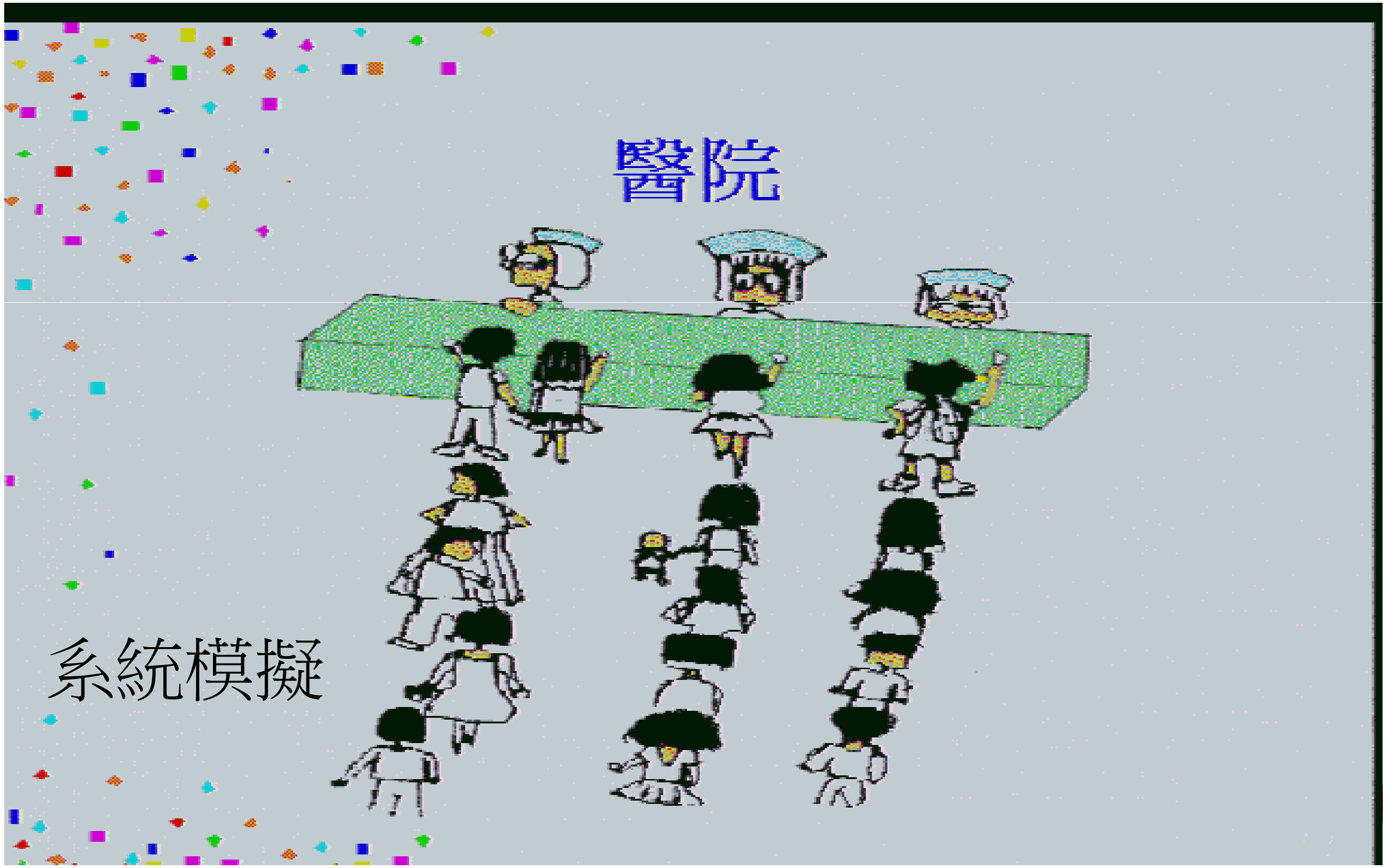
佇列的等候機率

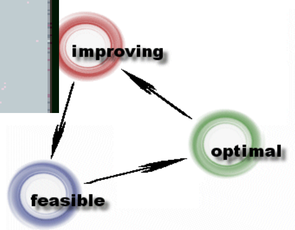
	M/M/2	v.s	M/M/1
P{waiting time > t min}	$\frac{\lambda^2}{\mu(2\mu+\lambda)}e^{-t(2\mu-\lambda)}$	$<$	$\frac{\lambda}{2\mu}e^{-t\left(\mu-\frac{\lambda}{2}\right)}$

$$\frac{\frac{\lambda}{2\mu}e^{-t\left(\mu-\frac{\lambda}{2}\right)}}{\frac{\lambda^2}{\mu(2\mu+\lambda)}e^{-t(2\mu-\lambda)}} = \left(\frac{\mu}{\lambda}+\frac{1}{2}\right)e^{t(\mu-\lambda/2)}$$

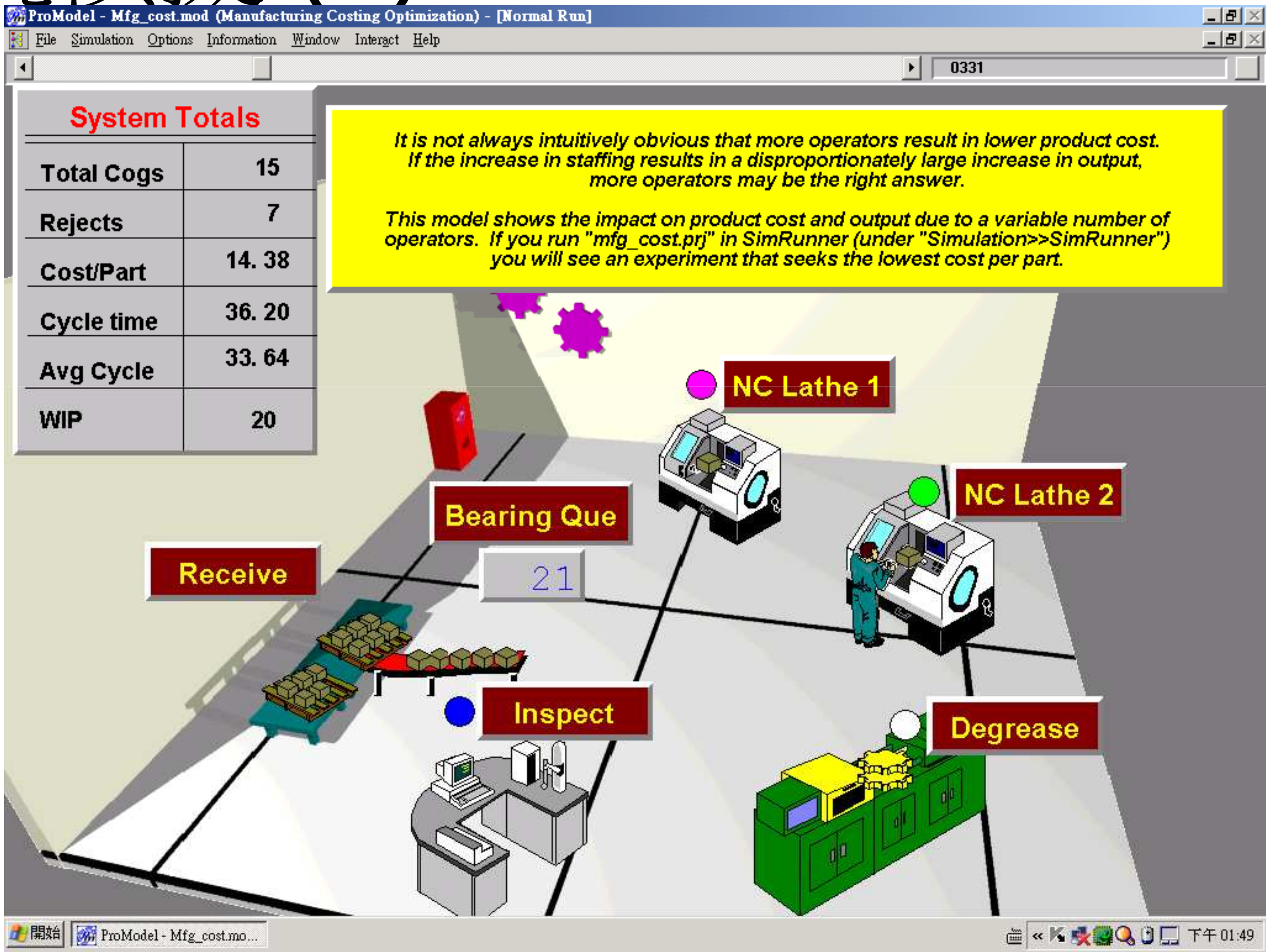
$$\lambda < 2\mu$$

時間延誤問題

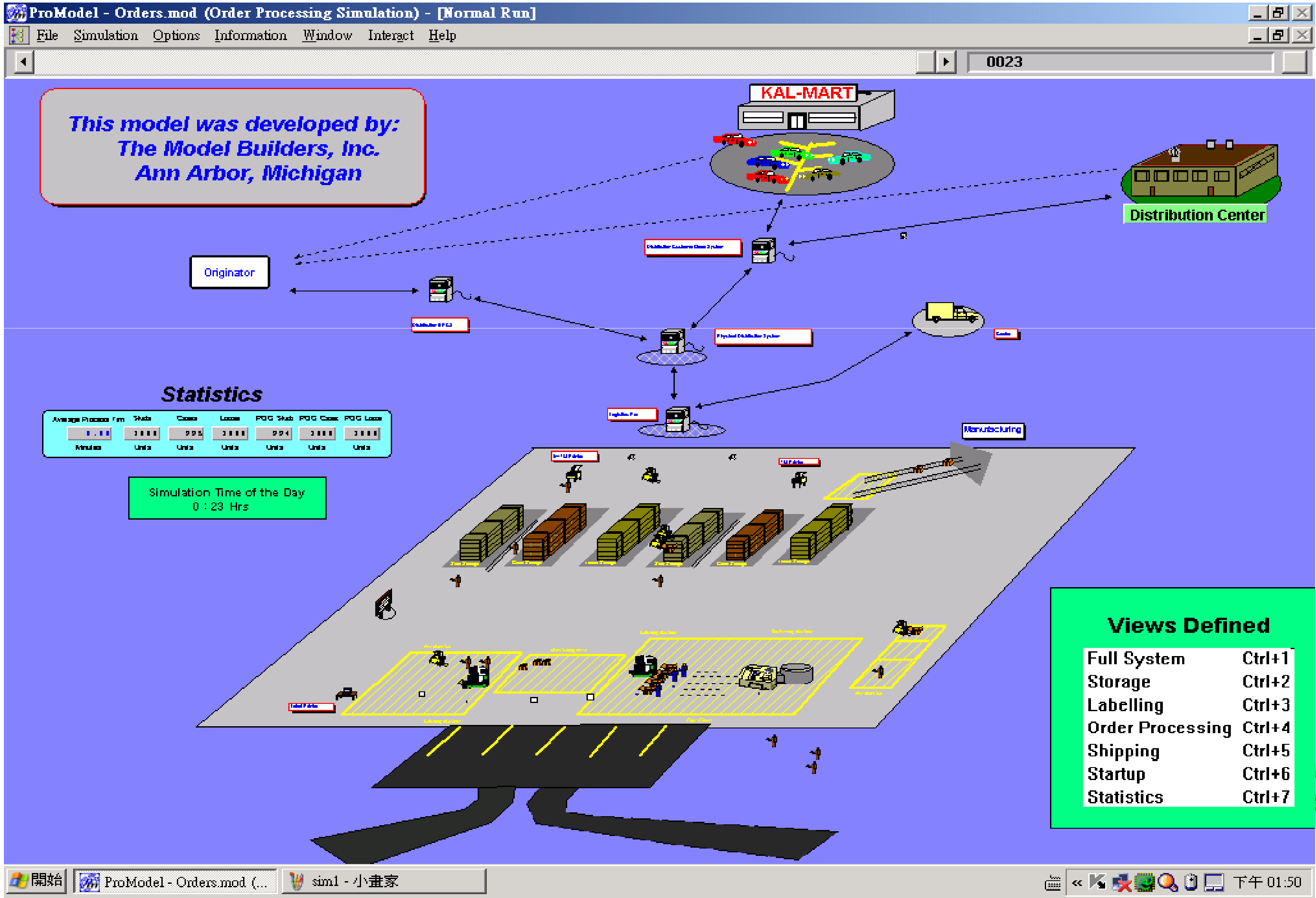




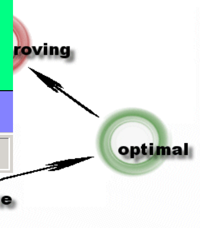
系統模擬 (1)


ProModel - Mfg_cost.mod (Manufacturing Costing Optimization) - [Normal Run]
File Simulation Options Information Window Interact Help
0331
System Totals
Total Cogs 15
Rejects 7
Cost/Part 14. 38
Cycle time 36. 20
Avg Cycle 33. 64
WIP 20
It is not always intuitively obvious that more operators result in lower product cost. If the increase in staffing results in a disproportionately large increase in output, more operators may be the right answer.
This model shows the impact on product cost and output due to a variable number of operators. If you run "mfg_cost.prj" in SimRunner (under "Simulation>>SimRunner") you will see an experiment that seeks the lowest cost per part.
NC Lathe 1
NC Lathe 2
Bearing Que
21
Receive
Inspect
Degrease
開始
ProModel - Mfg_cost.mo...
下午 01:49

系統模擬 (2)

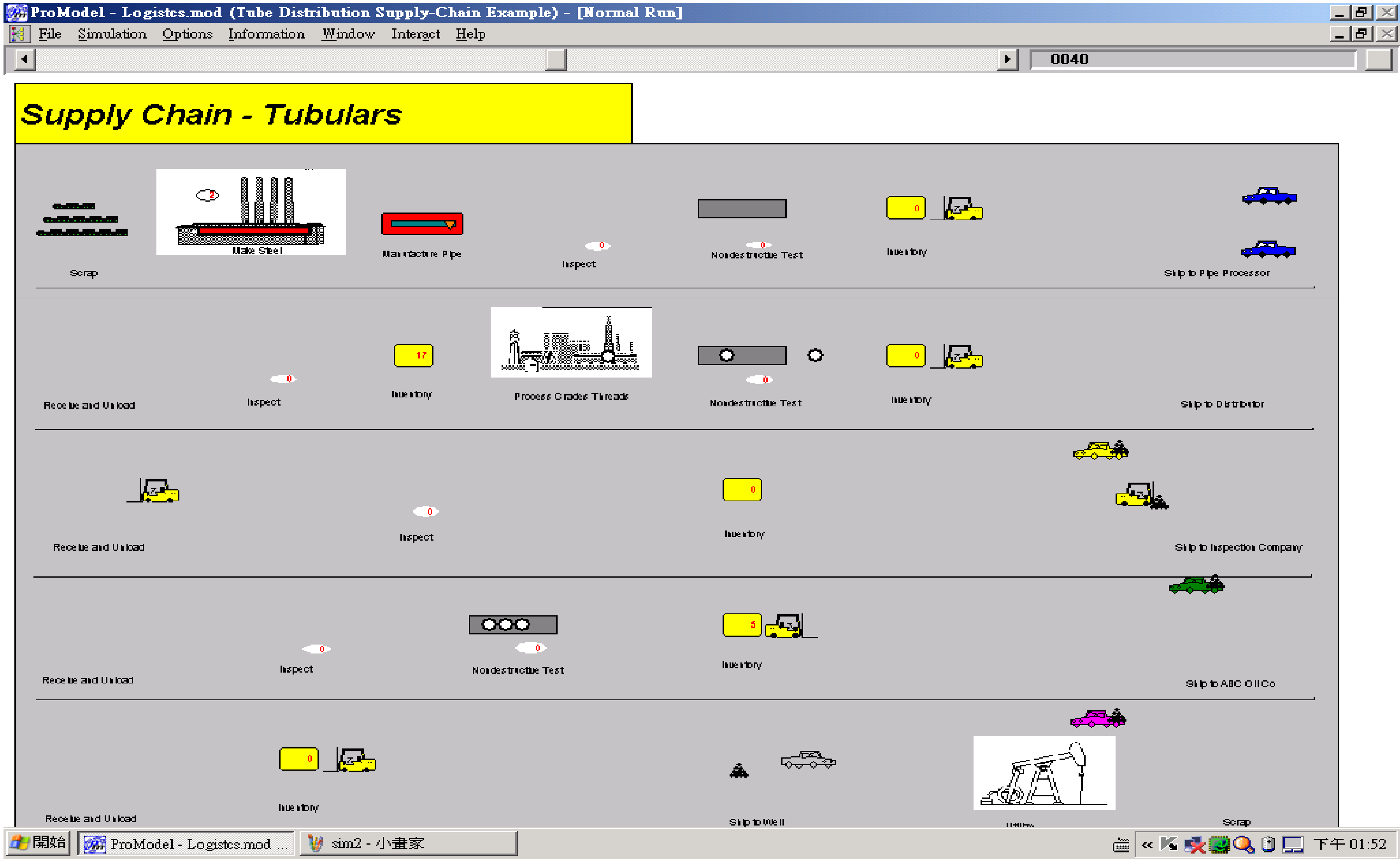


ProModel - Orders.mod (Order Processing Simulation) - [Normal Run]
File Simulation Options Information Window Interact Help
0023
This model was developed by:
The Model Builders, Inc.
Ann Arbor, Michigan
KAL-MART
Distribution Center
Originator
Manufacturing
Statistics
Simulation Time of the Day
0 : 23 Hrs
Views Defined
Full System Ctrl+1
Storage Ctrl+2
Labelling Ctrl+3
Order Processing Ctrl+4
Shipping Ctrl+5
Startup Ctrl+6
Statistics Ctrl+7
開始
ProModel - Orders.mod (...
sim1 - 小畫家
下午 01:50

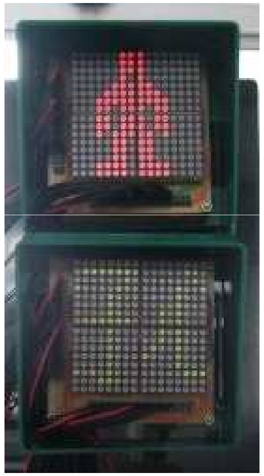


roving
optimal
feasible

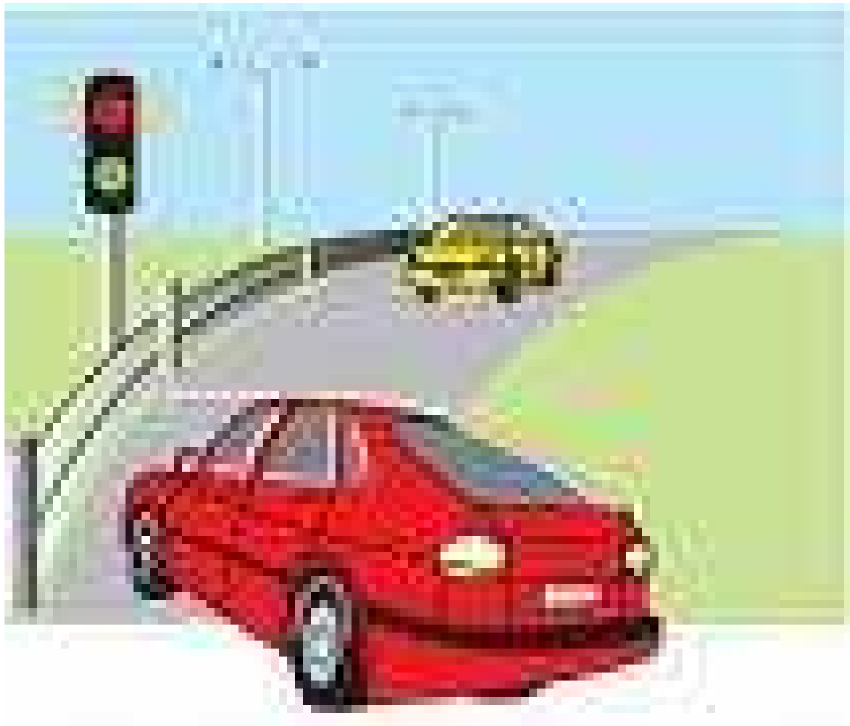
系統模擬 (3)

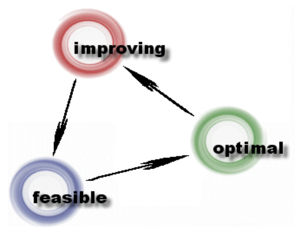

ProModel - Logistcs.mod (Tube Distribution Supply-Chain Example) - [Normal Run]
File Simulation Options Information Window Interact Help
0040
Supply Chain - Tubulars
Scrap
Make Steel
Manufacture Pipe
Inspect
Nondestructive Test
Inventory
Ship to Pipe Processor
Receive and Unload
Process Grades Threads
Ship to Distributor
Ship to Inspection Company
Ship to ABC Oil Co
Ship to Well
Scrap
開始
ProModel - Logistcs.mod ...
sim2 - 小畫家
下午 01:52

交通號誌

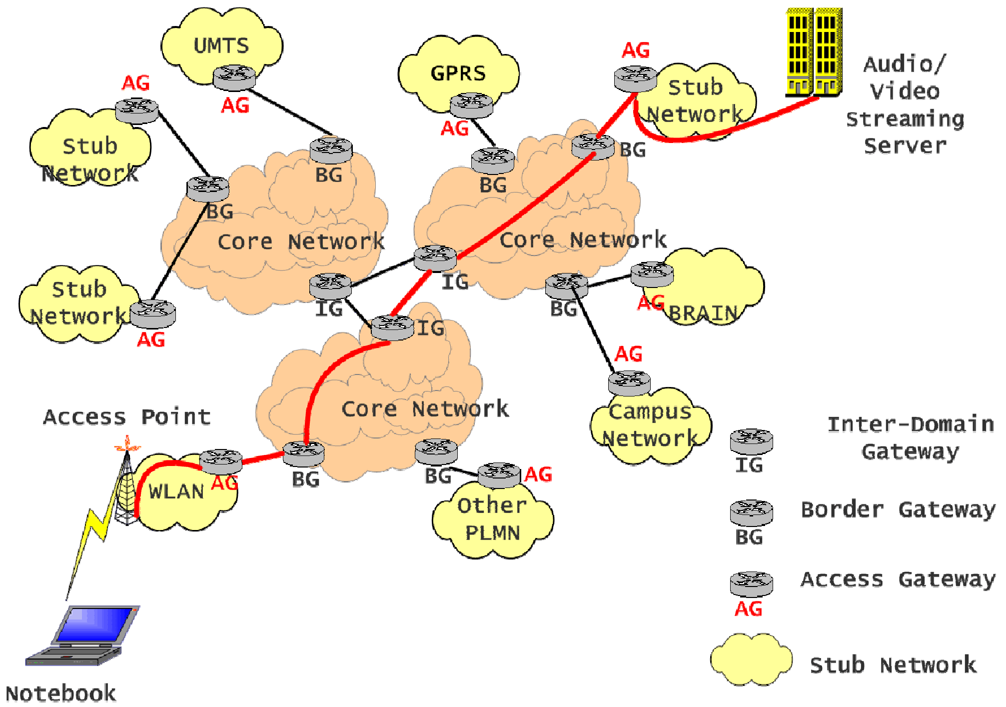


匝道管制

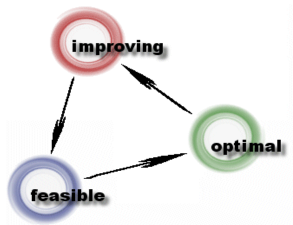




通訊服務



UMTS
GPRS
AG
BG
IG
Stub Network
Core Network
Audio/ Video Streaming Server
BRAIN
Campus Network
Access Point
WLAN
Other PLMN
Notebook
Inter-Domain Gateway
Border Gateway
Access Gateway
Stub Network



improving
optimal
feasible

最佳化模型

An optimization model

Minimize

Penalty made by unsatisfied traffic aggregates

Subject to

Fixed resource (Limited Short Path bandwidth)

Traffic aggregate QDF budget must be satisfied

Allocate resource to traffic aggregates

Trading off QDF with bandwidth in each Short Path

謝　謝　大　家